CHAPTER 2

THE EFFECTS OF SPATIAL VARIATION OF CERTAIN SOIL PROPERTIES ON THE WINTER YIELD OF A LUCERNE STAND

A.Venter[2*], D.J.Beukes[1], A.S. Claassens[2] & M. Van Meirvenne[3]

[1]ARC-Institute for Soil, Climate and Water, Private Bag X79, 0001, Pretoria, South Africa

[2]Department of Plant Production and Soil Science, University of Pretoria, 0002, Pretoria, South Africa

[3]Dept. Soil Management and Soil Care, Ghent University, Coupure 653, 9000, Ghent, Belgium

2.1 Abstract

In general, agricultural fields are managed as uniform units, ignoring spatial soil heterogeneity and its effects on growth and yield of field crops. This study was conducted from June 2001 – February 2002 and examines the effects of spatial variation of soil properties on the winter yield of a two-year-old lucerne stand on two soil types using geostatistical procedures. Seventy-two sampling points (nodes) were laid out on a 20 m square grid, with an additional 90 sampling points laid out on a 2.5 m square grid at six randomly selected node points to ensure that the total spatial structure would be identified. From initial soil sampling and analyses, the two experimental soils were classified as belonging to the Stella and Pyramid soil families with *inter alia* mean clay contents of 45% and 46%, pH(H_2O) values of 7.8 and 8.8, and mean P status (Ambic) contents of 18.3 and 6.4 mg kg^{-1}, respectively. Green biomass lucerne yield was determined on six occasions at all nodes, while soil sampling (0 - 300 mm layer) and analyses were done once in June 2001. Basic statistical analyses showed, for some soil properties, two distinct data populations, emphasizing the presence of two soil types. A yield prediction model ($R^2 = 0.55$) contained pH(H_2O), organic C, K and sand contents as variables. The geostatistical analyses of the yield model variables produced standard semi-variograms although with highly variable autocorrelation lengths. Making use of various kriging techniques, maps of soil properties and yield were compiled. These maps reveal that spatial variation of yield bears a fair resemblance to that of some soil properties and, therefore, supports the validity of the yield

prediction model. This study has shown that the scale of variation of lucerne yield can be related to that of soil properties, a finding which can be useful when designing sampling schemes.

2.2 Introduction

Plant nutrient management plays a vital role in the success or failure of modern lucerne (*Medicago sativa* L.) production. The production of high-yielding lucerne removes much larger amounts of nutrients from the soil than grain crops such as maize or wheat (Rhykerd & Overdahl, 1972). Lucerne has a high requirement for nutrient elements such as nitrogen (N), phosphorus (P), potassium (K), calcium (Ca), magnesium (Mg) and sulphur (S). Essential micronutrient elements are, *inter alia* boron (B) and molybdenum (Mo). Apart from desirable management practices, successful lucerne stands are obtained on deep, well-drained soils with $pH(H_2O) = 6.2 - 7.8$, P(Bray1) ≥ 25 mg kg^{-1} (P(Ambic) equivalent = 21 mg kg^{-1}), K ≥ 80 mg kg^{-1}, Ca ≥ 600 mg kg^{-1} and Mg ≥ 600 mg kg^{-1} (Fertilizer Society of South Africa, 1991). In South Africa, lucerne is produced under a wide range of conditions, according to the area of production. In the warmer regions, lucerne is produced throughout the year, which means that there is a continuous demand for nutrients under a wide range of environmental conditions.

Sensible fertilizer recommendations depend on factors such as yield level, cutting schedule and a thorough knowledge of the soil nutrient status. The precision of statements that can be made about soil properties at any location depends largely on the amount of variation within the area sampled. Soil scientists are restricted to limited observations, necessitating interpolation to estimate values at unsampled locations. The precision of such interpolations is strongly influenced by the variability of soil both within sampling units and between locations (Trangmar, Yost & Uehara, 1985). Traditionally, spatial variation and correlation of soil parameters were managed by grouping soils together in seemingly homogeneous units and assuming variability within the units to be purely random or spatially uncorrelated. That resulted in the field being managed by uniform practices such as sowing, fertilizer and pesticide applications and ignored the spatial variability of the soil and hence the site-specific crop requirements (McBratney & Pringle, 1997). Conceptually, geostatistics offers an alternative approach in that spatial correlations are quantified, and estimates for a property at an unsampled location principally determined by measurements made close by, rather than assuming a class or plot average (Warrick, Myers & Nielsen, 1986; Di, Trangmar & Kemp, 1989). The aim of this study was to quantify the variation and spatial correlations of selected soil properties that govern the yield of a lucerne stand and to predict yield using these soil properties.

2.3 Materials and Methods

2.3.1 Field and analytical methods

The study was conducted from June 2001 to February 2002 on an 18 ha lucerne stand in the Brits district in the North West Province of South Africa (27°49'47''E, 25°33'12''S). The area has a mean annual rainfall of 650 mm and the geology consists of ferrogabro and diorite of the Rustenburg Layered Suite. A rectangular area of 160 m X 140 m was demarcated as the study area. The latter comprised two soil units, which were classified (on the basis of a field survey) as a deep (1100 mm) Hutton form (Stella family) in the southwesterly corner and a deep (1000 mm) Shortlands form (Pyramid family) (Soil Classification Working Group, 1991) towards the northeasterly part of the field (Table 2.1) and covers approximately 80 % of the total area. The clay mineralogy of the two soil units was determined using the X-ray diffraction method. The lucerne stand was 2 years old when the trial commenced, and had been irrigated by a sprinkler irrigation system.

Seventy-two sampling points (nodes) were laid out on a 20 m square grid, with an additional 90 sampling points laid out on a 2.5 m square grid at six randomly selected node points to ensure that the total spatial structure would be identified. All sampling points were georeferenced using a Global Positioning System (GPS) and marked with flat metal discs. Figure 2.1 shows an aerial photograph of the field with the sampling points described as small black dots.

Table 2.1 Characteristics of the two experimental soils

Soil Family	**Stella**		**Pyramid**		
Horizons	A1	B1	A1	B2	B3
Depth (mm)	0 – 300	300 – 1000	0 – 250	250 – 550	550 – 1100
Properties					
pH(H_2O)	7.67	8.05	8.44	8.66	9.32
Org. C (%)	1.25	1.12	-	-	-
P (mg kg^{-1})	18.53	3.15	4.51	-	-
Ca (mg kg^{-1})	1618	3550	4326	4152	3068
K (mg kg^{-1})	78.2	70.4	160.3	109.5	93.8
Na (mg kg^{-1})	18.4	32.2	55.2	181.7	110.4
Mg (mg kg^{-1})	566	784	1270	2359	1914
Elec. cond. (mS m^{-1})	33	66	45	70	103
Clay (%)	42.0	48.8	45.9	54.3	37.7
Silt (%)	13.7	15.6	19.0	17.5	19.2
Sand (%)	42.3	33.4	33.7	25.7	41.4
Dominant clay mineral	Kaolinite	-	Smectite	-	-

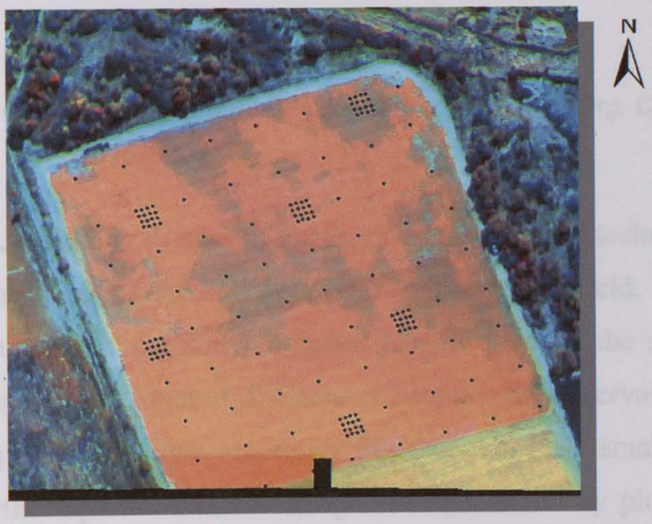


Figure 2.1 An aerial photograph of the field. The black dots depict the sampling points.

Harvesting was done by cutting and weighing the above ground plant parts within a 0.6 m square around each of the sampling points to determine green biomass yield. Starting in June 2001, yield sampling was done on six occasions, at approximate intervals of 5 weeks. At each of the sampling points three soil samples were taken in June 2001 within the 0.6 m square from the 0 – 300 mm soil layer and thoroughly mixed to serve as a composite sample. These samples were analyzed for K, Ca, Mg, sodium (Na) (ammonium acetate), P(Ambic) and organic C content, as well as for pH(H_2O), electrical resistance, particle size (hydrometer – 3 fractions) and water retention (at -33kPa) using the standard methods described in the Handbook of Standard Soil Testing Methods for Advisory Purposes (Non-Affiliated Soil Analysis Work Committee, 1990).

2.3.2 Statistical methods

For the purpose of this study only yield and soil data from the June 2001 sampling were analyzed. Basic analyses (Hintze, 1997) to obtain information on the frequency distribution, standard deviation and coefficients of variation were performed on all soil properties and lucerne yields (Table 2.2). An all-possible regression analysis (step-wise regression) was performed to identify the primary soil properties that govern the yield of the field. A model for the prediction of the yield from the soil properties was generated using a multiple linear regression analysis and is described by:

$$\text{Green biomass yield} = 512.7591 - 62.29047 \times (pH(H_2O)) + 202.8696 \times (Org.\ C) - 0.259602 \times (K) + 1.038987 \times (Sand) \qquad (R^2 = 0.55) \qquad (2.1)$$

Geostatistical analyses, including the use of the kriging interpolation technique to generate spatial presentations of the variation of the soil properties and lucerne yield, were performed (Hunt, 2002). The spatial structure of the soil properties is described by the semi-variance, which is estimated as the average of the squared differences between all observations separated by a lag distance. Consequently the points that are closer together will have smaller semi-variances than the points that are further apart. A semi-variogram is generated by plotting the semi-variance against the lag and is modeled by a mathematical function. Kriging interpolation is then used to estimate values at unsampled locations, which can be mapped (Webster & Oliver, 2000). In those cases where data populations were normally distributed, standard semi-variograms and ordinary kriging were used for estimation purposes. For bi-modal data populations the indicator kriging method was used (Goovaerts, 1997; Hunt, 2002). All estimates were contoured and mapped (Golden Software Inc., 1995) to illustrate the spatial variability of properties.

Table 2.2 Statistical descriptions of topsoil properties and yield.

	Minimum	Maximum	Mean	Median	Std. dev.	CV
pH(H_2O)	7.5	9.0	8.3	8.4	0.4	0.05
Org. C (%)	0.88	1.48	1.15	1.14	0.14	0.12
P (mg kg^{-1})	5.1	65.19	19.35	16.58	9.91	0.51
Ca (mg kg^{-1})	1565	8657	4798	5140	1570	0.33
K (mg kg^{-1})	94	468	222	221	65	0.30
Na (mg kg^{-1})	56	531	172	127	117	0.68
Mg (mg kg^{-1})	399	1917	1116	945	451	0.40
Resistance (ohm)	340	1800	699	440	488	0.70
Clay (%)	38.0	50.0	42.9	42.0	2.8	0.07
Silt (%)	12.1	35.5	22.4	21.9	4.0	0.18
Sand (%)	21.1	47.2	34.7	34.9	5.4	0.16
Water reten. (%) (at -33kPa)	16.4	37.5	27.3	26.9	3.9	0.14
Yield (t ha)	1.6	10.7	5.7	5.4	2.0	0.36

2.4 Results and Discussion

2.4.1 *Soil characteristics*

The two experimental soils have apedal (Stella) and blocky structured (Pyramid) B-horizons. Dominant clay minerals in the A horizon are kaolinite (approximately 80 %) and smectite (approximately 70 %), respectively. Both soils are deep (1000 – 1100 mm) and have high topsoil clay contents (43 – 54 % clay) with a clay texture. Soil chemical properties like pH(H_2O), Ca and Mg are markedly different between the two soils. Soil pH(H_2O) (too high) and P status (too low) (FSSA, 1991) are not conducive to optimal lucerne growth (Table 2.1), rendering P fertilization necessary.

2.4.2 *Statistical analyses*

Since the semi-variogram is based on variances, the statistical distribution of the data should ideally be close to normal to ensure that the variances are stable. However, the preliminary analyses indicated that most of the soil properties had a skew distribution (data not included) and had to be transformed. The histograms of pH(H_2O), Mg, Ca and resistance indicated that there are two distinct, relatively normally distributed populations of data. The latter is probably a result of the two soil types present in the experimental area. The histograms of the P status and exchangeable Na of the soil and that of lucerne yield were positively skewed. Several transformations (logarithmic, $\log_e$ and square-root) were performed to obtain symmetrical distributions and the best transformation for each soil property was selected. The correlation coefficients of the soil and plant properties were computed and are presented in Table 2.3.

Table 2.3 A correlation matrix for soil properties and lucerne yield.

	pH(H_2O)	C	P	Ca	K	Mg	Elec. res.	Sand	Measured Yield
pH(H_2O)	1								
C	-0.63	1							
P	-0.52	0.30	1						
Ca	0.77	-0.28	-0.57	1					
K	0.38	0.19	0.00	0.57	1				
Mg	0.78	-0.54	-0.50	0.83	0.40	1			
Elec. res.	-0.70	0.26	0.46	-0.65	-0.41	-0.56	1		
Sand	-0.32	-0.22	0.52	-0.53	-0.36	-0.30	0.41	1	
Measured Yield	-0.70	0.55	0.34	-0.57	-0.31	-0.62	0.43	0.18	1

The yield prediction model contained the variables pH(H_2O), organic C, K and sand content. In a similar study, Frogbrook, Oliver, Salahi and Ellis (2002) found that soil pH, P and K, amongst others, determined the spatial yield of a cereal crop. Although phosphorus is essential to lucerne plants in its involvement in adenosine triphosphate (ATP) associated with nitrogenase activity, the correlation between soil P and yield was relatively low ($r = 0.34$). This may be explained in that P status values were sub-optimal for good lucerne growth.

2.4.3 Spatial analyses

In this paper only those soil properties that were used in the yield prediction model are discussed (see equation 2.1). Of the four properties included in the model, only pH(H_2O) exhibited a bi-modal distribution reflecting the presence of two soil types within the experimental area. However, the southwesterly part had too few sampling points to compute a semi-variogram and indicator kriging (IK) was used to estimate this property. Indicator kriging is a non-linear, non-parametric form of kriging (Webster and Oliver, 2000) in which continuous variables are converted to binary indicators. This makes the approach suited to non-normal and crude data. The dataset was divided into nine percentile ranges (Isaaks & Srivastava, 1989) that served as the threshold values. An isotropic, indicator semi-variogram was computed for each of the percentile ranges and then used to do a multiple indicator kriging analysis.

No preferential long or short-range directions could be identified for the soil K content and thus an isotropic semi-variogram was modeled. Well-defined long and short-range an-isotropic semi-variograms were modeled for the organic C, sand content, yield and predicted yield using a double spherical model (Webster & Oliver, 2000) given by:

$$\gamma(h)=\gamma_0(h)+\gamma_1(h)+\gamma_2(h) \qquad \text{with :} \tag{2.2}$$

$$\gamma_0(h)=\begin{cases}0 & \text{if } h=0\\ C_0 & \text{if } h>0\end{cases}$$

$$\gamma_1(h)=C_1\left(\frac{3h}{2a_1}-\frac{1}{2}\left(\frac{h}{a_1}\right)^3\right) \qquad 0<h\le a_1$$

$$\gamma_1(h)=C_1 \qquad h>a_1$$

$$\gamma_2(h)=C_2\left(\frac{3h}{2a_2}-\frac{1}{2}\left(\frac{h}{a_2}\right)^3\right) \qquad h\le a_2$$

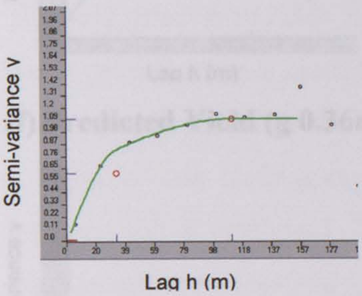
$$\gamma_2(h)=C_2 \qquad h>a_2$$

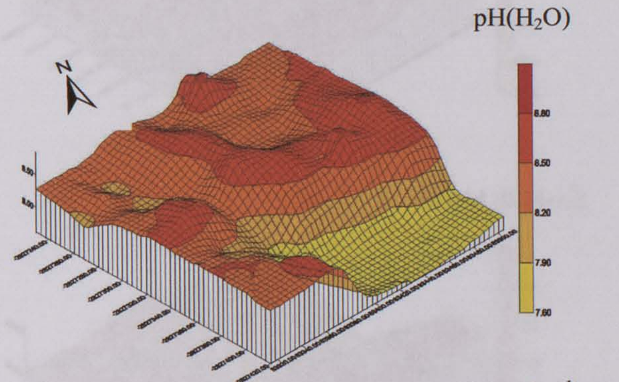
with C_0 the nugget effect, a_1 and a_2 the short and long ranges, respectively, C_1 and C_2 the sill coefficients of both structures, $C_0+C_1+C_2$ the overall sill and h the lag distance. The variograms and the maps of the kriged estimates are shown in Figure 2.2. The model parameters are given in Table 2.4. The estimation error of predicted yield was also calculated and mapped (Figure 2.3).

Table 2.4 Model parameters for soil properties and lucerne yield.

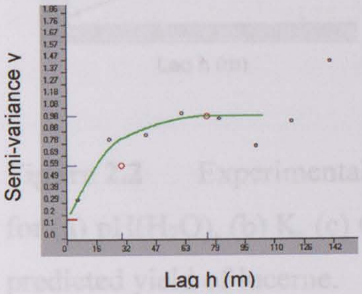
	Model	Nugget	Sill	Long range (m)	Short range (m)
pH(H_2O)	Spherical	0.01	0.98	112.3	-
K	Spherical	0.17	0.82	74.5	-
Org. C	Spherical	0.30	0.99	111.9	58.7
Sand	Spherical	0.21	0.78	220.0	75.4
Measured Yield	Spherical	0.24	1.17	90.1	48.1
Predicted Yield	Spherical	0.12	0.88	115.6	61.7

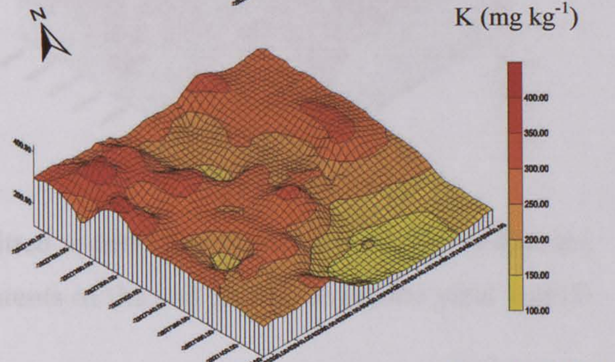
(a) pH(H_2O)



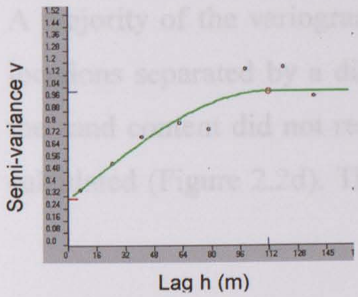


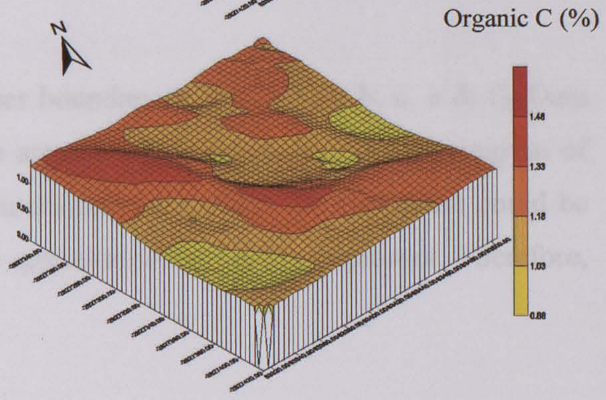
(b) K (mg kg^{-1})



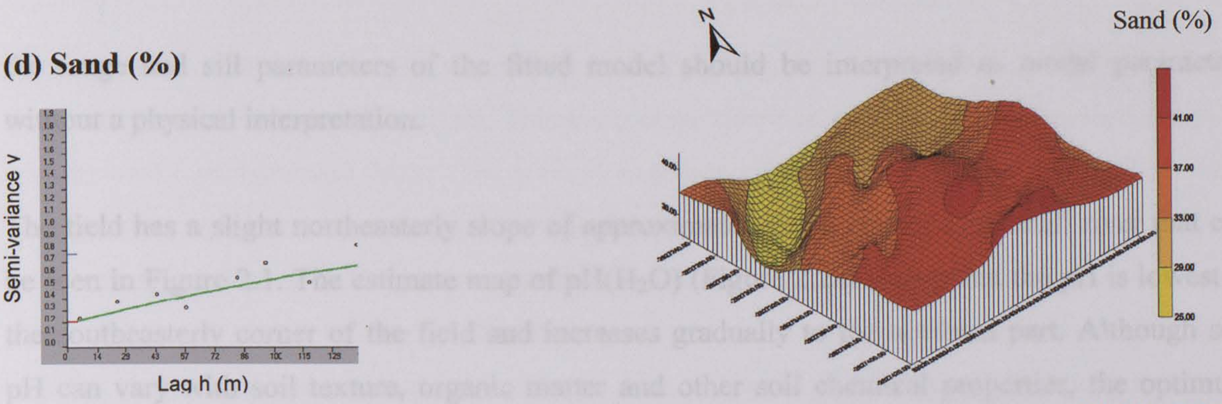


(c) Organic C (%)

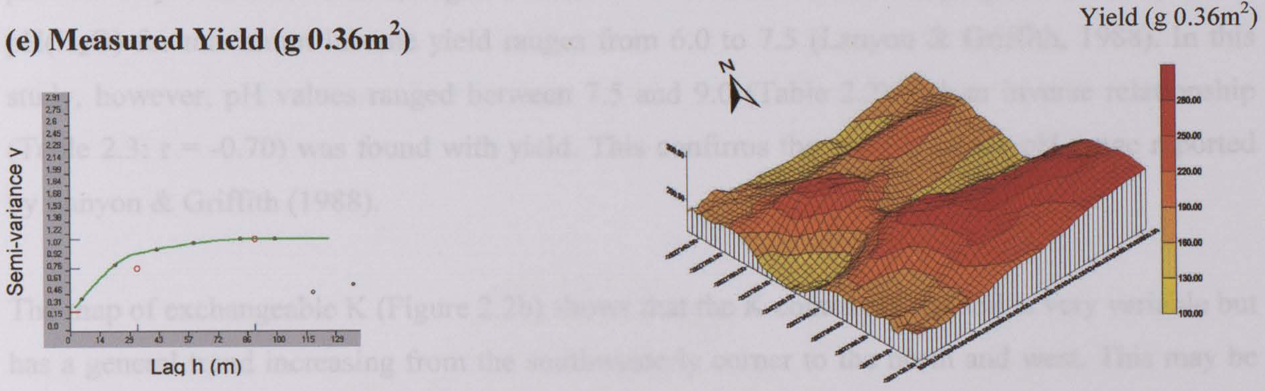




(d) Sand (%)



(e) Measured Yield (g 0.36m^2)



(f) Predicted Yield (g 0.36m^2)

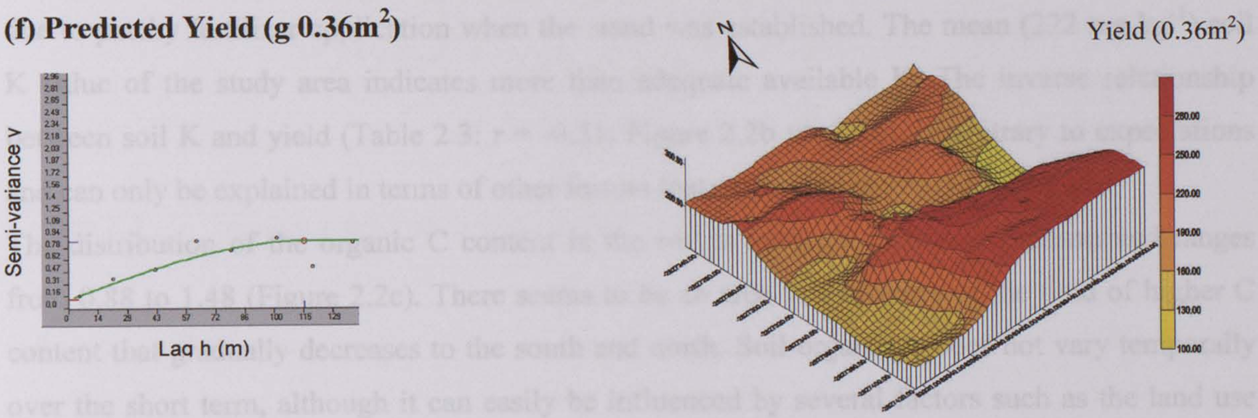


Figure 2.2 Experimental variograms with the fitted models and maps of the kriged estimates for (a) pH(H_2O), (b) K, (c) Org. C and (d) sand contents of the soil as well as (e) the yield and (f) predicted yield of lucerne.

A majority of the variograms reached a sill or upper boundary (Figures 2.2a, b, c, e & f). Data locations separated by a distance beyond the range are spatially independent. The variogram of the sand content did not reach its sill within the dimensions over which the variogram could be calculated (Figure 2.2d). This might indicate that a spatial non-stationarity is present. Therefore, the range and sill parameters of the fitted model should be interpreted as model parameters without a physical interpretation.

The field has a slight northeasterly slope of approximately 2 %, ending in a small river that can be seen in Figure 2.1. The estimate map of pH(H_2O) (Figure 2.2a) shows that the pH is lowest in the southeasterly corner of the field and increases gradually to the northern part. Although soil pH can vary with soil texture, organic matter and other soil chemical properties, the optimum pH(H_2O) for maximum lucerne yield ranges from 6.0 to 7.5 (Lanyon & Griffith, 1988). In this study, however, pH values ranged between 7.5 and 9.0 (Table 2.2) and an inverse relationship (Table 2.3: $r = -0.70$) was found with yield. This confirms the validity of the pH range reported by Lanyon & Griffith (1988).

The map of exchangeable K (Figure 2.2b) shows that the K content of the soil is very variable but has a general trend increasing from the southwesterly corner to the north and west. This may be due to patchy fertilizer application when the stand was established. The mean (222 mg kg^{-1}) soil K value of the study area indicates more than adequate available K. The inverse relationship between soil K and yield (Table 2.3: $r = -0.31$; Figure 2.2b vs. 2.2e), is contrary to expectations and can only be explained in terms of other factors that may determine yield response.

The distribution of the organic C content in the soil is spatially relatively uniform and ranges from 0.88 to 1.48 (Figure 2.2c). There seems to be an area in the centre of the field of higher C content that gradually decreases to the south and north. Soil organic C does not vary temporally over the short term, although it can easily be influenced by several factors such as the land use and management practices. It is, however, positively correlated with the lucerne yield (Table 2.3: $r = 0.55$), and is associated with higher nutrient concentrations.

Figure 2.2d shows the estimate map of the sand content of the soil. Sand content increases from the northwesterly corner across the field to the eastern corner. Sand content is an inherent soil property and cannot be manipulated by management practices. Sand content does not have a bounded semi-variogram, which means that the full extent of the spatial variation has not been encompassed at this scale of sampling. It also has a very weak correlation with yield (Table 2.3: $r = 0.18$).

The measured green biomass yield map (Figure 2.2e) shows that the values were generally larger in the southeastern corner of the field. There is a clear visual resemblance between biomass yield and the best correlated soil properties, soil pH and organic C. Although the nugget of the semi-variogram is less and the correlation range longer than that of measured yield, the map of predicted yield (Figure 2.2f) shows a good resemblance. This indicates that the green biomass yield of lucerne could be fairly accurately predicted from soil properties such as pH(H_2O), organic C, exchangeable K and sand content. The predicted yield map (Figure 2.3) showed a mean error of 21.2 %. The latter could possibly be minimized with the use of normalized differential vegetation index (NDVI) values and the inclusion of soil water features such as the water-holding capacity.

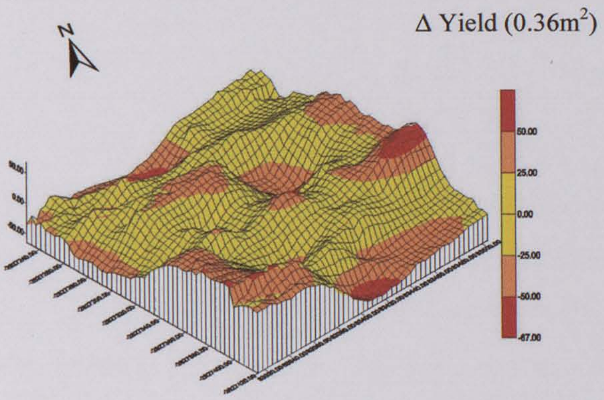


Figure 2.3 Estimation error map of predicted yield.

2.5 Conclusions

The two soils of the study site, although similar in certain aspects, exhibited differences in pH(H_2O), Ca, Mg and dominant clay minerals. These differences caused distinct bi-modal populations of data when subjected to statistical analysis. The majority of properties showed considerable variation and highly variable autocorrelation lengths. Simple linear regression analyses showed that the soil properties pH(H_2O), organic C, exchangeable Ca and Mg contents are individually well correlated with green biomass lucerne yield. A prediction model for lucerne yield (R^2 = 0.55) was obtained from stepwise multiple regression analyses. The model had pH(H_2O), organic C, exchangeable K and sand contents as variables. Although soil P status is a major nutrient element for lucerne growth, it did not feature in the prediction model. The geostatistical procedures allowed the construction of maps to demonstrate the spatial variability of soil properties and of lucerne yield. The fair resemblance between the measure and predicted yield maps supports the validity of the yield prediction model. The conclusion of Frogbrook *et al.* (2002) that the scale of variation of the yield can be related to that of soil properties is supported by this study. This can be useful in designing an appropriate sampling scheme for observing soil properties in future.