ELSEVIER

Journal of Arid Environments 58 (2004) 357–372

Journal of Arid Environments

www.elsevier.com/locate/jnlabr/yjare

Geostatistical analysis of soil moisture variability in grassland

Qi Feng[a,*], Yansui Liu[b], Masao Mikami[c]

[a] *Cold and Arid Regions Environmental and Engineering Research Institute, Chinese Academy of Sciences, No. 260 West Donggang Road, Lanzhou 730000, People's Republic of China*

[b] *Institute of Geographic Sciences and Natural Resources Research, CAS, Beijing 100101, China*

[c] *Department of Environment and Applied Meteorology Research, Meteorological Research Institute, Nagamine 1-1, Tsukuba Ibaraki 305, Japan*

Received 27 August 2002; accepted 20 August 2003

Abstract

This study sought to identify patterns of spatial variations in volumetric soil moisture content (θ) of a portion of the Tsukuba grasslands of Japan. The θ of the topsoil (0–0.15 m profile) of a square, 100 ha plot of grassland was sampled on a roughly rectangular grid, at the rate one sample per hectare. Spatial maps of θ were developed using ordinary kriging methods. The use of a kriging estimator in concert with a Gaussian semivariogram, the measurement error model, and exact interpolation within the grassland were deemed appropriate for the data collected. Estimated autocorrelation values begin at roughly 0.577 and gradually decreased towards zero when the distance between points reached roughly 750 m. Negative but non-significant correlations was noted for inter-sampling point distances ranging from 750 to 800 m. As the distance further increased, the autocorrelation value became alternatively positive and negative, suggesting that the soil and ecosystem system exhibit periodic variations through natural influences. Based on idealized spherical variograms, θ values tended to be dependent for inter-sampling point distances of less than 750 m, but independent for distances beyond 800 m.

© 2003 Elsevier Ltd. All rights reserved.

Keywords: Soil moisture; Semivariogram; Sequential Gaussian simulation; Spatial structure; Tsukuba grassland

*Corresponding author. Tel.: +86-931-232-5690; fax: +86-931-8821894.
E-mail address: qifeng@ns.lzb.ac.cn (Q. Feng).

0140-1963/$ - see front matter © 2003 Elsevier Ltd. All rights reserved.
doi:10.1016/j.jaridenv.2003.08.002

1. Introduction

Soil moisture content (θ) variations under different climatic and geological conditions and over different spatio-temporal scales strongly influence solute transport (Dirmeyer et al., 1999). However, there are very few locations where θ is measured routinely, and those measurements are most often conducted on level agricultural fields, which are not representative of regional land surfaces (Vinnikov and Yeserkepova, 1991; Vinnikov et al., 1996). Thus, it is often convenient to model spatial processes stochastically, allowing models to accommodate the measurement uncertainty and explain the process under observation (US National Research Council, 1994). It is commonly accepted that measurements closely proximate in time or space are much more likely to be alike than widely separated ones (Singh and Fiorentino, 1996). It is upon this assumption that spatial data analysis methods have been developed. To be of use, the soil moisture data need to be grouped, manipulated, or scaled to bring out spatial dependence and independence, as well as similarities and differences common to areas, or layers of soil (Wang et al., 2001). One way of detecting such attributes is by their geostatistical correlation structure in space (Clarke and Dane, 1991). Quantitative estimates of this structure are required for a number of purposes including interpolation strategies from point data and estimation of the catchments averages (Goodchild, 1992).

Geostatistics can characterize and quantify spatial variability, perform rational interpolation, as well as estimate the variance of the interpolated values. Kriging, a geostatistical technique, is a linear interpolation procedure which provides a best linear unbiased estimator (BLUE) for quantities that vary in space. The procedure provides estimates at non-sampled sites. Recently, kriging has been applied in the environmental field to analyze spatial variability and resolve site-specific problems (Polhmann, 1993) Log-normal kriging was developed to account for the frequently skewed distribution of the data under investigation (Zimmerman and Zimmerman, 1991). This technique transforms the data into a log-normal formation prior to kriging estimation. Stochastic conditional simulation techniques, such as sequential Gaussian simulation, can be applied to generate multiple realizations, including an error component, absent from classical interpolation techniques. These simulation techniques generate a set of values with a specified mean and covariance, and also reproduce measured data at several locations. Along with simulated values, measurements therein can be used to analyze the spatial distribution of the variable in question. Furthermore, these techniques have been recently applied to characterize the spatial variability of soil moisture (Wang et al., 2001).

Many researchers have studied the spatial variability of topsoil moisture content, using geostatistical methods based on remotely sensed and field-measured data (Good, 1989). Some studies have reported little spatial correlation of θ (Loague and Green, 1991; Schmugge and Jackson, 1996; Mohanty et al., 2000), whereas Fitzjohn et al. (1998) suggested that such spatial correlation was evident in the Puebla de Valles-Retiendas region of central Spain. Thus, while no consensus exists on the spatial dependence of θ, differences in spatial scales from 6 to 3500 m suggest that θ

is a time-dependent property and that an evolving spatial dependence may exist (Davidson and Watson, 1995).

Therefore, in order to comprehensively assess the spatio-temporal structure of θ, further investigation is required in various locations over a large range of scales. Thus topsoil θ, usually very heterogeneous for grid block scales typical of mesoscale or climatic models, has not been adequately represented in such models. In recent years, several approaches have been proposed to characterize small-scale land surface heterogeneity in modelling the land–atmosphere system. This has been the central focus of a number of recent studies (Hu and Islam, 1997). One approach is to determine the effect of land surface heterogeneity on the surface feature by using so-called scale-invariant and surface parameterizations. Another approach is to increase the resolution of the model grid, dividing the model grid into smaller, presumably homogenous, subgrid elements and then estimate the surface feature at the subgrid scale.

In this paper, we present an analytical evaluation of the required subgrid scale for the θ of the topsoil of grasslands at the Meteorological Research Institute (MIR) in Tsukuba, Japan. Because of soil formation processes and climatic conditions specific to the region, grasslands are widespread in Japan. The main objectives were to (i) compare the spatial structure of layer-averaged θ, (ii) investigate the implications of different interpolation strategies, (iii) estimate the variability of θ over a large scale, and (iv) develop a wide-ranging database of θ measurements for the grasslands. Such a database could then serve to evaluate the impact of uncertainties in climatic model parameters through a direct comparison to field studies or arrays of θ measurements.

2. Materials and methods

2.1. Site description

The 1.0 km × 1.0 km area of grassland under study was situated at the Meteorological Research Institute (36°3′6″N–140°7′49″E), southwest of the city of Tsukuba, Japan. This 100 ha area was situated within a grassland plain roughly 3 km from east to west and 2.5 km from north to south, and was surrounded by line-trees roughly (Fig. 1). Non-grazed, unfertilized, and pesticide-free short natural grass covered nearly 80% of the research area, trees another 10%, and roads and offices the remainder. Trees varied from 5 to 10 m in height and had been planted across prevailing slopes or, in the northwest corner, up and down prevailing slopes for erosion controlling. The mean height of the grass varied from 0.5 to 1.0 m over the year. The study area is on a gentle downward slope (1/2000) from a central upper contour line dropping towards both the southeast and northwest (Fig. 1). The soils consist loamy sand (loamy sand, 75% clay, 5% sand) and loam (45% sand, 22% clay) and clay soils (65% clay and 10% sand). Annual precipitation is 1300 mm and the annual mean atmospheric humidity is 78%.

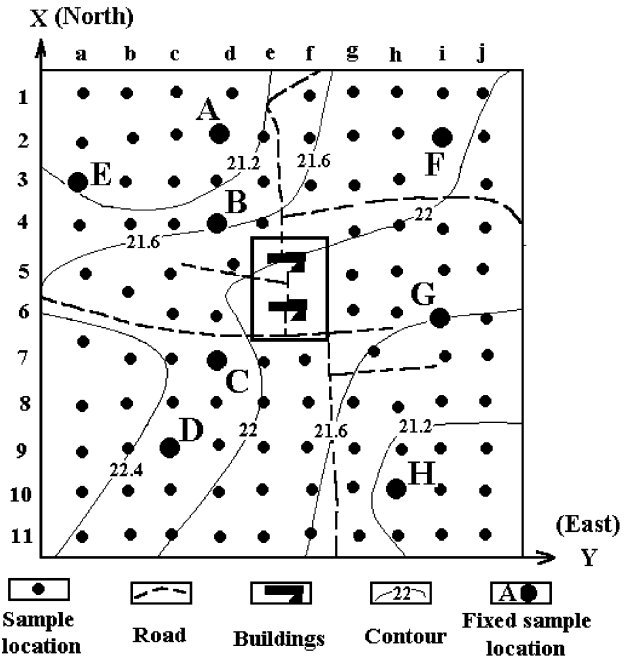

Fig. 1. Sketch of grid observation points in grassland in Tsukuba.

2.2. Soil parameter measurements

Measurements of gravimetric soil moisture content to a depth of 0.05 m were obtained for 100 locations distributed over the grassland, using the oven-drying method (Fig. 1). However, 8-fixed locations (point A, B, C, D, E, F, G and H) were arranged for continuous observation using TDR in the depth of 0.0–0.05, 0.05–0.10 and 0.10–0.20 through the year of 1999–2001. In October 1999, and January, June and October of 2000, volumetric soil moisture, θ (%), was measured at 100 locations in the study area, for soil profiles of 0–0.05 m. The mean θ data for each point across the four sampling dates without rain case during the sampling time were mapped for spatial analysis. The sampling method is described in greater detail in Feng et al. (2002).

The soils were sampled at depth increments of 0.05-m soil using a steel cylinder with a diameter and height of 0.05 m (volume $= 1 \times 10^{-4}$ m^3). Samples were dried for one day at 105°C, weighed and then heated for 4 h at 650°C, to measure the organic matter content. The percent weight difference between the dried and heated samples was taken as the organic matter content (Feng et al., 2002). Bulk density was determined using the known sample volume and the combined oven-dry weight of all sample material from a given site. Other soil property measurements were measured only once at permanent sampling points A, E, B, C, D, F, G, and H (Fig. 1) in depth of 0.0–0.05, 0.05–0.10 and 0.10–0.20 on April 8, 9, 10, 11, 12, 13, 14, and 15, 1999,

respectively. The detailed procedure is presented elsewhere (Bresler et al., 1969; Hillel, 1974, 1982). The results of all soil characteristics analyses are presented in Table 1.

The sampling standard deviations (across dates) of 100 sites of θ sampling from the 0–0.05 m surface profile of the MRI grassland are presented in Table 2, with sampling sites arranged as they are in Fig. 1. The development of a spatial model for predicting of θ at locations within the study area not already sampled in the study

Table 1
Properties of grassland soil in the top 0.20-m soil profile

Permanent sampling locations (Fig. 1)	Depth (m)	Soil characteristics[a]						
		ρ (Mg m^{-3})	Porosity (%)	Capillary height (m)	θ_{wp} (%)	θ_{fc} (%)	θ_{sat} (%)	Organic matter (g/kg)
A	0–0.05	1.55	35.46	0.75	1.34	2.65	54.83	11.2
	0.05–0.10	1.51	33.45	0.72	1.04	2.46	56.32	10.3
	0.10–0.20	1.47	29.03	0.75	1.43	2.02	58.51	8.9
E	0–0.05	1.54	28.57	0.73	1.35	2.61	52.43	5.6
	0.05–0.10	1.45	19.97	0.75	1.52	2.16	51.10	8.3
	0.10–0.20	1.53	27.99	0.73	0.88	2.62	53.31	7.6
B	0–0.05	1.59	38.91	0.78	0.57	2.19	44.00	10.5
	0.05–0.10	1.51	34.88	0.76	1.05	2.40	54.24	12.3
	0.10–0.20	1.54	35.73	0.67	0.61	2.63	53.80	10.2
C	0–0.05	1.65	33.85	0.85	1.17	2.79	53.16	8.6
	0.05–0.10	1.63	38.81	0.82	1.20	2.93	49.39	6.9
	0.10–0.20	1.62	29.25	0.85	1.23	2.78	59.44	7.3
D	0–0.05	1.69	37.60	0.83	1.25	2.88	48.84	10.2
	0.05–0.10	1.65	28.60	0.75	0.56	2.68	56.70	11.3
	0.10–0.20	1.67	35.39	0.78	0.58	2.80	50.42	10.5
F	0–0.05	1.57	38.99	0.72	0.89	2.81	51.73	5.6
	0.05–0.10	1.57	36.18	0.76	1.12	2.74	54.24	7.8
	0.10–0.20	1.57	41.15	0.77	1.02	2.85	49.66	8.3
G	0–0.05	1.68	37.65	0.85	1.03	2.78	50.58	10.5
	0.05–0.10	1.65	42.00	0.82	0.78	2.97	48.11	10.4
	0.10–0.20	1.63	38.05	0.85	0.77	2.81	51.73	8.4
H	0–0.05	1.61	38.46	0.74	0.89	2.70	50.84	11.0
	0.05–0.10	1.62	38.87	0.75	0.56	2.77	49.66	13.2
	0.10–0.20	1.61	38.70	0.73	0.45	2.80	49.57	9.7

[a] ρ, bulk density; θ_{wp}, percent volumetric soil moisture at permanent wilting point; θ_{fc}, percent volumetric soil moisture at field capacity; and θ_{sat}, percent volumetric soil moisture at saturation.

Table 2
Sampling standard deviation ($n = 4$; October 1999, January, June, October 2000) for 100 volumetric soil moisture content sampling sites (0–0.05 m depth) within the MRI grassland

North–south grid location	East–west grid location[a]									
	a	b	c	d	e	f	g	h	i	j
1	0.23	0.56	0.89	0.36	1.20	0.98	0.33	0.45	1.21	0.55
2	0.07	0.92	0.52	1.01	1.02	0.42	0.35	1.00	0.36	0.49
3	0.96	0.52	0.41	0.44	0.86	0.65	1.23	0.34	0.96	0.48
4	0.32	0.45	0.66	0.55	1.30	0.73	0.15	0.31	0.62	0.73
5	0.04	0.57	0.80	0.67	0.07	0.99	0.57	1.20	0.63	0.92
6	1.41	1.17	0.80	0.34	0.33	0.72	0.58	0.29	0.16	0.50
7	0.81	0.84	0.58	1.17	0.59	1.04	0.85	0.38	0.65	0.89
8	0.25	0.46	1.02	0.28	1.12	1.23	0.46	0.33	0.45	0.40
9	1.26	0.85	0.56	0.78	0.15	1.20	0.50	0.38	0.56	0.39
10	0.23	0.52	0.35	0.89	0.22	1.23	0.89	0.56	1.19	0.76

[a] a, b…j is shown in Fig. 1.

was the focus of further study. Rainfall was recorded by automatic datalogged rain gauges located at points C and F in Fig. 1.

2.3. Statistical analyses

In this study, the descriptive statistics of the frequency distributions, such as the mean, standard deviation and CV, were calculated using SAS software (SAS Institute Inc., 1992) and normality was assessed using the one-sample Kolmogorov–Smirnov (KS) test for goodness-of-fit. Semivariograms were constructed and fitted to curve types.

The simulation of the spatial variance in grassland θ, required the estimation of the number of samples required, N, given a specified probability, α,

$$N = t_\alpha^2 s^2 / D^2, \tag{1}$$

where D is the specified limit, expressing how close of an estimate is needed, s^2 are the sample variances, and t_α is the Student's t with $(N-1)$ degrees, at the α probability level (here 95%).

The sample covariance function was then:

$$\rho(h) = Cov[Z(x), Z(x+h)] / \{D[Z(x)]D[Z(x+h)]\}, \tag{2}$$

$$\rho(h) = Cov[Z(x), Z(x+h)] / \delta^2, \tag{3}$$

where Cov is the covariance for any two values of Z at a distance h apart, δ^2 is the variance of Z, h is the lag distance, $\rho(h)$ is the sample correlogram, and $Z(x)$, $Z(x+h)$ are the measured soil moisture contents at points x and $x+h$.

The semivariance, $\gamma(h)$, was estimated as

$$\gamma(h) = [\tfrac{1}{2}n(h)]\sum[Z(x_i) - Z(x_i + h)]^2 \quad [i = 1, 2, \ldots, n(h)], \tag{4}$$

$$\gamma(h) = \delta^2[1 - \rho(h)], \tag{5}$$

where $n(h)$ is the number of pairs of locations separated by a lag distance h.

The spherical model, $\gamma_1(h)$, was calculated as follows:

$$\gamma_1(h) = C_0 + C_1[1.5(h/a)^2 - 0.5(h/a)^3] \quad \text{for} \quad 0 < h < a, \tag{6}$$

$$\gamma_1(h) = C_0 + C_1 \quad \text{for} \quad h = a, \tag{7}$$

$$\gamma_1(h) = C_0 + C_1[1 - e^{(-h/a)}] \quad \text{for} \quad h > a, \tag{8}$$

where C_0, a, and $C_0 + C_1$ are the nugget, range, and sill in the idealized variogram. The nugget is a limiting value of $\gamma_1(h)$ for small h values; the sill is a maximum as h becomes larger; and the range is a distance for which $\gamma_1(h)$ has appreciably reached the sill.

The quality of the semivariogram fit to the data was indicated using a regression coefficient, R^2, and an F-test calculated as (Wang et al., 2001)

$$F = R^2/[1 - R^2] \times [n - k]/[k - 1], \tag{9}$$

where k is the number of variables in the regression model, and n is the number of samples.

Several functional models are available for idealized variogram analysis. For the three and two-dimensional simulation, the variogram model of Eqs. (6)–(8) is assumed to accurately depict spatial interdependence. The estimate Z_{est} at a non-sampled point (x_0, y_0) is then:

$$Z_{\text{est}}(x_0, y_0) = \sum\lambda_i Z(x_i, y_i) \quad \{i = 1, 2, \ldots, m\}, \tag{10}$$

where $Z_{\text{est}}(x_i, y_i)$ are the known values. $\lambda_1$andλ_m are weights chosen such that Z_{est} is unbiased.

Even though $Z(x_0, y_0)$ is unknown, the minimum variance can be evaluated in a manner consistent with the variogram function. The equation for evaluating λ is

$$\sum\lambda_j\gamma_{ij} + \mu = \gamma_{i0} \quad \{\mathrm{i} = 1, 2, \ldots, m\}, \tag{11}$$

where γ_{ij} is

$$\gamma_{ij} = \gamma[(x_i - x_j)^{0.5} + (y_i - y_j)^{0.5}], \tag{12}$$

i.e., the value of the variogram function evaluated between points (x_i, y_i) and (x_j, y_j) with i and j ranging from 1 to m. There are $m + 1$ unknowns consisting of the m weights λ_1, λ_2andλ_m, and the Lagrange multiplier μ. In addition to $Z_{i[\text{est}]}$, the variance of $Z_{\text{est}}(x_0, y_0) - Z(x_0, y_0)$ gives an idea of the quality of the estimate.

In comparing two series, one measured and one predicted, with differing variances, an efficiency factor, E, was selected, where the goodness-of-fit was based

on E approaching 1.0 (James and Burgess, 1982):

$$E = 1 - \frac{\sum_{i=1}^{i=n}(p_i - o_i)^2}{\sum_{i=1}^{i=n}(o_i - \bar{o})^2}, \tag{13}$$

where n is the total number of observations, o_i is the ith observed value, $\bar{o}$ is the mean of observed values, and p_i is the ith predicted value.

3. Results and discussions

3.1. Sample number estimates

Horizon boundaries may be more distinct than are the surface boundaries of a soil classification unit. Consequently, the sampling plan chosen is important, with the best design being one that provides the maximum precision for a given cost or vice versa. The principles of soil sampling were outlined by Cline (1944, 1989) and have not changed materially. One can use Eq. (1) to determine the number of observations required in future sampling of a population to estimate the mean within $\sim 5.00\%$, i.e. at the 95% probability level. The number of equally distributed θ samples to be taken over the Tsukuba grassland area were calculated based upon a mean θ of 39.23, and sample variance of 27.15 over the entire study area. Assuming confidence levels of 95%, 99% or 99.9% the $Z_{0.5\alpha}$ values would be 1.96, 2.326, or 3.091, respectively. If the estimate is to be within 5 [e.g. 39.23% $\sim$ 5.00%], the calculated number of samples, according to the above confidence levels would be 4.78, 6.74 and 11.9, respectively. These were rounded up to $n = 5, 7, 12$ and used as an estimated number of samples required to be within $5.0\,m^3$ (water) per $100\,m^3$ (dry soil) of the correct θ with a 95%, 99% and 99.9% confidence, respectively. In the study, we chose eight fixed locations for continuous soil water measurement, thus, our confidence in correctly assessing θ was slightly greater than 99%, with an estimate within $\sim 5\,m^3$ (water) m^{-3} (dry soil).

3.2. Autocorrelation analysis

The experimental mean and standard deviation of the 100-sample values, sampled in the order a1, a2, and a10; b1, b2, and b10; and j1, j2, and j10 (Table 2) were 39.23

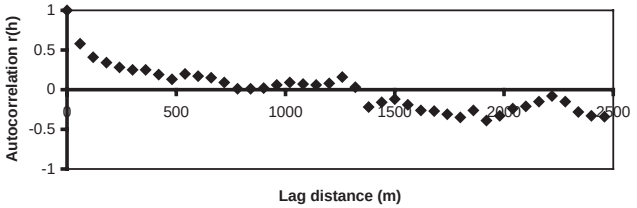

Fig. 2. Sample correlogram of the entire area.

and 4.6, respectively. Eq. (2) was used to calculate the correlogram (Fig. 2) or autocorrelation values $\rho(h)$. The value of 1.0 obtained at $h = 0$ was included for completeness, but for a small positive non-zero value of h, the $\rho(h)$ value was 0.577, gradually decreasing towards zero at about $h = 750$ m. For h values of 750–800 m, estimated autocorrelation values were slightly negative but not significant. As the distance h further increased, the correlogram became alternatively positive and negative. Physical systems exhibiting such a correlogram, include sediment deposits developed under periodic flooding or under anthropogenic influences. For h values of 0–750 m, the covariance decreased gradually until eventually, at the limit of 750 m the samples were alike and correlation or spatial dependence no longer existed. This result was obtained regardless of whether the analysis was being done from north to south or west to east.

3.3. Semivariance analysis

If strong stationary conditions are met, then both $n(h)$ and $\gamma(h)$ exist and Eq. (4) must be employed (Perrone and Madramootoo, 1997). Returning to the above 100 samples, we now calculate a sample variogram value for each lag distance, h (Fig. 3). For small separations, values of θ tend to be more alike and, consequently, $\gamma(h)$ tends to be small, on the order of 7.55%. For greater distances, the value of $\gamma(h)$ tends to be larger and eventually reaches about 35% for $h = 710$ m (Fig. 3).

An idealized spherical variogram of $\gamma(h)$ is presented for further illustration (Fig. 3). The idealized variogram parameters for the θ of MRI grasslands with spherical, exponential, linear and linear to sill models were calculated from Eqs. (6)–(8). The spherical and exponential models best satisfied the hypothesis according to the F-test (Eq. (9)) and have R^2 values ranging from 0.725 to 0.732 (Table 3), though the linear model was also acceptable with $R^2 = 0.645$ (Table 3). However, the linear model, with a finite sill, was not a valid variogram in the true sense, and when applied can only be used with confidence for points within the range limited by the sill (30.981). The higher sill (30.981) of the linear model than those of the spherical (27.216) or exponential (27.893) models shows that there would be a significant error if the linear model were used to estimate the θ. The spherical and

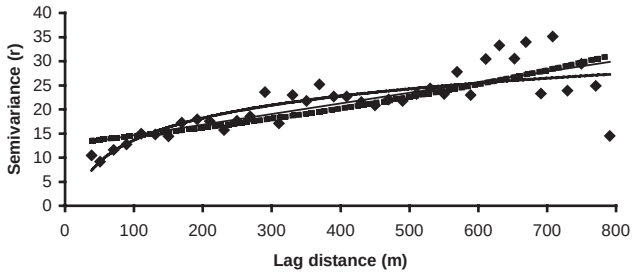

Fig. 3. Idealized variogram model. The light line is linear model; dotted line represents exponential model and heavy solid line is the spherical model.

Table 3
Isotropic variogram parameters for soil moisture content of the MRI grassland

Parameters	Nugget (C_0)	Sill ($C_0 + C_1$)	Range ($3A_0$)	Proportion $[C_1/(C_0 + C_1)]$	Correlation (R^2)
Spherical	7.550	27.216	682.00	0.723	0.725
Exponential	3.447	27.893	711.00	0.876	0.732
Linear	10.933	30.981	811.00	0.647	0.645
Linear to sill	6.831	25.715	404.00	0.734	0.684

exponential models could explain the spatial process very well, especially for small distances.

The linear variogram was also fit to the data for comparison (Fig. 3) but shows some lack of spatial fit relative to the others. Wang et al. (2001) suggested that the ratio of nugget variance to sill variance expressed as a percentage is an indication of the spatial dependence of the variable concerned. Cline et al. (1989) stated that ratios between 25% and 75% represented moderate spatial dependence, those below 25% strong spatial dependence, and all others weak dependence. Given these ranges of relative spatial dependence, the values of the ratio of nugget variance to sill variance of 27.74, 12.35, and 35.28 for the spherical, exponential and linear analyses, respectively, indicate a strong to moderate spatial dependence for semivariograms of averaged θ. Further, Fig. 3 suggests that for sampling points separated by less than 750 m, observations will be dependent. From 700 to 800 m the observations are gradually more independent. When the observation distance exceeds 800 m the observation points are fully independent.

3.4. Anisotropic axis orientation

In this analysis, we used principal axes with angles of 0°, 45°, 90°, and 135° clockwise from the base axis, North. Points aligned sufficiently close to one or another of these angles (i.e. within $\sim$22.30°) were included in the anisotropic analysis for that angle. In the 0° direction the semivariance values changes clearly, whereas at the 90° direction changes are small, while at the 45° and 135° directions change is gradual (Fig. 4). The nugget ranged from 700 to 800 m. These analyses show that no foolproof procedure exists to best-fit every situation. However, comparing semivariances, $n(h)$ (Eq. (2)) and the correlation analysis (Fig. 2) suggests that the value of both $n(h)$ and $\gamma(h)$ are similar and Eq. (3) is correct (Table 4). So Eq. (4) gives a reasonable criterion to judge whether strong stationary conditions are met or not. The analysis of the 100 samples showed that the variance of soil moisture content in the MRI grassland was stationary.

3.5. Two-dimensional analysis

A primary motivation for sampling is to make meaningful estimates of values at nearby positions in space and time. If we first assume the values to be independent,

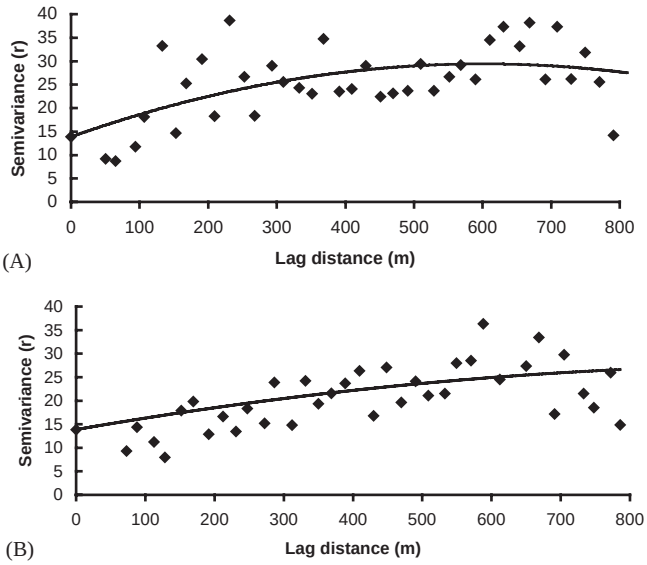

Fig. 4. Principal axis variogram for the spherical values of MRI field: (A) 0° semivariance and distance; (B) 45° and 135° semivariance and distance.

Table 4
Comparison of $\rho(h)$ values obtained from Eq. (2) or from the sample correlation analysis (Fig. 2). The coefficient of efficiency, E, for all 40 pairs = 0.5756385

$\rho(h)$ Fig. 2	Eq. (2)	Fig. 2	Eq. (2)	Fig. 2	Eq. (2)	Fig. 2	Eq. (2)	Fig. 2	Eq. (2)
0.504	0.512	0.153	0.142	−0.03	0.02	−0.097	−0.099	−0.606	−0.661
0.566	0.567	0.169	0.156	−0.192	−0.185	−0.153	−0.136	−0.097	−0.058
0.45	0.443	0.256	0.223	0.071	0.042	−0.09	0.002	−0.659	−0.669
0.397	0.385	0.165	0.164	−0.071	−0.43	−0.313	−0.330	−0.129	−0.139
0.29	0.285	0.124	0.112	−0.018	−0.001	−0.087	−0.058	−0.39	−0.445
0.298	0.302	−0.115	−0.109	0.015	0.01	−0.441	−0.442	−0.177	−0.233
0.317	0.3156	0.191	0.120	−0.043	−0.012	−0.573	−0.573	0.313	0.441
0.185	0.178	−0.088	−0.063	−0.028	−0.028	−0.443	−0.443	0.228	0.228

the best estimate for an unmeasured point will be the mean. However, for correlated values, nearby values are estimated by interpolation. Nowadays, one uses the kriging estimate methods of Krige (1966a, b), based largely on random function theory (Matheron, 1973; Journel and Huijbregts, 1978). Predicting unknown θ based on $n-1$ sampling points and comparing it to the removed measured value using the root mean square (RMS) can allow one to compare two or more methods of interpolation, and select the best one. If the values are independent of their neighbors, the best prediction value is the sample mean, and RMS will be approximately the same as the sample standard deviation, s, which can also be used as a standard for the interpolation.

Based on kriging analysis, and the use of Eqs. (6)–(8), a two-dimensional kriging map of θ (Fig. 5) and the associated variance map (Fig. 6) were generated for the MRI grassland. A high θ ($>40\%$) region existed through the middle-north portion of the field, extending from west to east, while a sizable low θ ($<40\%$) area was present in the south central region. The kriging estimates produced smooth contour (Fig. 5) that approximated the general form of the original contours. Standard errors of predicted values were greatest, as expected, around the borders of the field (Fig. 6), with coefficients of variation of 2–3%, suggesting nonetheless a good prediction

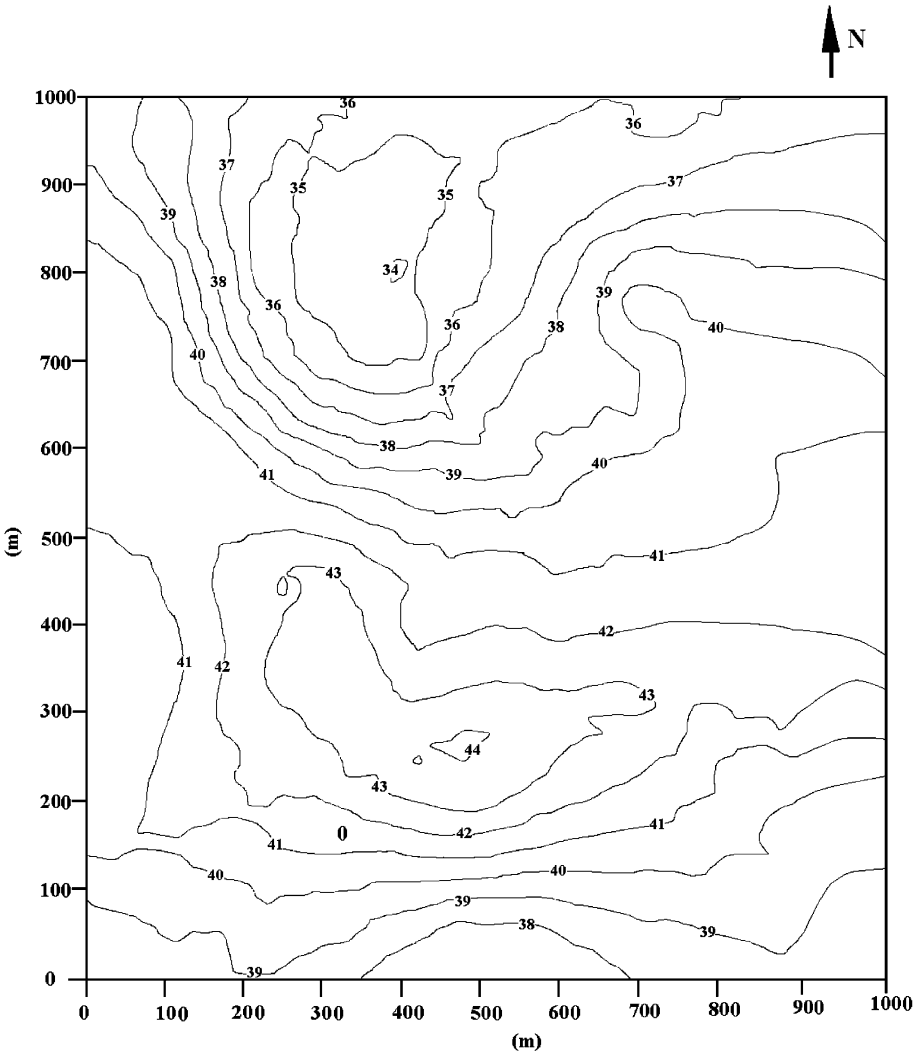

Fig. 5. Map of observation points contours by kriging.

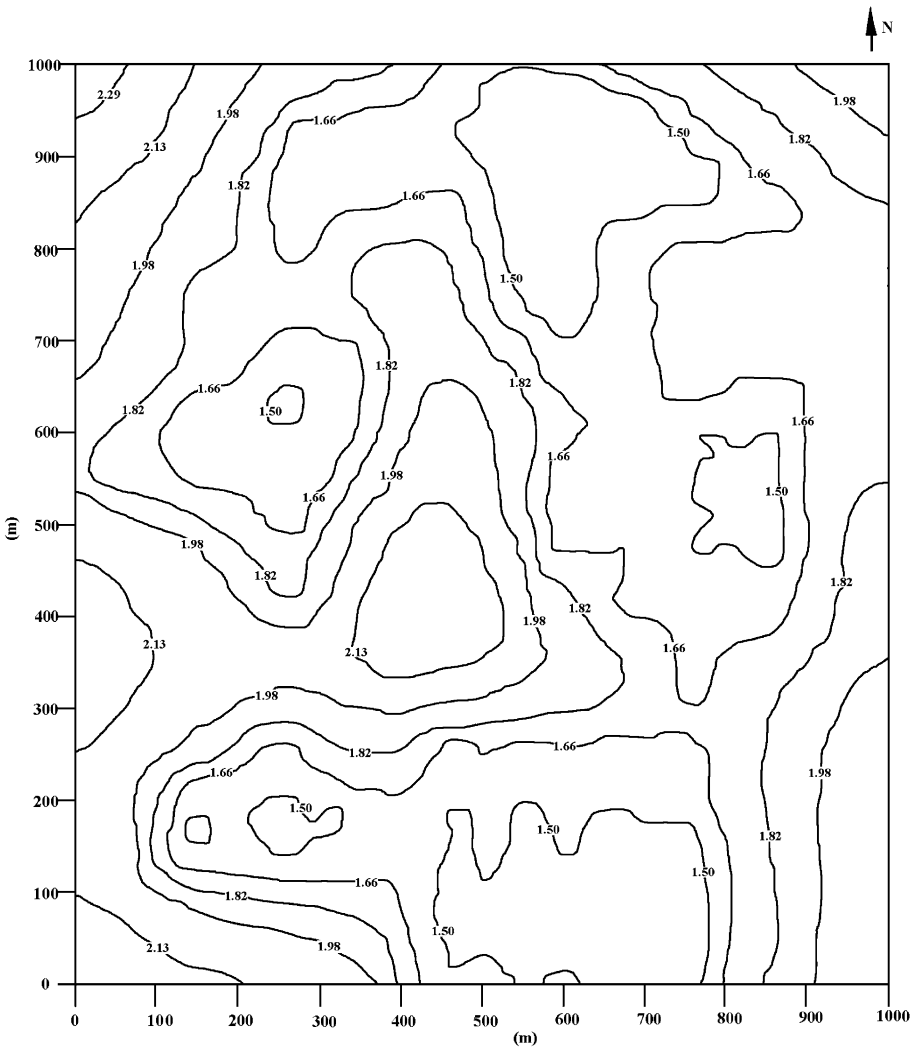

Fig. 6. Map of observation points contours by the kriged variance.

ability. The kriging variance indicates which regions' θ tends to be predicted with the greatest confidence. The mean error over the study area was 1.71, which is well below 5%. The results presented imply that the kriging estimator with Gaussian semivariogram and measurement error models was appropriate and that exact interpolation can be used in this grassland.

In order to check the difference between θ predicted by kriging and field observations of θ, three sets of eight θ samples were obtained within the study area, one in each of January, June and October 2001 (Table 5). In Table 5 an E value of 1 represents a perfect match, a value of 0 (zero) represents a match no better than

Table 5
Coefficient of efficiency, E, site-wise across sampling dates, sampling date-wise across sampling sites, and overall for the difference between predicted (pred.) and observed (obs.) values of volumetric soil moisture content, in the year of 2001

Permanent sampling location[a]	Observed and predicted θ						E
	January		June		October		
	Obs.	Pred.	Obs.	Pred.	Obs.	Pred.	
A	33.46	33.08	34.46	35.02	35.33	35.46	0.729
E	38.17	41.34	41.02	38.27	40.48	38.50	−3.697
B	35.91	36.21	35.30	35.91	36.71	35.97	−0.010
C	43.85	42.85	43.50	43.66	42.37	43.05	−0.243
D	37.61	37.10	42.5	42.6	43.16	42.62	0.969
F	37.09	37.21	38.41	38.09	38.19	37.99	0.843
G	36.65	37.65	40.22	39.81	38.25	37.63	0.757
H	39.70	38.90	40.17	39.89	39.67	38.98	−6.595
E	0.796		0.881		0.879		0.882 (all)

[a] Location of sites shown in Fig. 1.

substituting the mean of measured values, and progressively more negative numbers represent poorer and poorer matches. The relative error between the predicted and observed is less than 8.3 in the whole 8 fixed locations, except for point E, the relative error in the other points is less than 2.7 in three seasons (Table 5). The higher relative error at sampling point E may be related to the low and variable bulk densities of the different soil layers within the top 0.20 m at that site soil layer (Table 1). This will require a further investigation. The comparison suggests that kriging estimates are appropriate and that the interpolation method can be used in the study area.

4. Conclusions

The analyses performed to establish the spatial variance in θ for the MRI grassland and derived conclusions can be generalized as: A sample variogram value suggests that for small separations, values of θ tend to be more alike and, consequently, $\gamma(h)$ tends to be small, on the order of 7.55%. But for greater distances, the value of $\gamma(h)$ tend to be larger and eventually reaches about 35% for $h = 710$ m. Increasing the value of the lag distance, h, from 0 to 750 m, the θ covariance decreased gradually until eventually, at the upper limit, samples were alike and the correlation or spatial dependence non-existent. An idealized spherical variogram also showed that for a distance of less than 750 m between observations, these would be dependent. Over the range of lag distance, h, from 0 to 750 m, along the base direction, North, the semivariance value changed quickly, in the 45° and 135° directions change was more gradual, whereas in the 90° orientation changes were small. From 750 to 800 m the observations became gradually independent (Fig. 4). When the lag distance exceeded 800 m, observation points were independent.

The spherical and exponential models best satisfied the hypothesis according to the F-test. The high sill of the linear model compared to that of the spherical or exponential models showed that there would be a significant error were the linear model used to simulate θ. The spherical and exponential models could explain the spatial process very well, especially for small distances. A linear variogram was also fit to the data for comparison purposes but showed some spatial lack of fit relative to the other two. The analysis of the 100 samples showed that the variance of θ in the MRI grassland was stationary and indicates strong to moderate spatial dependence for the semivariogram of mean θ in the study.

Given that the mean error over the study area is only 1.71%, the use of the kriging estimator with Gaussian semivariogram and measurement error models was deemed appropriate, and would allow exact interpolation over the entire grassland area. This suggests that θ exhibits letter changing spatial dependence in the direction of north–south or east–west, however, the final choice of model should be based upon the investigators knowledge of hydrology, of the physical process at work, and the eventual intended use of the model. The main reason for this is that spatial analysis covers a very broad spectrum of statistic methods, and to do a comprehensive statistical analysis requires more than simply the fitting of a model.

If Kriging methods are to be used, then empirical semivariograms should be constructed, probably using different binning techniques and different variogram estimators. The variograms presented an irregular pattern for $700 > h > 600$ m and this was perhaps associated with the small number of pairs of observations for this range of h values. This is an important point given that the range of dependence was near this value.

Acknowledgements

This research was supported by a grant from the Hundred Talent Scholar Foundation (2003401), Key Project (KZCX3-SW-329; KZCX1-09-03) and (KZCX1-10-06) of Chinese Academy of Sciences, and National Nature Sciences Foundation of China (40171007).

References

Bresler, E., Kemper, W.D., Hanks, R.J., 1969. Infiltration, redistribution, and subsequent evaporation of water from soil as affected by wetting rate and hysteresis. Soil Science Society of America Proceedings 33, 832–840.

Clarke, G.P.Y., Dane, J.H., 1991. A simplified theory of point kriging and its extension to co-kriging and sampling optimization. Bulletin 609, Alabama Agricultural Experiment Station, 44pp.

Cline, M.D., 1944. Principles of soil sampling. Soil Science 58, 275–288.

Cline, T.J., Molinas, A., Julien, P.Y., 1989. An auto-CAD-based watershed information system for the Hydrologic Model HEC-1. Water Resources Bulletin 25 (3), 641–652.

Davidson, D.A., Watson, A.I., 1995. Spatial variability in soil moisture as predicted from airbore thematic mapper (ATM) data. Earth Surface Process Landform 20, 219–225.

Dirmeyer, P.A., Dolman, A.J., Sato, N.B., 1999. The pilot phase of the global soil wetness project. Bulletin of the American Meteorological Society 80 (5), 851–872.

Feng, Q., Endo, K.H., Cheng, G.D., 2002. Soil water and chemical characteristic of sandy soils and their significance to land reclamation. Journal of Arid Environment 51, 35–54.

Fitzjohn, C., Ternan, J.L., Willianms, A.G., 1998. Soil moisture variability in a semiarid gully catchment: implications for runoff and drosion control. Catena 32, 55–65.

Good, C., 1989. Surface interpolation, spatial adjacency and GIS. In: Raper, J. (Ed.), Three Dimensional Application in GIS. Taylor and Francis, London, UK, pp. 21–35.

Goodchild, M., 1992. Geographical data modeling. Computers and Geosciences, UK 18 (4), 401–408.

Hillel, D., 1974. Methods of laboratory and field investigation of physical properties of soils. In: Transactions of the 10th International Soil Science Congress, Vol. 1, Moscow, pp. 302–308.

Hillel, D., 1982. Negev: Land, Water and Life in a Desert Environment. Praeger, New York.

Hu, Z.L., Islam, S.F., 1997. Effects spatial variability on the scaling of land surface parameterizations. Boundary-Layer Meteorology 83, 441–461.

James, L.D., Burgess, S.J., 1982. Selection, calibration and testing of hydrologic models. In: Haan, C.T., Johnson, H.P., Brakensiek, D.L. (Eds.), Hydrological Modeling of Small Watersheds. American Society of Agricultural Engineers, St. Joseph, MI, pp. 215–257.

Journel, A.G., Huijbregts, C.H., 1978. Mining Geostatistics. Academic Press, London, pp. 21–32.

Krige, D.G., 1966a. A study of gold and uranium distribution patterns in the Klerksdorp gold field. Geoexploration 4 (1), 43–53.

Krige, D.G., 1966b. Two-dimensional weighted moving average trend surface forore evaluation. Journal of the South African Institute of Mining and Metallurgy 67, 13–79.

Loague, K., Green, R.E., 1991. Statistical and graphical methods for evaluating solute transport models: overview and application. Journal of Contaminant Hydrology 7, 51–73.

Matheron, G., 1973. The intrinsic random functions and their applications. Advances in Applied Probability 5, 239–465.

Mohanty, B.P., Skaggs, T.H., Famiglietti, J.S., 2000. Analysis and mapping of field-scale soil moisture variability ising high-resolution ground-based da aduring the Southern Great Plains 1997 (SGP97) Hydrology Experiment. Water Research Resource 36 (4), 1023–1028.

Perrone, J., Madramootoo, C.A., 1997. Use of AGNPS for watershed modeling in Quebec. Transactions of the American Society of Agricultural Engineers 40 (5), 1349–1354.

Polhmann, H., 1993. Geostatistical modeling of environmental data. Catena 20, 191–198.

SAS Institute Inc., 1992. The MIXED procedure. In Technical Report: SAS/STAT Software-Changes and Enhancements, Release 6.07, Cary, NC, SAS Institute, Inc., 229pp.

Schmugge, T.J., Jackson, T.J., 1996. Soil moisture variability. In: Stewart, J.B., Engman, E.T., Feddes, R.A., Kerr, Y. (Eds.), Scaling up in Hydrology Using Remote Sensing. Institute of Hydrology/Wiley, Chichester, UK.

Singh, V.P., Fiorentino, M., 1996. Geographical Information Systems in Hydrology. Kluwer Academic Publishers, Netherlands, pp. 91–113.

US National Research Council, 1994. GOALS Global Ocean-atmospheric-land System for Predicting Seasonal-to-interannual Climate. National Academy Press, Washington, DC, 103pp.

Vinnikov, K.Y.A., Yeserkepova, I.B., 1991. Soil moisture: empirical data and model results. Journal of Climate 4, 66–79.

Vinnikov, K.Y.A., Robock, A., Speranskaya, N.A., Schlosser, C.A., 1996. Scales of temporal and spatial variability of midlatitude soil moisture. Journal of Geophysical Research 101 (D3), 7163–7174.

Wang, J., Fu, B.J., Qiu, Y., Chen, L.D., Wang, Z., 2001. Geostatistical analysis of soil moisture variability on Da Nangou catchments of the loess plateau, China. Environmental Geology 41 (1), 113–116.

Zimmerman, D.L., Zimmerman, M.B., 1991. A comparison of spatial semivariogram estimators and corresponding ordinary kriging predictors. Technometrics 33, 77–91.