J1.1

Analysis of Large Scale Spatial Variability of Soil Moisture Data Using Geostatistical Method

Tarendra Y. Lakhankar *, Andrew S. Jones, Cynthia L. Combs, Manajit Sengupta and Thomas H. Vonder Haar

Cooperative Institute for Research in the Atmosphere (CIRA), Colorado State University, Fort Collins, CO 80523

ABSTRACT

In this study, a variogram is used to analyze the spatial structure of the Air Force Weather Agency's (AFWA) Agricultural Meteorology (AGRMET) soil moisture model output and in-situ Oklahoma Mesonet soil moisture data. The spatial variability information is then used by a Kriging method to estimate soil moisture at unsampled locations. The spatial decorrelation length scale of soil moisture is critical for the initialization in 4-dimensional variational (4DVAR) data assimilation research. The decorrelation length of soil moisture is seen to vary according to precipitation. Pre-precipitation regimes have a higher length than post-precipitation regimes indicating that precipitation storm-scales drive soil moisture spatial structures. The In-situ measurement systems used in this study were originally considered to be research-grade soil moisture networks, but were found to be susceptible to quality control issues. Under such conditions, techniques such as the Kriging method described in this study mitigate some of the quality control errors with appropriate geostatistical information. The effect of precipitation events on the spatial geostatistical structure (decorrelation length and sill) was observed.

Keyword: Soil moisture, Kriging, Variogram, AGRMET, Oklahoma Mesonet, Data assimilation

1 INTRODUCTION

Spatially and temporally varying soil moisture is being increasingly used as input to hydrological and meteorological models. In meteorology and climate studies, soil moisture directly affects the partitioning of energy at the surface between latent and sensible heating. Evaporation predominates at higher soil moisture levels adding to atmospheric moisture content. Weather prediction models therefore require extensive information about the interaction between atmosphere and land surface processes. Several research groups are examining new methods, including satellite remote sensing and data assimilation, to provide this key information to the forecast models (Jones et al. 2004; Reichle et al. 2007). In addition, global climate change can be monitored through broader knowledge of accurate soil moisture content (Engman; Chauhan 1995)

Corresponding author address: Tarendra Y Lakhankar, Cooperative Institute for Research in the Atmosphere (CIRA), Colorado State University, Fort Collins, CO 80523
Email: tarendra@cira.colostate.edu

The variogram structure consists of the nugget (the variance at zero lag distance), sill (the variance to which the variogram asymptotically rises), and decorrelation length (range of spatial dependence). The decorrelation length varies based on minimum distance between sampling locations and size of sampled area (Western; Bloschl 1999). In this study the average distance between adjacent in-situ (Oklahoma Mesonet network) soil moisture sites is 51 km and is comparable to the grid resolution (47 km) of AGRMET soil moisture data as well as the spatial resolution of the WindSat satellite microwave data, which is the data source of the 4DVAR methodology related to this study. This distance is also approximately equal to precipitation storm-scales, which drive the soil moisture spatial structures (Hoff 2001), and therefore demonstrates that WindSat and other satellite-based microwave soil moisture sensors are useful even at these relatively crude resolutions.

The geostatistical studies for soil moisture variability (Anctil et al. 2002; Bardossy; Lehmann 1998; Herbst; Diekkruger 2003; Wang et al. 2001; Western; Bloschl 1999; Western et al. 1999) are carried out at the scales of small catchments areas (1-5 km^2). Thus, the areal extent of these studies is too small for robust soil moisture analysis at precipitation scales as well as spatial scales of soil moisture retrieval from passive microwave satellite data. Also a rigorous quality control tool is developed in this study to eliminate non-responsive sensors. This quality control is based on the autocorrelation relationship between precipitation and soil moisture change. Additionally AGRMET model output is compared against Oklahoma Mesonet in-situ soil moisture data using geostatistical methods for a period of 30 days during September 2003.

This research is part of the overall goal (Fig. 1) of development and validation of 4DVAR systems to bring new data into the forecast models. The 4DVAR system will be used to retrieve soil moisture using WindSat and future National Polar-orbiting Operational Environmental Satellite System (NPOESS) satellite data. There are 3 primary components to this study: 1) a subjective intercomparison between modeled (AGRMET) and Oklahoma Mesonet soil moisture data sources, 2) estimation of the decorrelation length before and after a precipitation event, and 3) an objective spatial intercomparison analysis that employs a Kriging geostatistical technique.

2 DATA

For the study, the time period of September 1-30, 2003 was chosen. A strong front with associated precipitation crossed the Midwest during the month, allowing observation of soil moisture both before and after a heavy rain event, including observations during the drying period. Many locations experienced at least two significant rain events.

2.1 AGRMET Model

The AGRMET soil moisture model output was used in this study (AWFA 2002). AGRMET is a near real-time global land surface analysis model at 47 km resolution. One of its unique features is that it produces a 3 hourly Special Sensor Microwave Imager (SSM/I) rain estimate as one of several sources of estimated precipitation. One of the products is soil moisture at four soil layer depths: 0-10 cm, 10-40 cm, 40-100 cm and 100-200 cm. In order to compare model output with in-situ data, a 110 km by 63 km grid box was centered over each site and averaged. This is shown in section 3.3. The soil moisture data from the AGRMET model is also used for the Kriging analysis of section 4.

2.2 In situ Oklahoma Mesonet

In-situ measurements of soil moisture was collected from the Oklahoma Mesonet (Brock et al. 1995). This state-wide monitoring network was originally set up by Oklahoma State University agricultural scientists to expand the use of weather data in agricultural applications and the needs of University of Oklahoma scientists to plan and implement a flood warning system in Tulsa. Currently, the network has at least one site in each county, measuring a variety of weather parameters. Since these datasets were not collected with satellite or model calibration in mind, we must understand the strengths and limitations of each dataset in order to properly compare in-situ measurements with satellite-based or model-based results.

Soil moisture is measured at each site at depths of 5 cm, 25 cm, 60 cm and 70 cm. The documentation mentions data quality issues at the two lower depths, so we concentrated on the top two depths. There are also local issues with respect to soil texture and vegetation. A site having more clay tends to have lower soil moisture change (thus lower infiltration rates) compared to a site which is sandier. Also, clay sites typically generate more runoff than a site with higher sand content. Similarly, areas with more vegetation hold more water within their roots than do bare soil. Due to all these variations, we are primarily concerned with signal response than with absolute calibration or 'truth'.

2.3 Oklahoma Mesonet Data Screening and Quality Control

Because of its extensive network the Oklahoma Mesonet has the number of stations and spatial extent for meaningful use of the Kriging technique. While analyzing the Oklahoma Mesonet soil moisture data, we observed that some of the data at some of the sites were either unrealistic or had little soil moisture change throughout the month of study. Some of the sensors did not respond to high precipitation events that occurred during the study period. This could be attributed to saturated soil, sensors malfunctioning underground or calibration issues. In this study, it is a prerequisite that such sites (sensors) be removed from further spatial analysis. Therefore, we developed a response function to eliminate the bad soil moisture sensors based on threshold limits. The response function is based on the cross-correlation of change in soil moisture with respect to precipitation events occurring during the study period.

The locations of filtered sites after applying the response function to all Oklahoma Mesonet sites are shown in Fig. 2. For sake of simplicity in the geostatistical analysis, the four sites in panhandle area of Oklahoma State were not selected. The application of this response function has been limited by available soil moisture data (30 days) for each Oklahoma Mesonet location. The average number of precipitation events occurring at each site varies between 2 to 7. However, some of the precipitation events have less than 5 mm of total rainfall during events, which may not lead to changes in soil moisture measured at 5 cm below the surface. Hence such events were not considered when developing the response function. An improved evaluation of this response function would require additional months of soil moisture data.

Mean soil moisture values are higher after precipitation events with higher variance being observed during wet periods after precipitation (Fig. 3). This could be due to spatially varying soil hydraulic properties creating differential infiltration rates during wet periods following rainfall, causing larger variation in soil moisture. Smaller variation is observed during dry periods where soil-related variability becomes minimal (Reynolds 1970). The AGRMET model underestimate the soil moisture compared to in situ measurements. However, the differences between average soil moisture values are smaller after precipitation or in wet soil conditions between both datasets. The variance of soil moisture was increasing during precipitation events for Oklahoma Mesonet data compared to more stable variance in the AGRMET data. Based on the trend (Fig. 3), it can be concluded that higher drying rates have been considered in the AGRMET model.

3 GEOSTATISTICAL ANALYSIS OF SOIL MOISTURE DATA

After performing an initial quality assessment of the various in situ data sets and AGRMET output, a more detailed geostatistical analysis was performed to better quantify important spatial statistics, as well as to provide a more objective performance assessment that can be used directly within the 4DVAR analysis as constraints on the background information provided by the AGRMET output. The AGRMET and in-situ data available in latitude and longitude coordinates were converted to distance in km from the reference coordinates set at 100° W longitude and 33.5° N latitude.

The experimental variogram characterizes the spatial variability in the measured data. This variogram is used in Kriging to determine soil moisture values at unsampled locations. This section first describes the effect of selection of active separation distance for best model fitting. The second part compares the variogram and its elements for AGRMET and Oklahoma Mesonet soil moisture data. The last subsection covers the performance of the Kriging approach in determining soil moisture at unsampled locations.

One of the major issues in variographic analysis is the selection of total lag distance for variogram fitting to experimental data. As the separation distance increases, after half of the total separation distance, the variogram starts to decompose at larger separation distances due to the reduced availability of pairs. Thus to obtain robust estimation of the variogram, we ignored pairs at larger separation distances that usually have smaller variance. The separation distance is selected based on the criterion that 95% pairs should have been used for variogram model fitting. The effect of removing data pairs at larger separation distance significantly improves variogram model fitting to the Oklahoma Mesonet and AGRMET soil moisture data.

3.1 Variogram analysis of soil moisture data

The fitting of the appropriate model to the experimental variogram data is the most important step in geostatistical analysis. The fitting of the model can be done by personal judgment, or an automatic procedure can be followed to reduce subjectivity and to increase reproducibility. Different models can be fitted to the experimental semi-variance values. The most commonly used models, linear, spherical, exponential and Gaussian, were chosen to fit the experimental semi-variance plot, generated from soil moisture data, using least squares curve fitting. The elements of each variogram model and the regression coefficient R^2 of the fitting procedure were determined. The model with the higher value of R^2 was selected as an appropriate model to represent the sample semi-variogram. Based on the data, the Gaussian variogram model was the best fit for the AGRMET soil moisture data while the spherical model was better suited to the Oklahoma Mesonet in-situ soil moisture data. This is due to the smoothing may have been already occurred in AGRMET data due to its information source, i.e. Special Sensor Microwave/Imager (SSM/I), having a resolution of ~50 km. Most of the variograms fit with non-zero nuggets.

The Oklahoma Mesonet variogram contains an outlier with exceedingly high semi-variance resulting from a large difference in soil moisture between the two closest pairs of observation (PORT and HASK; PERK and STIL). Although these stations show appropriate variations in soil moisture with precipitation sensor calibration and soil type variation can have a significant influence on the semi-variance. Additionally the number of samples at this small separation distance is not statistically significant as there are only two data points available. Similar observations for nearby samples were also made by Hollingsworth and Lönnberg (1986).

The largest change in isotropic variogram properties for AGRMET and Oklahoma Mesonet soil moisture data was observed after heavy precipitation on day 254 compared to day 253. Variograms were generated for all days of September 2003 for AGRMET and Oklahoma Mesonet soil moisture data. The comparison of decorrelation length and average soil moisture shows the effect of precipitation on change in decorrelation length (A_0). The decorrelation length is higher for dry periods before precipitation and decreases with increasing soil moisture during and after precipitation events.

The time series comparison of decorrelation length estimated for AGRMET and Oklahoma Mesonet (Fig. 4), shows that decorrelation length is higher in the case of AGRMET soil moisture data. However, the decorrelation length matches at two instances, specifically after precipitation events. The larger decorrelation in the case of AGRMET compared to Oklahoma Mesonet data is due to AGRMET's spatial averaging versus the point sampling for the Oklahoma Mesonet data.

3.2 Kriging Performance Assessment

Kriging provides optimal interpolation of soil moisture at grid points in a spatial domain based on autocorrelation in the variograms. The theoretical variogram model (Gaussian, spherical, exponential, or linear) that best fits the experimental variogram of AGRMET and Oklahoma Mesonet data was selected for soil moisture mapping using the block Kriging technique (Webster; Oliver 2001).

A jack-knifing method was applied to evaluate the performance of the Kriged data at different locations of the Oklahoma Mesonet when compared with true soil moisture values. The jack-knifing method is a process where a small set of stations that have been selected for the comparison study are left out and not used to generate the variogram and Kriging soil moisture estimates. This method ensures unbiased validation of the Kriging estimates by examining and quantifying the errors associated with estimating soil moisture using the Kriging process. The estimated value at the selected Oklahoma Mesonet stations was obtained by creating a semi-variogram and Kriging estimated using information from the rest of the sites. This procedure provided measured and estimated values for each sample location, so that actual estimation errors could be computed and compared. To test the performance we selected 10 sites (15% of total sites) out of 74 sites randomly distributed across the Oklahoma Mesonet area (Table 1). Thus, a list of measured values and interpolated values was obtained for the set of stations, and the distribution of errors was analyzed. The measured values were compared with the interpolated values and the bias, and root mean square errors (RMSE) were calculated.

No trends in bias and RMSE, specific to the wet and dry periods were observed at the locations. During the month, mostly positive biases were observed at KING, MARE, MAYR, and MEDI. The negative bias observed at KETC, MIAM, MINC, NOWA, and OKEM. LAHO and OKMU sites were small. The average RMSE of 10 jackknifed sites was found to be 3.4% through September 2003 (Table 1). Larger RMSE (~5.5%) were observed at KETC, MEDI and MIAM; and lower RMSE (about 1-1.5%) were observed at LAHO, MAYR and OKMU Mesonet sites. The RMSE values could potentially be lowered through use of a co-kriging analysis (Webster; Oliver 2001) by including precipitation as an additional variable.

The average root mean square of the difference (RMSD) between kriged Oklahoma Mesonet and AGRMET soil moisture maps for the study area was 4.6% of soil moisture. Higher RMSE was observed during drying period which could be due to a higher drying rate in the AGRMET model. Bias is lower than RMSD, though the study period follows a similar trend to RMSD.

4 CONCLUSIONS

We have compared several independent in situ soil moisture measurements with the AFWA AGRMET model output for areas around the Oklahoma Mesonet site for a period of 30 days during September 2003. Results indicate a tendency for the AGRMET precipitation input estimates to bias the model soil moisture results. This can result in entire rain events being omitted or added to the AGRMET output. When the AGRMET precipitation estimate is more realistic, the AGRMET soil moisture estimate improves. In addition, some Oklahoma Mesonet sites performed better than others as compared to in situ precipitation measurements. The variance of precipitation in AGRMET is observed to be smaller than in situ precipitation measurements from the Oklahoma Mesonet. A response function for quality control of Oklahoma Mesonet data is used to eliminate the soil moisture measuring sites which did not respond well to precipitation. Results from our studies indicate that in situ soil moisture data are of various quality levels. Some data networks experienced > 30% sensor failure rates using our more detailed quality control analysis procedures. Remaining quality-controlled data sets indicated that precipitation inputs were the primary cause of discrepancies between the AGRMET model output and in situ soil moisture measurements. However, in some circumstances soil texture and possibly other AGRMET model parameters or inputs were suspected of causing inconsistent soil moisture output results. In addition, due to the spatial representation errors we do not expect perfect model versus in situ agreement, although application of downscaling methods may be able to partially mitigate these errors (Merlin et al. 2006). We expect that improved in situ sensor calibration and quality control methods would increase the reliably of the soil moisture measurements.

In the future, we intend to perform a more detailed spatial analysis to automate the detection of false signals within a dispersed soil moisture network. In addition, the quantification of co-variances will be used to advance satellite data assimilation experiments (Jones et al. 2007). In particular, the horizontal decorrelation lengths determine the sharing of information within the 4DVAR cost function which directly impacts the data assimilation system performance. Previously, without this information, educated guesses are typically employed for the horizontal scale lengths (Zupanski et al., 2002). This can result in two possible data assimilation behavior errors: 1) an under-estimate of the horizontal decorrelation length scale results in unrealistic decoupling of the data assimilation spatial effects, thus needlessly increasing the data assimilation system errors thereby requiring a stronger remote sensing signal strength to compensate for the errors, and 2) an over-estimate of the horizontal decorrelation length scale results in overly-smooth data assimilation output results resulting in lost high resolution soil moisture information data. Thus, accurate spatial correlation length scale information minimizes the loss of data assimilation method accuracy and data signal strength. The results contained in this study are fundamental to the performance and behaviors of future 4DVAR assimilation for soil moisture retrieval using WindSat and future NPOESS satellite data, and will also direct future research activities toward areas requiring additional improvements.

ACKNOWLEDGEMENTS

Dr. Lakhankar is supported by NOAA through a post-doctoral position at Cooperative Institute for Research in the Atmosphere, Colorado State University. Part of this research was supported by the DoD Center for Geosciences/Atmospheric Research at Colorado State University under Cooperative Agreements DAAD19-02-2-0005 and W911NF-06-2-0015 with the Army Research Laboratory. The authors also thank Dr. George Mason from the Army Corps of Engineers and Mr. Gary McWilliams and Dr. Michael Mungiole from the Army Research Laboratory for many helpful discussions. We thank Dr. Jeff Basara and the rest of the Oklahoma Climatological Survey for supplying the Oklahoma Mesonet data and algorithms necessary to derive the in situ volumetric water content information used in this research.

REFERENCES

Anctil, F., R. Mathieu, L.-E. Parent, A. A. Viau, M. Sbih, and M. Hessami, 2002: Geostatistics of near-surface moisture in bare cultivated organic soils. *Journal of Hydrology*, **260,** 30-37.

AWFA, 2002: Data format handbook for AGRMET. http://www.mmm.ucar.edu/mm5/documents/DATA_FORMAT_HANDBOOK.pdf.

Bardossy, A., and W. Lehmann, 1998: Spatial distribution of soil moisture in a small catchment. Part 1: geostatistical analysis. *Journal of Hydrology*, **206,** 1-15.

Brock, F. V., K. C. Crawford, R. L. Elliot, G. W. Cuperus, S. J. Stadler, H. L. Johnson, and M. D. Eilts, 1995: The Oklahoma Mesonet: A Technical Overview. *Journal of Atmospheric Ocean Technology*, **12,** 5-19.

Engman, E. T., and N. Chauhan, 1995: Status of microwave soil moisture measurements with remote sensing. *Remote Sensing of Environment*, **51,** 189-198.

Herbst, M., and B. Diekkruger, 2003: Modelling the spatial variability of soil moisture in a micro-scale catchment and comparison with field data using geostatistics. *Physics and Chemistry of the Earth, Parts A/B/C*, **28,** 239-245.

Hoff, C., 2001: Spatial and temporal persistence of mean monthly temperature on two GCM grid cells. *International Journal of Climatology*, **21,** 731-744.

Hollingsworth, A., and P. Lönnberg, 1986: The statistical structure of short-range forecast errors as determined from radiosonde data. Part I: The wind field. *Tellus*, **38A,** 111-136.

Jones, A. S., T. Vukicevic, and T. H. Vonder-Haar, 2004: A microwave satellite observational operator for variational data assimilation of soil moisture. *Journal of Hydrometeorology*, **5,** 213-229.

Jones, A. S., and Coauthors, 2007: NPOESS soil moisture satellite data assimilation research using WindSat data. *Third Symposium on Future National Operational Environmental Satellite Systems—Strengthening Our Understanding of Weather and Climate*.

Merlin, O., A. Chehbouni, Y. H. Kerr, and D. C. Goodrich, 2006: A downscaling method for distributing surface soil moisture within a microwave pixel: Application to the Monsoon '90 data. *Remote Sensing of Environment*, **101,** 379-389.

Reichle, R. H., R. D. Koster, P. Liu, S. P. Mahanama, E. G. Njoku, and M. Owe, 2007: Comparison and assimilation of global soil moisture retrievals from the Advanced Microwave Scanning Radiometer for the Earth Observing System (AMSR-E) and the Scanning Multichannel Microwave Radiometer (SMMR). *Journal of Geophysical Research*, **112,** D09108, doi:09110.01029/02006JD008033.

Reynolds, S. G., 1970: The gravimetric method of soil moisture determination, part III, An examination of factors influencing soil moisture variability. *Journal of Hydrology*, **11,** 288–300.

Wang, J., B. Fu, Y. Qiu, L. Chen, and Z. Wang, 2001: Geostatistical analysis of soil moisture variability on Da Nangou catchment of the loess plateau, China. *Environmental Geology*, **41,** 113-120.

Webster, R., and M. A. Oliver, 2001: *Geostatistics for Environmental Scientists.* John Wiley & Sons Ltd, Chichester.

Western, A. W., and G. Bloschl, 1999: On the spatial scaling of soil moisture. *Journal of Hydrology*, **217,** 203-224.

Western, A. W., R. B. Grayson, and T. R. Green, 1999: The Tarrawarra Project: High resolution spatial measurement, modelling and analysis of hydrological response. *Hydrological Processes*, **13,** 633-652.

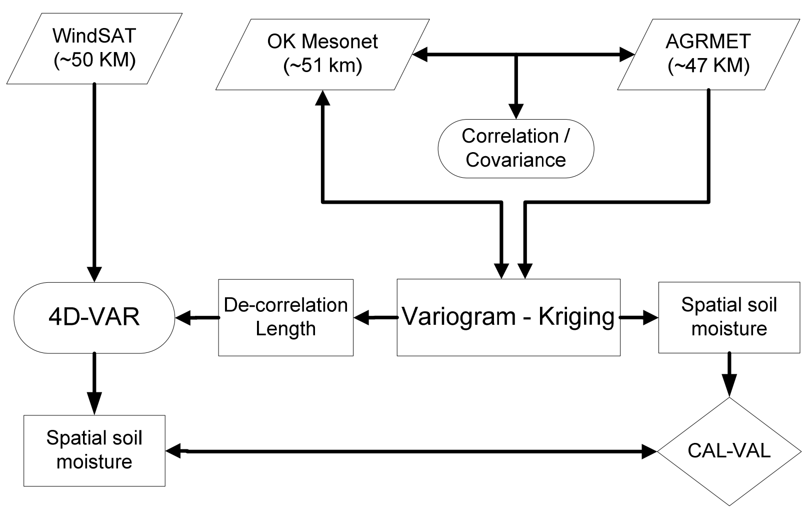

Fig. 1: This flow chart shows the application of variogram and Kriging analysis in calibration and validation of soil moisture information for data assimilation.

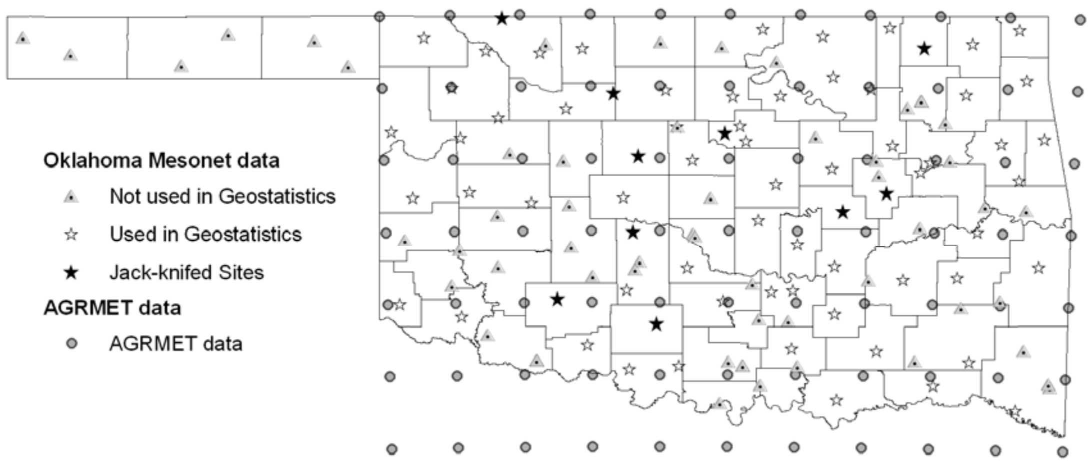

Fig. 2: The distribution of AGRMET grid points and Oklahoma Mesonet sites used in the geostatistical analysis. The data is described in detail in section 3.4.

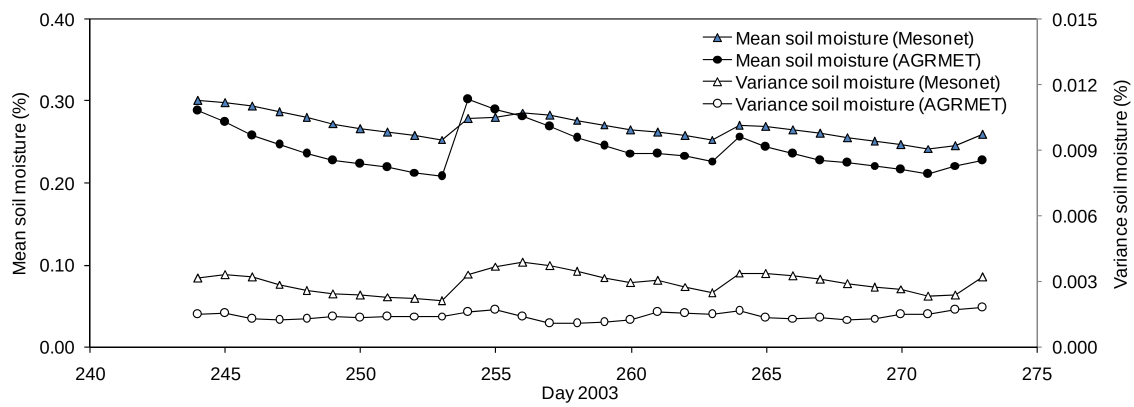

Fig. 3: The mean and variance of soil moisture measured at Oklahoma Mesonet sites and AGRMET data for study area shows peaks after precipitation events on day 244, 254, 264.

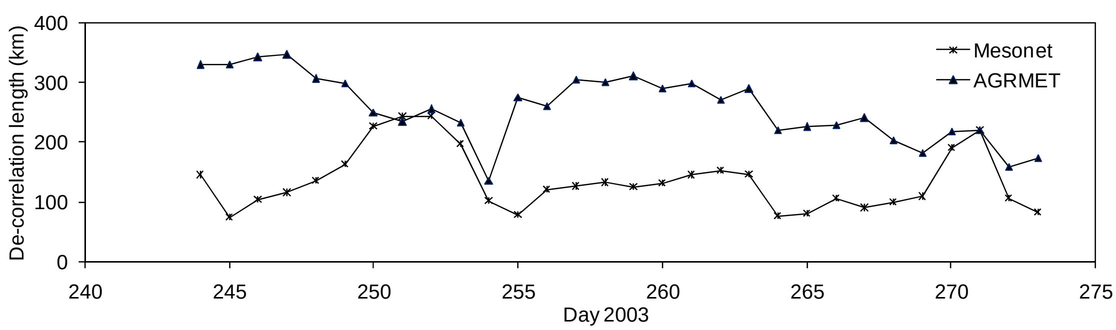

Fig. 4: Decorrelation lengths are higher for AGRMET compared Oklahoma Mesonet soil moisture data through study period.

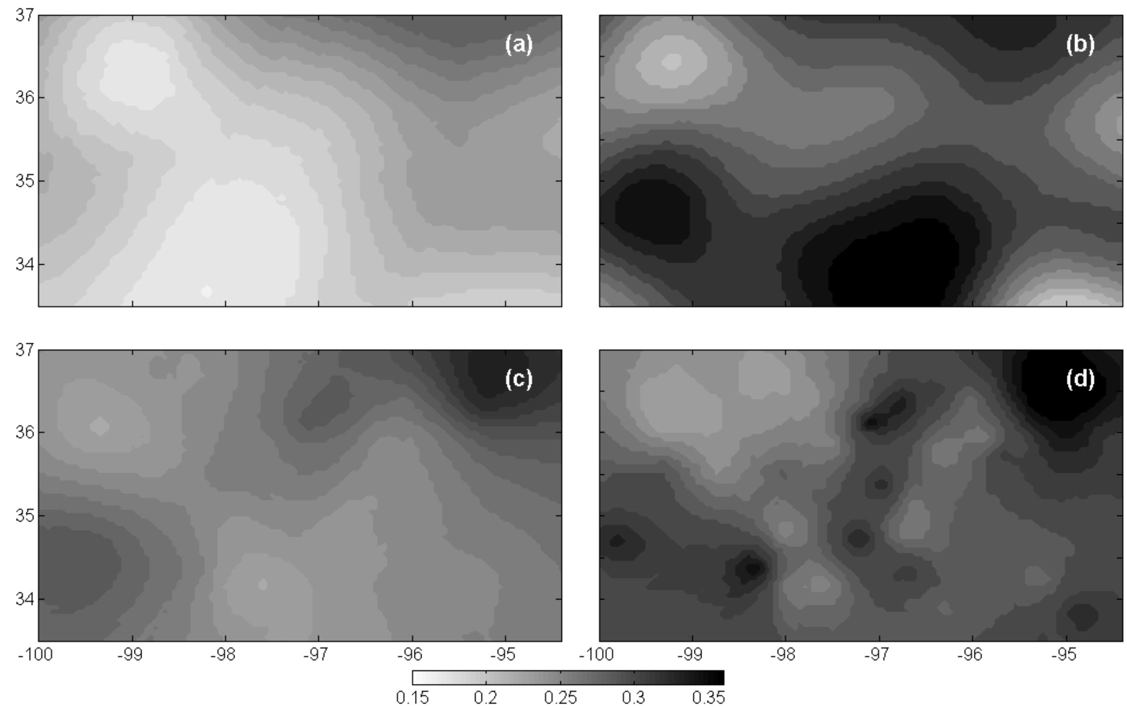

Fig. 5: Kriged map of soil moisture for AGRMET data (a and b) and Mesonet data (c and d) generated using semi-variograms. Figures (a) and (c) are before precipitation event (253 day), and (b) and (d) for after precipitation event (253 day).

Table 1: This table shows the performance of Kriging in terms of volumetric soil moisture at each jack-knifed Mesonet site for September 2003. The jack knifing procedure is outlined in Section 3.2.

Site Name	**Latitude**	**Longitude**	**Absolute Bias**	**RMSE**	**Correlation Coefficient**
KETC	34.529	-97.765	0.059	0.060	0.81
KING	35.881	-97.911	0.027	0.030	0.97
LAHO	36.384	-98.111	0.006	0.008	0.94
MARE	36.064	-97.213	0.047	0.048	0.86
MAYR	36.987	-99.011	0.012	0.013	0.80
MEDI	34.729	-98.567	0.060	0.061	0.94
MINC	35.272	-97.956	0.023	0.026	0.86
NOWA	36.744	-95.608	0.027	0.032	0.46
OKEM	35.432	-96.263	0.025	0.026	0.86
OKMU	35.581	-95.915	0.011	0.013	0.91
All Sites	**--**	**--**	**0.030**	**0.032**	**0.84**