Dr. LARISA GOLOVKO **+1-281-942-8850** larisa@landviser.com

PEER-REVIEWED PUBLICATIONS

Here I include five peer-reviewed publications, most of them are from my graduate and post-graduate work at Rutgers University and University of Wyoming (under maiden name, Pozdnyakova Larisa). While working in industry, at RiceTec and Bayer, publishing in the scientific journals was not a priority and I mostly presented at grower conferences, field days and symposiums and published in proceedings and trade magazines.

Trubin, Aleksei, Yuri Manstein, and Larisa Golovko. 2022. "Electrical Geophysics for Agronomic Soil Characterization." *Modern Concepts & Developments in Agronomy* 11 (1): 1085–86. http://crimsonpublishers.com/mcda/fulltext/MCDA.000754.php.

Pozdnyakov, A. I., Pozdnyakova, L. A. and Karpachevskii, L. O. 2006. "Relationship between Water Tension and Electrical Resistivity in Soils," *Eurasian Soil Science*, 39, S78–S83

Pozdnyakova, Larisa, Daniel Giménez, and Peter V. Oudemans. 2005. "Spatial Analysis of Cranberry Yield at Three Scales." *Agronomy Journal* 97 (1): 49–57.

Pozdnyakova, L., P.V. Oudemans, M.G. Hughes, and D. Gimenez. 2002. "Estimation of Spatial and Spectral Properties of Phytophthora Root Rot and Its Effects on Cranberry Yield." *Computers and Electronics in Agriculture* 37: 57–70.

Pozdnyakova, Larisa, and Renduo Zhang. 1999. "Geostatistical Analyses of Soil Salinity in a Large Field." *Precision Agriculture* 1: 153–65. http://dx.doi.org/10.1023/A:1009947506264.

Crimson Publishers
Wings to the Research

Short Communication

Electrical Geophysics for Agronomic Soil Characterization

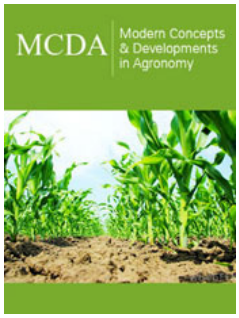
Aleksei Trubin[3], Yuri Manstein[2] and Larisa Golovko[1]*

[1]Landviser LLC, Houston, TX, USA and Landviser s.r.o., Prague, Czech Rep

[2]SiberGeo OU, Tallinn, Estonia

[3]Faculty of Forestry and Wood Sciences, Czech University of Life Sciences in Prague, 16500 Kamycká 129, Suchdol, Prague 6, Czech Republic

ISSN: 2637-7659



***Corresponding author:** Larisa Golovko, Landviser LLC, Houston, TX, USA and Landviser s.r.o., Prague, Czech Republic, Email: larisa@landviser.com

Submission: June 08, 2022

Published: July 01, 2022

Volume 11 - Issue 1

How to cite this article: Aleksei Trubin, Yuri Manstein, Larisa Golovko*. Electrical Geophysics for Agronomic Soil Characterization. Mod Concep Dev Agrono. 11(1). MCDA. 000754. 2022. DOI: 10.31031/MCDA.2022.11.000754

Abstract

Soil properties are often non-uniformly distributed within the fields. The key soil characteristics (moisture, salinity, acidity, texture, total fertility) vary across the field and dictate variable rates of agricultural inputs. Geophysical methods of Electrical Resistivity (ER) or Conductivity (EC) have been used to map soil properties since 1930 in the USA and since 1960 in Europe, as summarized in Allred et al. [1] and Pozdnyakov [2]. Methods of Direct Current (DC) such as four-electrode probe mapping and Vertical Electrical Sounding (VES) provide actual values of ER at multiple depths but slower than methods of Electromagnetic Induction (EMI). However, the depth of penetration and ER values measured with EMI methods varies depending on soil type and field condition. Farm fields were mapped with AEMP-14 in Rostov, Russia and Montrose, IA, USA, in 2021; ER at low frequencies (15-25kHz) correlated with available water at 50-70cm depth. Electrical resistivity at higher frequencies (60kHz) correlated with pH at a 10-20 cm depth. A Methodic of Calibrating EMI (AEMP-14) measurements with DC VES (Land Mapper) at a few locations is proposed to quantify ER values at multiple depths in outlined areas.

Keywords: Vertical electrical sounding; Agronomy; Soil; AEMP-14; Land mapper

Introduction

The combination of soil characteristics is the main factor affecting yields. Soil properties are often distributed unevenly over the field area. This leads to heterogeneous soil characteristics (moisture, salinity, acidity, texture, total fertility) in different parts of the field. The absence of accurate soil maps increases fertilizer application costs and irrigation water consumption, resulting in reduced yields and profits. The first studies on applying resistivity methods to soil moisture measurement are known from the 1930s. In the 1960s, the same methods were increasingly being applied in the USA and Europe to determine the salinity of soils Allred et al. [1].

Work on the electrophysical characterization of soils has been carried out in Russia since the 1990s under Prof. Pozdnyakov of Moscow State University Pozdnyakov [2]. He used the resistivity method and obtained typical VES curves for different soil types. VES is not the best methodology for investigating large areas due to the low measurement speed. Modern geophysical electromagnetic survey equipment can be successfully used for the zoning of agricultural soils.

Material and Methods

In July 2021, fieldwork was carried out in the Rostov region. AEMP-14 induction electromagnetic frequency sounding equipment was used, mounted on plastic skids 3m long,

the height of the device above the ground surface 20cm. The skids were towed by vehicle and tractor, generator of the device forward, the distance from the generator to the tow was not less than 2m (Figure 1).

Figure 1: AEMP-14 on skids.

Point datasets with resistivity values were exported with iiSystem (Version 4.04) and plotted on maps in the ArcGIS Pro (Version 2.8) environment, using the Kriging plugin (3D Analyst Tools). As a result of the research, maps of the apparent resistivity distribution were plotted, zoning was carried out, and eight sampling points were identified (Figure 2).

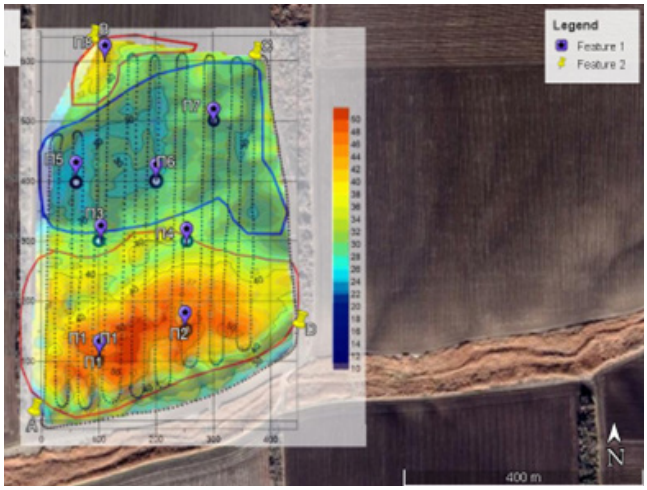


Figure 2: Resistivity map at 61 kHz with sampling points, Rostov region.

The results of the laboratory tests of the samples were compared with the apparent ground resistivity values. Correlation indices were revealed. It was shown that the moisture content in the soil at a depth of 50-100cm has the most significant influence on ground resistivity at frequencies 9-11kHz, and acidity of the upper part (20-30cm) of the section influences resistivity at frequencies 60-100kHz. Similar studies were carried out in November 2021 in the USA (Montrose, Iowa). A post-harvest cornfield was surveyed using an AEMP-14 on a skid towed by an ATV. A resistivity map with soil map element boundaries is shown in Figure 3.

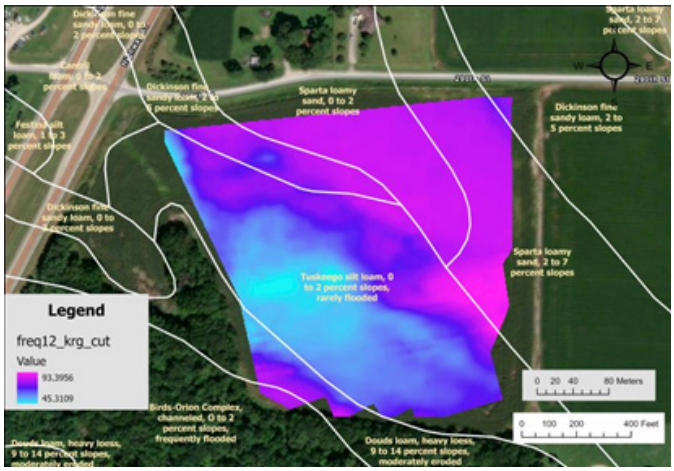


Figure 3: Apparent resistivity map, Montrose, Iowa.

Results, Interpretation, Conclusion

The data obtained show that the soil characteristics of agricultural soils can be effectively derived from the electrophysical characteristics measured with modern electrical survey equipment.

A limitation of this method is its application to soils with high resistivity. This will cause high deviations between the apparent resistivity calculated from the induction equipment's signal and the true resistivity values.

Further development of the method, according to the authors, consists of the combination of induction equipment and the Land Mapper portable resistivity method equipment. This technique can be effectively used to measure the resistivity at various spacings, plotting the VES curve. Thus, it is possible to achieve calibration of the apparent resistivity using Land Mapper data by applying it at points selected from induction electromagnetic mapping results.

Acknowledgment

This research and sensor development was supported by EU HORIZON 2020 research and innovation program, PARSEC Consortium, under grant agreement #824478 "Crop Predictions Take Flight – Linking Genomics and Geophysics".

In addition, geo-analytical services development was supported by grant no. 43950/1312/3128 "Green attack identification with the use of multi and hyperspectral data" financed by Internal Grant Agency FFWS CULS in Prague.

References

1. Allred B, Daniels JJ, Ehsani MR (2008) Handbook of Agricultural Geophysics. In: Allred B, Daniels JJ, et al. (Eds.), (1st edn), CRC Press Boca Raton, Florida, USA.
2. Pozdnyakov AI (2001) Field electrophysics of soils. Nauka/ Interperiodika, Moscow.

ISSN 1064-2293, Eurasian Soil Science, 2006, Vol. 39, Suppl. 1, pp. S78–S83. © Pleiades Publishing, Inc., 2006.

SOIL PHYSICS

Relationship between Water Tension and Electrical Resistivity in Soils

A. I. Pozdnyakov[a], L. A. Pozdnyakova[b], and L. O. Karpachevskii[a]

[a] *Faculty of Soil Science, Moscow State of University, Leninskie gory, Moscow, 119899 Russia*
e-mail: antpozd@bk.ru, karpach@soil.msu.ru

[b] *Landviser Inc., USA, e-mail: info@landviser.com*

Received April 15, 2006

Abstract—An increase in the soil water content results in a decrease in the electrical resistivity of soil. This relationship obeys an exponential equation. When the water content is below the range of capillary-film water, the clearest relationship between the electrical resistivity and water content is observed, which allows the soil and ground waters to be studied. At a high water content, the dependence of the electrical resistivity on the content of exchangeable anions, the pH, and some other properties is detectable.

DOI: 10.1134/S1064229306130138

Water is retained in the soil by Van der Waals forces, electrostatic forces between molecules of different solutes and the solid phase, and interfacial tension in the capillaries. At low water contents, the forces of molecular attraction are predominant [11, 29]. Changes in the basic mechanisms of water retention in the soil under varying water content were described by numerous authors [3, 11, 12, 17, 20, 23, 30].

Electrical parameters (such as resistivity) are exponentionally related to the density of mobile electrical charges in soils by the Boltzmann distribution law

$$\sum_{i=1}^{i=m} N_i/N_{i0} = \exp\left(-\varphi \sum_{i=1}^{i=m} \nu_i e/kT\right), \quad (1)$$

where $\sum_{i=1}^{i=m} N_i/N_{i0}$ is the relative density of the mobile electrical charges in a local volume under standard conditions, ν_i is the valence of the i-th ion, e is the electron charge, k is the universal gas constant, and T is the absolute temperature. Therefore, the density of the mobile electrical charges is exponentially related to the electrical potential. According to Ohm's law, the electrical potential is directly proportional to the electrical resistivity. If a change in some soil properties such as the water content, bulk density, or salt content results in a proportional change in the density of the mobile electrical charges, the relationship between the electrical parameters and the soil properties (SPs) can be expressed as follows:

$$SP = a_1 \exp(-b_1\varphi) = a_2 \exp(-b_2 ER), \quad (2)$$

where a_1, a_2, b_1, and b_2 are empirical parameters; φ is the electrical potential; and ER is the electrical resistivity of the soil. Some relationships between the soil properties and the bulk density of the electrical charges can be nonlinear and complicated.

In soil studies, especially in the in situ measurements of electrical parameters, it is difficult to separately study the relationships between the soil properties and electrical parameters. Therefore, the relationship (2) can be less strict when many different soil properties are measured simultaneously. However, general exponential relationships were obtained for many soil properties both under laboratory and field conditions.

Different models are reported in the literature for describing the relationships of the electrical parameters with the soil water content, temperature, and salt content (table). The electrical conductivity and resistivity are the electrical parameters usually measured under laboratory and field conditions. The relationships between the soil water content and electrical parameters were measured under field and laboratory conditions, and linear models were mainly derived [4, 6]. Linear equations were also proposed to describe the relationships between the electrical resistivity and temperature. It was shown experimentally that an exponential relationship exists between the electrical resistivity, soil temperature, and soil water content.

Many authors obtained nonlinear relationships between the electrical parameters and soil salinity [26, 27]. However, other authors showed that nonlinear relationships are derived when the soil salinity varies in a wide range.

Thus, if strongly saline samples are included in the analysis, the correlations are usually nonlinear. The shape and closeness of the correlations between the salinity and electrical parameters are given in the table. Other authors proposed polynomial functions of different orders to describe the nonlinear relationships

Some relationships between soil properties and electrical parameters (from literature data)

Author	Method	Relationship	Property series
Nizenkov, 1932	4-electrode	$W = a/ER^3 + b/ER^2 + c$	$NW - S_w$
Davydov, 1936	4-electrode	$W = a/ER^2 + b$	*N/S*
Archine, 1942	4-electrode	$EC = EC_w S_w^a \phi^b$	*N/S*
Anyang, 1961	4-electrode	$ER = e^{a+bt}W^c$	*WP – FC*
Gupta and Hanks, 1972	4-electrode	$EC = aEC_sW^b$	*HW – FC*
Rogozov, 1977	4-electrode	$W = 0.01ER + 2.1$	0.021–0.1 g g^{-1}
Trotskii, 1979	4-electrode	$ER = b - aW$	*HW – FC*
Semenov, 1980	4-electrode	$\varphi = b - aW$	*N/S*
Borovinskaya et al., 1981	4-electrode	$ER = a/W^b$	*WP – FC*
Seyfried, 1993	2-electrode	$\theta = a + b\log(ER)$	0.1–0.4 m^3 m^{-3}
Ferre et al., 1998	2-electrode	$EC = EC_w\theta^a\phi^{b-a}$	0.13–0.35 m^3 m^{-3}
Temperature (t) relationship			
Anyang, 1961	4-electrode	$ER = e^{a+bt}W^c$	*N/S*
Raisov, 1973	4-electrode	$ER_{t2} = ER_{t1}[1 + a(t_2 - t_1)]$	0–50°C
Wells, 1978	4-electrode	$\ln EC_{t1} = \ln EC_{t2} + b_1(t_2 - t_1) - b_2(t_2^2 - t_1^2)$	*N/S*
Salinization relationship			
Campbell et al., 1948	2-electrode	$TC = aEC_w^b$	*N/S*
Halvorson and Rhoades, 1976	4-electrode	$EC = aEC_w - b$	*N/S*
Chang et al., 1982	4-electrode	$TC = aEC_w^b$	*N/S*

between the electrical resistivity or conductivity and the salinity of the soils, solutions, and groundwater. Many authors [15–17] compared different relationships for calculating the ionic composition of soil solutions and found that the exponential model was the best fit. In the study of the simultaneous effect of salts and water on the electrical resistivity of soil pastes, it was found that the electrical conductivity of soil pastes is the sum of the exponential functions of the water content and soil solution conductivity [13–18].

Most relationships (table) are nonlinear or linear in a specific range of soil properties. This fact can be considered as indirect evidence of the applicability of the Boltzmann distribution law for describing the relationships between the electrical parameters and the soil properties affecting the bulk charge density. According to the Boltzmann distribution law, these relationships should be exponential in shape. The statistical differences between the exponential, power, and polynomial functions are usually insignificant, especially in a limited range of arguments.

Let us dwell on the characterization of the relationship between the electrical resistivity and the soil water content. These are the key relationships for choosing the conditions of using field electrophysical methods for soil, reclamation, or agricultural purposes. These relationships should be primarily assessed by field electrophysical techniques, because it is important to know the effect of water on, e.g., the resistance, in different ranges of water content. By selecting a specific range of water content where it affects the resistivity, we can estimate its changes; study its distribution in the soil; and, hence, assess the zone of desiccation, the depth of the groundwater table, etc. On the other hand, using electrophysical methods in the range of water content where its effect is insignificant, other soil characteristics (e.g., genetic features or chemical properties) can be studied.

The soil water content affects the mobility of electrical charges in a complicated way. Electrical charges are mobile when they occur in the free soil solution or form a double electric layer on the surface of soil particles. When the soil water content increases from air-dry to water-saturated soil, ions adsorbed on the surface of solid soil particles are released, which affects the formation of a double electric layer, and large soil pores are filled with free soil solution. Therefore, the mobility of electrical charges usually increases with increasing soil water content. However, the mobility of electrical charges is also affected by the mobility of soil water itself, because water also forms ions, including hydroxonium ions. Water is retained in the soil by molecular attraction forces; Van der Waals forces; electrostatic forces between water molecules, solute molecules, and solid surfaces. At low water contents, most soil water is strongly retained in the form of films typical for the predominant forces of molecular attraction. At high water contents, most water is retained by relatively weak capillary forces between soil particles and capillaries [11, 29]. The contributions of different forces to the reten-

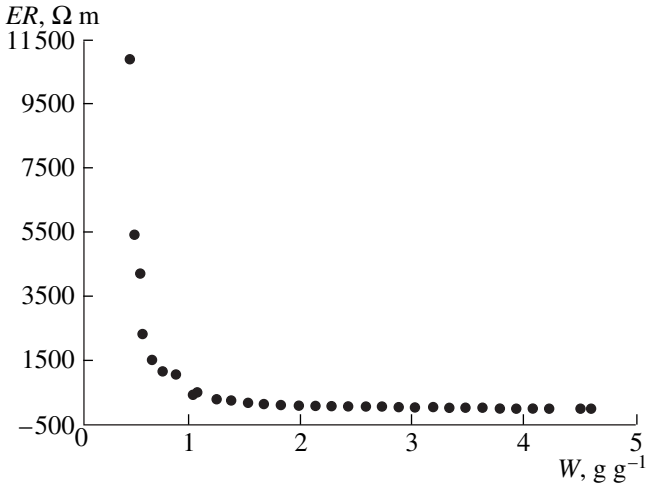


Fig. 1. An example of experimental relationship between the electrical resistivity and water content of peat soils.

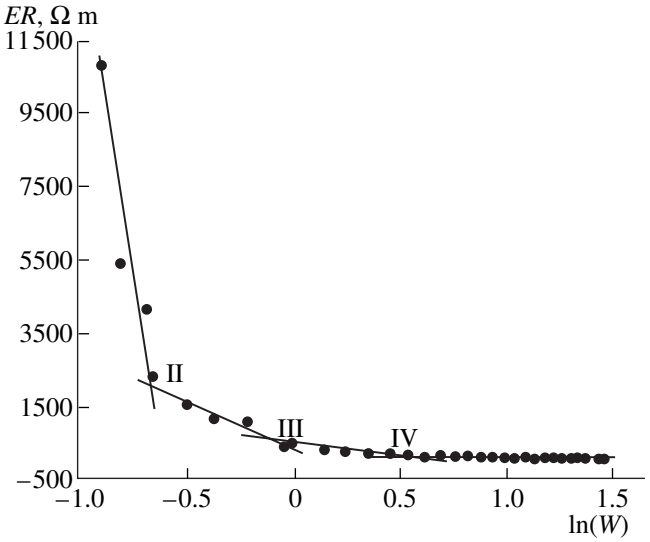


Fig. 2. An example of linearization of the relationship between electrical resistivity and water content.

tion of water vary with changing soil water content [3, 19, 23, 30]. To understand the effect of different forces on the retention of water in the soil and on the stability of electrical charges, we studied the interactions between the electrical resistivity and the soil water content under laboratory conditions.

Nonlinear relationships between the electrical resistivity and soil water content were observed for all 31 soil samples. The relationship was not linear, but it consisted of four linear fragments (Fig. 1). All together, the linear segments represented an exponential relationship between the water content and electrical resistivity in the characteristic ranges of water content and indicated similarity in the relationships of the soil water properties in these ranges (Fig. 2).

The break points between the segments correspond to the specific water contents separated into the ranges of film, film-capillary, capillary, and gravitational water according to the Voronin concept [2]. The water content in break point II is very close to the characteristic water content between the ranges of film and film-capillary water. Break point III corresponds to the water content between the ranges of film-capillary and capillary water, and break point IV corresponds to the field capacity.

Break point I was difficult to identify, because the measurements of electrical resistivity by the direct-current technique could not provide a sufficient number of experimental points in the range of the adsorbed water. The break point between the ranges of the adsorbed and film water can be determined at a high measuring frequency by electrometric methods.

The Boltzmann distribution law was used as the theoretical basis of the model, because it determines the relationship between the number of molecules and the system's energy:

$$N_i/N_j = \exp[-(E_i - E_j)/kT], \quad (3)$$

where N_i and N_j are the numbers of molecules in the states i and j, respectively, and E_i and E_j denote the system's energy in these states. Equation (3) considers all kinds of energy existing in the system, while Eq. (4) describes a system with predominant electric energy. The energy of the soil solution is affected by electrical Maxwell forces resulting from the changes in the density of the mobile electrical charges in the solution and the matrix potential [3]. Therefore, for two states i and j of the soil solution with different water contents (different energy states within the same infinite soil volume), the difference between their energies can be determined as follows [23]:

$$E_i - E_j = ze(\varphi_i - \varphi_j) + V(p_i - p_j), \quad (4)$$

where the indices i and j denote the different states of the soil solution, E is the energy of the soil solution, p is the matrix potential, and z is the total valence of the mobile ions in the solution.

Taking into account Eq. (4), Eq. (3) can be divided into the electric-potential and matrix-potential fragments as follows:

$$N_i/N_j = \exp[(-V(p_i - p_j))/kT] \times \exp[-ze(\varphi_i - \varphi_j)/kT]. \quad (5)$$

The N_i/N_j ratio represents the relative content of mobile electrical charges in the soil solution.

The content of mobile electrical charges in the soil is proportional to the water content, because electrical charges in soils are mobile only in the hydrated form, in capillaries filled with water, or in water films around the

solid soil phase (soil matrix). The difference between the electrical potentials $\varphi_i - \varphi_j$ is proportional to the electrical resistivity according to Ohm's law. Thus, from Eq. (5), the relationship between the soil water content W and the electrical resistivity ER can be expressed as follows:

$$W = a\exp(-bER). \quad (6)$$

The parameter a accounts for the effect of the soil water retention on the electrical properties of the soil solution and can depend on the physical properties of the soil that affect the water retention in the soil, such as the porosity and others. The parameter b can be related to the chemical properties of the soil and the soil solution as can the content of dissolved inorganic and organic components affecting the valence of the mobile electrical charges in the soil solution. Since the nature and magnitude of the forces influencing the soil water retention vary considerably for different water content conditions, the parameters a and b change with the soil water content. It is assumed that a and b in Eq. (6) are constants within a certain water content range but may vary for different ranges of soil water content. This assumption yields a piecewise-linear relationship between the electrical resistivity and the logarithm of the soil water content.

Different parts of the piecewise-linear relationship between the electrical resistivity or conductivity of the soil and its water content were reported in the literature [8, 10]. The break points of the relationship are related to some soil water constants such as the hygroscopic water, wilting point, and field capacity separating different soil water categories [1, 7].

Voronin [2] developed a concept of soil water categories based on the theory of molecular attraction and capillary forces. The contribution of molecular attraction to the soil matrix potential increases with increasing water content from air-dry to water-saturated soil, because the contribution of capillary forces increases. Voronin showed that the nature and magnitude of the forces contributing to the soil matrix potential change abruptly at some specific conditions separating the different ranges of soil water such as adsorbed, film, film-capillary, capillary, and gravitational water. The Voronin concept was applied to characterize the effect of soil water retention forces on the relationship between the soil water content and electrical resistivity. The soil water retention restricts the mobility of electrical charges in the soil and, hence, affects the a and b parameters in Eq. (6). Since Voronin proposed an abrupt change in the nature and magnitude of the water retention forces between the soil water ranges, the parameters a and b in Eq. (6) vary among the ranges of adsorbed, film, film-capillary, capillary, and gravitational water.

The electrical resistivity decreases rapidly with increasing soil water content in the adsorption water range (Fig. 2). Although the solute and water molecules

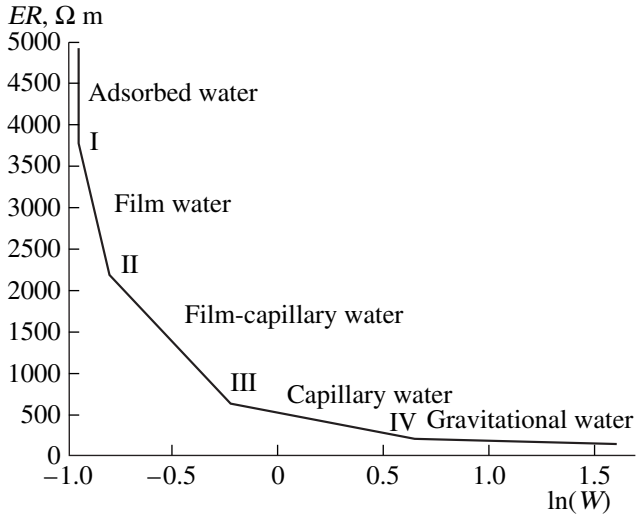


Fig. 3. A piecewise-linear relationship between the natural logarithm of water content and electrical resistivity.

on the surfaces of the soil matrix are immobile in the adsorbed water range, the dipolar molecules of adsorbed water create a conductive path in the soil matrix, which rapidly decreases the resistance when the water content increases. The adsorbed water is strongly bound to the soil matrix by the force of molecular attraction, which restricts its effect. The molecular attraction forces include Van der Walls' attractions between solid surfaces and water molecules; attractions between water and solute molecules; and electrostatic interactions between solid surfaces, water dipoles, and solute molecules. Electrostatic forces are weaker than Van der Waals forces, which retain adsorbed water in the soil; therefore, the electrical resistivity decreases in the film water range slower than in the range of adsorbed water (Fig. 3).

After the maximal possible thickness of the water film is reached, water moves from the films to fissures. In the range of film-capillary water, the relative portion of film water decreases and the content of water held in the wedges and capillaries increases [2]. Since the forces of molecular attraction are much stronger than capillary forces, the film-capillary water is still held in the soil, mainly by the molecular attraction. However, the magnitude of the capillary forces becomes increasingly pronounced within the film-capillary water range.

Since the film-capillary water is partly influenced by capillary forces, which are weaker than molecular attraction forces, electrical charges are more mobile in the film-capillary water range than in the adsorbed and film water ranges. Therefore, the electrical resistivity decreases less dramatically in the range of film-capillary water (Fig. 3). Electrical charges are even more mobile in the range of capillary water, which is more affected by capillary forces.

In the gravitational water range, the mobility of water molecules does not affect the mobility of electri-

cal charges (solute ions); therefore, the electrical resistivity is almost independent of the water content in this range [26, 27]. Nevertheless, a small decrease in the electrical resistivity can still occur in the gravitational water range due to the continuous dissolution of the adsorbed and precipitated ions from the soil solid phase (Fig. 3).

In sum, the exponential model is appropriate to describe the relationship between the electrical resistivity and the soil water content at all possible water contents. However, different model parameters should be used in the different ranges of water content (such as the adsorption, film, film-capillary, capillary, and gravitational water ranges) with different mechanisms of water retention. Theoretically, the relationship between the electrical resistivity and the logarithm of the water content should represent a series of linear segments with consequently decreasing slopes corresponding to different ranges of the water content (Fig. 3).

Thus, the interpretation based on the fundamental Boltzmann distribution law showed that the exponential model adequately describes the relationship between the water content and electrical resistivity. The model is applicable for all water conditions from air-dry to water-saturated soil. Different parameters can be used to characterize the exponential relationship under specific water content ranges, where the water retention in the soil is controlled by different mechanisms. The model parameters vary among the water content ranges such as the adsorption, film, film-capillary, capillary, and gravitational water. The proposed laboratory method is suitable for the rapid measurement of changes in water retention and electrical resistivity under varying soil water conditions.

The exponential relationship between the water content and resistivity leads to important conclusions related to the application of field electrophysical methods.

CONCLUSIONS

(1) In the low water content region, especially in the ranges of adsorbed, film, and film-capillary water, the effect of water on the resistivity is significant and can be used for its evaluation, as well as for the assessment of related parameters such as the water distribution, the zones and depths of desiccation, and the groundwater table depths in arid regions or in the humid zone during dry seasons and dry years.

(2) In the high water content region (the capillary and gravitational water ranges), where the effect of water on the resistivity is insignificant, it is advisable to use resistivity methods for studying the genetic and chemical properties of soil, e.g., for soil mapping and surveying for different soil, reclamation, and other purposes.

ACKNOWLEDGMENTS

This work was supported in part by the Russian Foundation for Basic Research, project nos. 06-04-48461 and 04-04-48580.

REFERENCES

1. L. B. Borovinskaya, V. P. Samsonova, and L. M. Plohih, "Relationship between Electrical Resistivity and Soil Water Content," Nauchn. Dokl. Vyssh. Shk., Biol. Nauki, No. 3, 23–30 (1981).
2. A. D. Voronin, *The Bases of Soil Physics* (Mosk. Gos. Univ., Moscow, 1986) [in Russian].
3. S. V. Nerpin and A. F. Chudnovskii, *Physics of the Soil* (Nauka, Moscow, 1970) [in Russian].
4. A. I. Pozdnyakov, L. A. Pozdnyakova, and A. D. Pozdnyakova, *Stationary Electrical Fields in Soils* (Moscow, 1996) [in Russian].
5. A. W. Adamson, *Physical Chemistry of Surfaces*, 2nd ed. (Interscience, New York, 1967), p. 126.
6. O. Banton, M.-K. Seguin, and M.-A. Cimon, "Mapping Field-Scale Physical Properties of Soil with Electrical Resistivity," Soil Sci. Soc. Am. J. **61**, 1010–1017 (1997).
7. G. J. Bouyoucos, "Nylon Electrical Resistance Unit for Continuous Measurement of Soil Moisture in the Field," Soil Sci. **67**, 319–330 (1948).
8. N. E. Edlefsen and A. B. C. Anderson, "The Four-Electrode Resistance Method for Measuring Soil-Moisture Content under Field Conditions," Soil Sci. **51**, 367–376 (1941).
9. D. M. Farrar, "The Use of Vapor-Pressure and Moisture-Content Measurements to Deduce the Internal and External Surface Area of Soil Particles," J. Soil Sci. **14**, 303–320 (1963).
10. C. S. Hirtzel, and R. Rajagopalan, *Colloidal Phenomena: Advanced Topics* (Noyes, Park Ridge, 1985).
11. S. Iwata, T. Tabuchi, and B. P. Warkentin, *Soil–Water Interactions: Mechanisms and Applications* (Dekker, New York, 1995).
12. W. D. Kemper, "Water and Ion Movement in Thin Films and Influenced by Electrostatic Charge and Diffuse Layer of Cations Associated with Clay Mineral Surfaces," Soil Sci. Soc. Am. Proc. **24**, 10–16 (1960).
13. W. D. Kemper, D. E. L. Maasland, and L. K. Porter, "Mobility of Water Adjacent to Mineral Surfaces," Soil Sci. Soc. Proc. **28**, 164–167 (1964).
14. D. Kirkham and W. L. Powers, *Advanced Soil Physics* (Krieger, 1972).
15. M. R. Laverdiere and R. M. Weaver, "Charge Characteristics of Spodic Horizons," Soil Sci. Soc. Am. J. **41**, 505–510 (1977).
16. D. S. McIntyre, "Basic Relationships for Salinity Evaluation from Measurements on Soil Solution," Aust. J. Soil Res. **18**, 199–206 (1977).
17. I. S. McQueen and R. F. Miller, "Approximating Soil Moisture Characteristics from Limited Data: Empirical Evidence and Tentative Model," Water Resour. Res. **10** (3), 521–527 (1974).

18. *Methods of Soil Analysis*, Part 1. *Physical and Mineralogical Methods*, 2nd ed. (ASA, Madison, Wisconsin, 1986).
19. B. N. Michurin and I. A. Lytayev, "Relationship between Moisture Content, Moisture Tension, and Specific Surface Area in Soil," Soviet Soil Sci. **8**, 1093–1103 (1967).
20. R. F. Miller, I. S. McQueen, F. A. Branson, et al., "An Evaluation of Range Floodwater Spreads," J. Range Manage. **22**, 246–257 (1969).
21. C. B. Monk, *Electrolytic Dissociation* (Academic, New York, 1961).
22. A. Nadler, "Estimating the Soil Water Dependence of the Electrical Conductivity Soil Solution/Electrical Conductivity Bulk Soil Ratio," Soil Sci. Soc. Am. J. **46**, 722–726 (1982).
23. *Soil Water*, Ed. by D. R. Nielsen, R. D. Jackson, J. W. Cary, and D. D. Evans (ASA–SSSA, Madison, WI, 1972).
24. R. M. Pashley and J. P. Quirk, "Co-ion Exclusion by Clay Surfaces: I. Equation for 1 : 1, 2 : 1, and 3 : 1 Electrolyte Solutions," Soil Sci. Soc. Am. J. **61**, 58–63 (1997).
25. B. V. Raij and M. Peech, "Electrochemical Properties of Some Oxisols and Alfisols of the Tropics," Soil Sci. Am. Proc. **36**, 587–593 (1972).
26. J. D. Rhoades, N. A. Manteghi, P. J. Shouse, and W. J. Alves, "Soil Electrical Conductivity and Soil Salinity: New Formulations and Calibrations," Soil Sci. Am. J. **53**, 433–439 (1989).
27. J. D. Rhoades, P. A. C. Raats, and R. J. Prather, "Effects of Liquid-Phase Electrical Conductivity, Water Content, and Surface Conductivity on Bulk Soil Electrical Conductivity," Soil Sci. Soc. Am. J. **40**, 651–655 (1976).
28. D. L. Sparks, *Environmental Soil Chemistry* (Academic, San Diego, 1995).
29. S. A. Taylor and G. L. Ashcroft, *Physical Edaphology: The Physics of Irrigated and Nonirrigated Soils* (Freeman, San Francisco, 1972).
30. M. Tschapek, L. Boggio, C. Wasowski and R. M. Torres Sanchez, "The Undrainable Water in Sand," Aust. J. Soil Res. **19**, 209–216 (1981).

Spatial Analysis of Cranberry Yield at Three Scales

Larisa Pozdnyakova, Daniel Giménez,* and Peter V. Oudemans

contact

ABSTRACT

Cranberry (*Vaccinium macrocarpon* Ait.) is an intensively managed perennial crop. Patches of disease, local variation in soil properties, and regional changes in soil type and hydrology cause its yield to vary spatially at several scales. We evaluated the spatial variability of cranberry yield with two support sizes and covering three scales: (i) 500 contiguous 0.09-m^2 samples covering a 6 by 7.5 m area (small scale, SS), (ii) an average number of 100 variably spaced 0.09-m^2 samples from each of 21 fields (medium scale, MS), and (iii) 534 fields (16 830 m^2 average area) each characterized with a single value of total yield (large scale, LS). Differences in yield calculated from points separated by incremental distances h were raised to power values q (from 0 to 4 in steps of 0.1). The $q = 2$ data were fitted to either spherical (SS and LS) or exponential (MS) semivariogram models. The logarithm of average differences plotted vs. log h were characterized by their slope, $\zeta(q)$. Structure functions [$\zeta(q)$ vs. q] were fitted with the universal multifractal model containing three parameters (C, α, and H). Small scale and LS data had nonlinear structure functions typical of multiscale phenomena. Spatial properties of cranberry yield at MS were: (i) better defined in cranberry fields with more than 12 yr in production (small range and nugget variance), and (ii) influenced by multiscale factors (nonlinear structure functions). Younger fields had greater range and nugget variance and a linear structure function. Precision agriculture in perennial crops should consider temporal changes in the spatial structure of crop yield.

Perennial crops are likely to develop persistent spatial variability in yield in response to biological and edaphic factors by developing genotypic heterogeneity over the life span of a plant (Novy et al., 1994, 1996). However, very limited research was conducted on spatial variability of yield and diseases of perennial crops (Timmer et al., 1989; Horner and Wilcox, 1996; Turechek and Maddan, 1999; Oudemans et al., 2002; Pozdnyakova et al., 2002). Cranberry (*Vaccinium macrocarpon* Ait.) is a perennial high value horticultural crop that grows on wetlands as a complete coverage of vines. Although commercial cranberry production is intensively managed, a high degree of variability has been identified within cranberry fields (beds), both by remote sensing as well as by ground based measurements (Pozdnyakova et al., 2002). Considering the characteristics of cranberry production, enhancement of profitability should be achieved through precision management rather than hectarage expansion. A necessary first step for developing a precision agriculture system for cranberries is to adequately understand the spatial variation of soil properties and crop characteristics and any possible changes of those spatial patterns over time.

Preliminary studies of the spatial variability in cranberry yield indicate that factors influencing variability differ with scale. At the small scale, those factors are likely to be soil-born fungal diseases, such as Fairy Ring and Phytophthora Root Rot (PRR). Several causal pathogens have been described in the literature as related to Fairy Ring, including basidiomycetous fungus *Psilocybe agrariella* as well as *Phialophora* and *Rhizoctonia* spp. (Caruso and Ramsdell, 1995; Zuckerman et al., 1968; Chang, 1989). *Phytophthora cinnamomi* Rands is a primarily causal agent of PRR in New Jersey (Caruso and Ramsdell, 1995). Fairy Ring expands outward from the point of infection in all directions at a rate of 0.3 to 0.4 m per growing season, whereas PRR is found as discrete patches ranging in size from <1 m^2 to about 100 m^2 (Caruso and Ramsdell, 1995). The diseases reduce root biomass and vine density leading to lower yields (Oudemans et al., 2000). Vine density can also be influenced by the genetic diversity found within a cultivar. For example, following planting, the vines will "fill in" the planted area and as the field matures (beds may remain in production for as much as 100 yr) specific spatial patterns of genetic variation in cultivar can develop. Development of rogue genotypes with low yield potential can spread into die-back patches during the life span of the field. At the field (medium) scale, cranberry yield is correlated with vine density and multiple soil properties, such as water content, infiltration rate, and temperature, as well as with elevation (Pozdnyakova et al., 2002). At the regional (large) scale yields from entire fields are correlated ($R^2 = 0.27$–0.39) with field geometry, namely with the ratio of bed perimeter to bed area (Pozdnyakova, unpublished data, 2001). The correlation between field geometry and cranberry yield might be caused by better drainage in beds with large perimeter/area ratios since all the beds typically have drainage ditches at the edges. Beds of similar geometry also tend to be clustered (Fig. 1), resulting in a spatial dependence in yield. Other factors that might contribute to yield autocorrelation at the large scale are soil type, microclimate, and regional hydrology.

Geostatistical methods—and more recently fractal and multifractal techniques—are often used to characterize soil properties and to describe their spatial variation (Burrough, 1983; Zhang et al., 1997; Kravchenko et al., 1999; Eghball et al., 1999). These methods are particularly well developed for field crop systems, such as wheat (*Triticum aestivum* L.), corn (*Zea mays* L.), and soybean [*Glycine max* (L.) Merr.] (Brownie et al., 1993; Cassel et al., 1988; Blackmore, 2000; Kravchenko

L. Pozdnyakova and P.V. Oudemans, P.E. Marucci Center for Blue/Cranberry Res. and Ext., Rutgers Univ., 125A Lake Oswego Rd., Chatsworth, NJ 08019-2006; and D. Giménez, Dep. of Environ. Sciences, Rutgers Univ., New Brunswick, NJ. Received 17 Dec. 2003. *Corresponding author (gimenez@envsci.rutgers.edu).

Published in Agron. J. 97:49–57 (2005).

677 S. Segoe Rd., Madison, WI 53711 USA

Abbreviations: Lng_Rng, long range; PRR, Phytophthora Root Rot; Srt_Rng, short range.

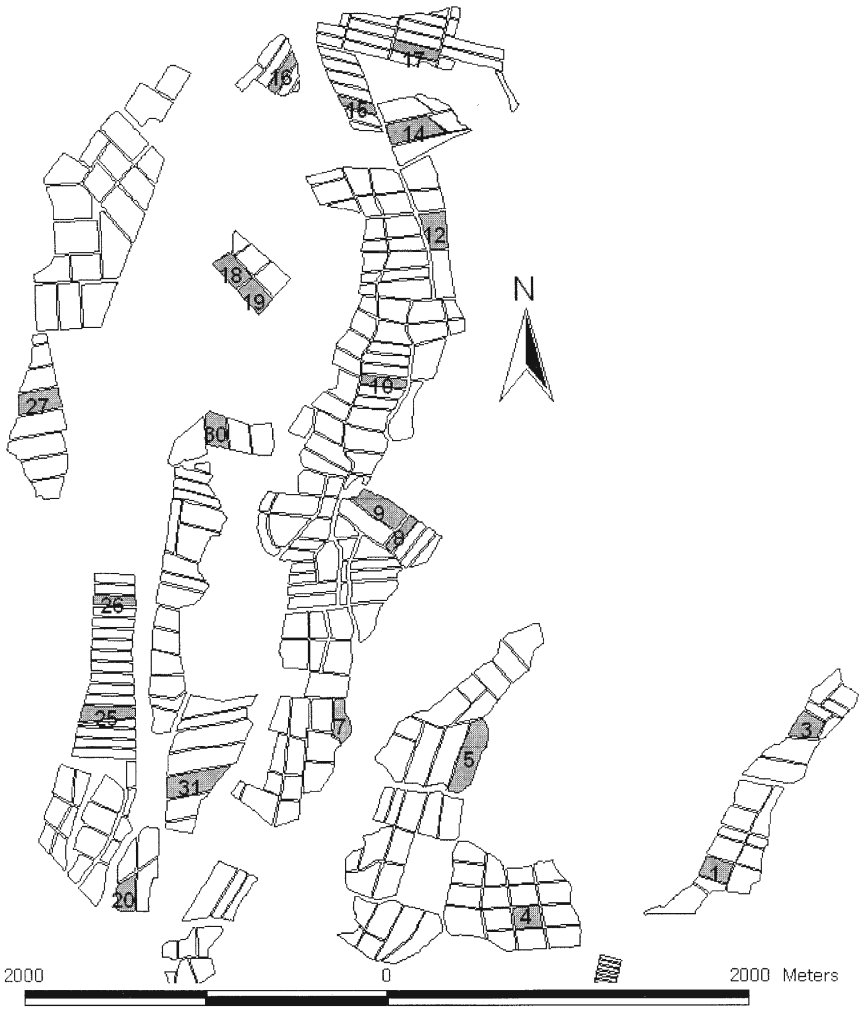


Fig. 1. Major part of the study area near Chatsworth, NJ. Total yield for entire fields was used to evaluate large scale variability. Shaded are 21 fields with dense sampling to account for within field variability (medium scale).

et al., 2000; Eghball et al., 2003), but they need to be tested in perennial systems that cover smaller areas.

The objective of the present study was to evaluate spatial variability of cranberry yield at three scales with semivariogram models and with a generalized semivariogram (structure function) approach. The results can be applied to develop sampling and spatial analysis strategies for the precision agriculture practices on cranberries, as a model system for perennial crops.

Theory

Semivariogram

A semivariogram represents the scaling of the second moment (variance) of the fluctuations (or increments) of measured values, z_i, taken at increasingly longer distances, h:

$$\hat{\gamma}(h) = \frac{1}{2n}\sum_{i=1}^{n}[z_x - z_{x+h}]^2 \qquad [1]$$

where $\hat{\gamma}(h)$ is the semivariance, whose values typically increase with an increase in h reaching a maximum (sill) at a value of h defined as the *range* (Isaaks and Srivastava, 1989). The intercept of the semivariance at $h = 0$ is defined as the *nugget*. Positive values of the nugget represent a combination of experimental error and of unresolved spatial variability occurring at scales smaller than inter-sampling lag distance (Burrough, 1993). The fraction of the sill that is not nugget variance constitutes the structural variance with values approaching unity in a strongly spatially structured system (Robertson and Freckman, 1995), or zero in a system that has little spatial structure at the scale of observation (Morris, 1999). In other cases, $\hat{\gamma}(h)$ values can increase without evidence of reaching a sill. Unbounded semivariograms are typically the result of spatial trends, and in some cases result in a power-law semivariogram of the form (Burrough, 1983; Neuman, 1994):

$$\hat{\gamma}(h) = \frac{1}{2n}\sum_{i=1}^{n}[z_x - z_{x+h}]^2 \propto h^{2H} \qquad [2]$$

where H is the Hurst exponent, a real number with values between $0 < H < 1$. The value of $H = 0.5$ indicates noncorrelated increments (white noise), whereas values of $H < 0.5$ indicates negative correlation (phenomena dominated by short-range variability), and $H > 0.5$ represent positive correlation (phenomena dominated by long-range variability).

Semivariograms spanning wide range of scales may exhibit different domains of variation each dominated by a local condition. This situation is referred as to a *multiscaled pattern of variation* (Burrough, 1983; Neuman, 1994).

Structure Function

A generalization of Eq. [2] consists of extending the analysis of the spatial dependence of the increments to higher and lower moments of order q (Davies et al., 1994):

$$\sum_{i=1}^{n} \langle |Z_X - Z_{X+h}|^q \rangle \propto h^{\zeta(q)} \qquad [3]$$

where $\zeta(q)$ is an exponent, and $\langle . \rangle$ and $|.|$ indicate statistical average and absolute value, respectively. The various moments of order q weight the distribution of increments differently and provide information on the spatial dependence of increments with small and large values. For instance, high order moments emphasize the largest values in a distribution of increments, whereas $q = 1$ reproduces their original distribution. The shape of the structure function $\zeta(q)$ vs. q defines the complexity of a distribution. If $\zeta(q)$ is a linear function of q, its value reduces to qH (single scaling), and the Hurst exponent H (see Eq. [2]) is sufficient to characterize it. On the other hand, a nonlinear structure function (multiscaling) requires more than one parameter for its characterization (Davies et al., 1994).

The structure function can be characterized with three parameters using the universal multifractal model (Schmitt et al., 1995):

$$\zeta(q) = qH - \frac{C}{\alpha - 1}(q^{\alpha} - q) \qquad \alpha \neq 1 \qquad [4]$$

where H, defined as $\zeta(1) = H$, characterizes the spatial correlation of the average absolute increment, C is the codimension that characterizes the inhomogeneity of the mean of the process, and α describes the degree of multifractality; $\alpha = 0$ is single scaling and $\alpha = 2$ is the lognormal multifractal case. In the case of single scaling either C or α is equal to zero and the structure function is a linear function of q, and only defined by H. Larger H values result in better defined spatial correlation (Tennekoon and Boufadel, 2003). Nonlinear structure functions typical of multifractal processes are characterized by values of C and α larger than zero. The larger the value of C the sparser the occurrence of any given increment, whereas larger values of α imply than few high increment values dominate a distribution (Seuront et al., 1999).

To date, the semivariogram remains as a standard method to quantify spatial structure of soil/crop properties and a large body of knowledge on its application exists, making this technique also appropriate as a standard for comparison of new approaches. However, using the same raw data, Eq. [3] provides a better insight on the complexity of a field than either the semivariogram or the power law semivariogram (note that Eq. [2] cannot discriminate between single and multiscaling spatial patterns because it only quantifies the second order moment). The disadvantage of Eq. [3] resides in the potentially large number of parameters needed to quantify the heterogeneity, which can be partially overcome by using the universal multifractal model (Eq. [4]), which represents the structure function with three parameters (Tennekoon and Boufadel, 2003).

MATERIALS AND METHODS

Data Collection

We analyzed the spatial variation of cranberry yield data obtained with two sample support sizes at three different scales. For the first two sampling scales (small and medium) yield was collected using a square support of 0.3 m side (0.09 m^2). Sampling squares were placed on the vines, and berries within a square were counted. Yield data included number of harvestable fruit, number of rotted fruit, as well as nonharvestable fruit. In addition, average berry weights were determined so that yields could be estimated and expressed in kg/m^2.

Small scale sampling comprised 500 contiguous (0.09 m^2) squares arranged in a 6 by 7.5 m area, where yield was recorded in 2001 from a location with Fairy Ring disease in a field planted with Ben Lear cultivar. The medium scale sampling was designed to characterize intrafield variability in 21 fields (Fig. 1) planted with three cultivars (Ben Lear, Early Black, and Stevens), which were sampled in the fall of 2002. In each field, an average number of 100 locations, separated by distances ranging from 3 to 200 m, were set before harvest using a stratified random sampling method (Sample 3.03 extension, ArcView 3.2, ESRI, Redlands, CA), and later sampled on a support of 0.09 m^2. Sampling density varied from a minimum of 21 to a maximum of 91 sampling areas per hectare, with an average of 57. Field 31 (Fig. 1) planted with Stevens cultivar and having a large area (1118 m^2) of PRR disease was sampled in 2000 by combining sampling strategies for the small and medium scales. A total of 148 randomly spaced points were set up within the field and sampled on a 0.09-m^2 support (Table 2, F31-1). In addition, 23 locations were sampled with 25 contiguous 0.09-m^2 samples arranged in a 1.5 by 1.5 m area. The small and medium scale data were pooled and analyzed together (Table 2, F31-2).

The large scale sampling quantifies yield variability among cranberry fields in New Jersey. The sample support comprised entire cranberry fields with an average field area of 16 830 m^2. Crop information for the individual fields (Year 2000) was provided by a growers' cooperative (Ocean Spray, Lakeville-Middleboro, MA), and included total harvested fruit, together with hectarage and the cultivar grown in each field. In our analysis we used the total harvested fruit value for each field, which was divided by field hectarage and the resulting value in kg/m^2 was assigned to the coordinates of the center point of the field. Distance between center points varied between 200 and 3000 m because fields were of irregular shape and often clustered at the different farms (Fig. 1).

Data Analysis

Descriptive statistics, which included mean, median, standard deviation, minimum and maximum values, were calculated with STATISTICA (StatSoft, Tulsa, OK). Skewness, kurtosis, and Shapiro-Wilk's index were used to test for normality. Shapiro-Wilk's parameter was estimated with the algorithm proposed by Royston (1982), as implemented in STATISTICA (StatSoft, Tulsa, OK).

Semivariogram

Semivariograms (Eq. [1]) were estimated using GS+ (Gamma Design Software, Plainwell, MI) (Robertson, 2000). Sample semivariograms were standardized by the sample variance for comparison across data sets. Omnidirectional semivariograms were calculated for each of the data sets because of the relatively limited number (about 100) of sampling locations (Isaaks and Srivastava, 1989). The maximum lag distance varied from 5 m with 0.5-m increments (small scale) to 100 m with 10-m increments (medium scale), resulting in 10 points to which to fit a semivariogram model for these two scales. The large-scale semivariogram was calculated for a maximum lag distance of 3000 m with 150-m increments, resulting in 20 points to which to fit a semivariogram model.

Spherical and exponential models were fitted to the experimental semivariograms (Isaaks and Srivastava, 1989). The spherical semivariogram model is defined by:

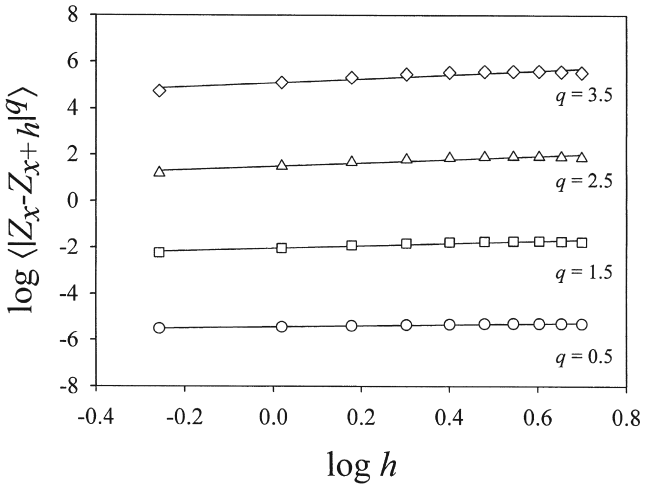


Fig. 2. Example of the fit of the log $\langle|Z_x - Z_{x+h}|^q\rangle$ vs. log (h) for different q, data for small scale yield variability.

$$\gamma(h) = \begin{cases} C_0 + C_1[1.5(h/a) - 0.5(h/a)^3] & h \le a \\ C_0 + C_1 & h > a \end{cases} \quad [5]$$

where C_0 is the nugget, $C_0 + C_1$ is the sill, and a is the range. The exponential semivariogram is similar to the spherical in that it approaches the sill, but it never reaches it (the range is usually assumed to be the point at which the model reaches 95% of the sill). The equation for an exponential model is:

$$\gamma(h) = C_0 + C_1[1 - \exp(-3h/a)] \quad [6]$$

The best model for a data set was chosen based on the maximum coefficient of determination and minimum residual sum of squares for the fit as well as through cross-validation, in which every known data point is estimated using only neighboring values (Isaaks and Srivastava, 1989). The regression coefficient between actual and estimated values and the proportion of the variation explained by the model should be as close to unity as possible.

Structure Function

The program GAMV provided in GSLIB (Deutsch and Journel, 1998) was modified to calculate multifractal spectra using Eq. [3], by raising the absolute value of the fluctuations (or increments) to q values between 0 and 4 at increments of 0.1. The output of the program included pairs of log $\langle|Z_x - Z_{x+h}|^q\rangle$ and log(h) values over the q values considered. Values of $\zeta(q)$ were estimated from the slope of the linear regression of the output pairs over the linear region of the function (Fig. 2) using the same maximal and incremental lag distances than for the corresponding semivariogram models. The functions $\zeta(q)$ vs. q were fitted with Eq. [4] using nonlinear regression to obtain values of parameters C and α, while values of H were obtained as $\zeta(q = 1) = H$ (Liu and Molz, 1997).

Table 1. Descriptive and spatial statistics for yield (10^{-3} kg/m^2) data at different scales.

Data set	Small scale	Medium scale	Large scale
No. of samples	500 (284/216)†	2407	534
Cultivar	Ben Lear	all‡	all
Support size, m^2	0.09	0.09	16 830§
Sampling density¶	107 573	57	NA
	Descriptive statistics		
Mean	3054 (4755/819)	2931	1911
Median	3344 (4802/639)	2728	1858
Min	0 (2532/0)	0	154
Max	8180 (8180/2461)	9445	5256
SD	2159 (1095/630)	1737	908
Skewness	0.04 (0.08/0.82)	0.69	0.53
Kurtosis	−1.41 (−0.41/−0.26)	0.31	0.55
Shapiro-Wilks parameter	0.91*** (0.99*/0.91***)	0.97***	0.98***
	Semivariogram		
Model	Eq. [5]	Eq. [6]	Eq. [5]
Nugget, C_0	0.00	0.33	0.27
Sill, $C_1 + C_0$	1.23	0.72	0.93
Structural variance#	1.00	0.55	0.71
Range, a	3.45	172	2297
R^2††	0.999	0.980	0.950
RSS‡‡	0.002	0.002	0.044
Total lags	10	10	20
	Structure function		
Model Eq. [4]			
H	0.50	0.12	0.25
α	1.29	NA§§	1.54
C	0.12	NA	0.06
RSS	1.0×10^{-3}	4.1×10^{-4}	3.0×10^{-3}
R^2min¶¶	0.983	0.894	0.874
R^2max	0.998	0.951	0.949
Lag skipped beginning/end	0/3	no skip	no skip

† Descriptive statistics for small scale presented as Y_{total} (Y_{max}/Y_{min}), where Y_{total} is the whole data set, Y_{max} is the group of high yield (>2500 10^{-3} kg/m^2), and Y_{min} is the group of low yield (<2500 10^{-3} kg/m^2) (See Fig. 3).
‡ Cultivars Stevens, Ben Lear, and Early Black.
§ Range from 1046 to 130 133 m^2.
¶ Sampling density in points per hectare.
# Structural variance, $C_1/(C_1 + C_0)$.
†† R^2, coefficient of determination of the fits.
‡‡ RSS, residual sum of squares of the fits.
§§ Structure function was linear, i.e., single scaling.
¶¶ R^2min/R^2max, smallest/largest R^2 of the linear fits log $\langle|Z_x - Z_{x+h}|^q\rangle$ vs. log (h) (see Fig. 2).

The points included in the fitting were selected to maximize the correlation coefficients (Eghball et al., 1999). The values used for fitting in each data set and those excluded in each case are given in Tables 1 and 2.

RESULTS AND DISCUSSION

Descriptive Statistics

Except for data at the small scale, most of the data sets could be approximated by the normal distribution (Tables 1 and 2). The Shapiro-Wilks parameter approached 1.0 for large and medium scales, and was relatively small for the small scale (Table 1). The data characterizing small scale variability had a bimodal distribution (Fig. 3), probably because the sampling area was selected to compare yield within a diseased area with that of the surrounding healthy area. Since neither a log-normal nor a square root transformation produced a normal distribution, nontransformed data were used for all of the data sets.

Sampling support size affected the statistics of yield distribution more than the spatial scale at which measurements were made (Table 1). The mean and the standard deviation of yield decreased markedly with the increase in support size. The mean of the observations taken with a 0.09-m^2 sampling support was similar at small and medium scales, but data were more variable at the former than at the latter scale, probably because measurements at the small scale included approximately equal areas with high and low yield.

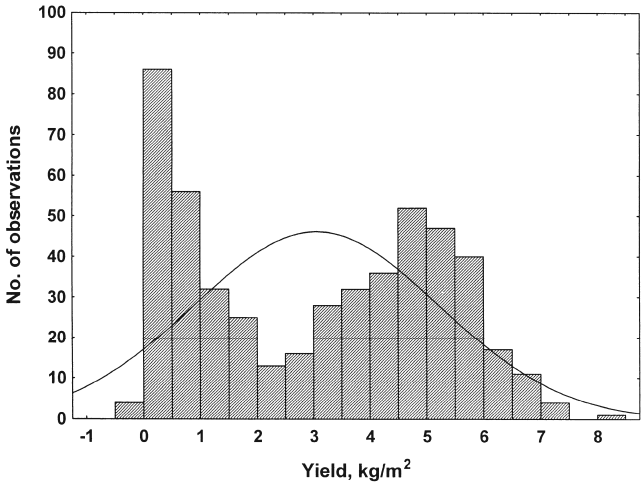


Fig. 3. Histogram of cranberry yield data for the small scale variation.

Semivariogram

Spatial variability of cranberry yield showed differences at the three scales of the study (Tables 1 and 2). At the small and large scales, semivariances reached

Table 2. Descriptive and spatial statistics for medium scale. Data from the individual fields are grouped according to short (Srt_Rng) and long (Lng_Rng) range values from the semivariogram models. Results from Field 31 (Fig. 1) when excluding (F31-1) or including (F31-2) close-spaced data.

Data set	Srt_Rng	Lng_Rng	F31-1	F31-2
No. of samples	1184	803	148	723
Cultivar	all†	all	Stevens	Stevens
Sampling density‡	59	54	43	216
		Descriptive statistics		
Mean	2738	3179	1276	1140
Median	2553	3002	971	1007
Min	0	0	0	0
Max	9222	9445	6255	6255
SD	1583	1768	1164	1040
Skewness	0.55	0.59	1.06	0.08
Kurtosis	−0.05	0.15	1.31	0.49
Shapiro-Wilks parameter	0.97	0.97	0.89	0.896
		Semivariogram		
Model	Eq. [6]	Eq. [6]	Eq. [6]	Eq. [5]
Nugget, C_0	0.14	0.36	0.51	0.16
Sill, $C_1 + C_0$	0.73	1.15	1.03	1.16
Structural variance§	0.93	0.69	0.5	0.86
Range, a	34	535	63	77
R^2¶	0.928	0.978	0.524	0.675
RSS#	0.007	0.005	0.076	0.682
Total lags	15	15	10	10
		Structure function		
Model Eq. [4]				
H	0.09	0.17	0.12	0.47
α	1.70	NA††	0.78	0.74
C	0.03	NA	0.06	0.215
RSS	1.8×10^{-6}	1.2×10^{-3}	1.1×10^{-5}	2.0×10^{-3}
R^2min‡‡	0.592	0.809	0.571	0.964
R^2max	0.954	0.840	0.890	0.978
Lag skipped beginning/end	0/4	0/4	1/1	1/1

† Cultivars Stevens, Ben Lear, and Early Black.
‡ Sampling density in points per hectare.
§ Structural variance, $C_1/(C_1 + C_0)$.
¶ R^2, coefficient of determination of the fits.
# RSS, residual sum of squares of the fits.
†† Structure function was linear, i.e., single scaling.
‡‡ R^2min/R^2max, smallest/largest R^2 of the linear fits log $\langle|Z_x - Z_{x+h}|^q\rangle$ vs. log (h) (see Fig. 2).

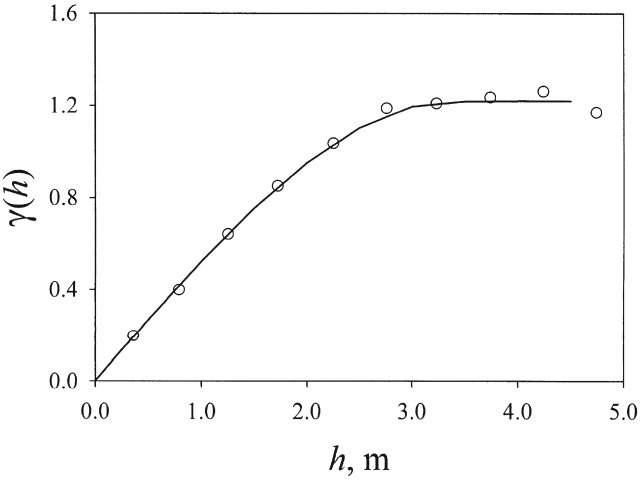


Fig. 4. Semivariogram of cranberry yield at the small scale.

a well-defined sill and were best characterized by the spherical semivariogram model (Table 1). The difference in the range at the small and large scales suggests that different factors operate at each of these scales (Gajem et al., 1981; Burrough, 1993; Gelhar, 1993; Neuman, 1994). At the small scale, yield variability was largest but its spatial structure was completely resolved by the semiovariogram function, which is to be expected given the high sampling density (continuous sampling) used at this scale (Fig. 4, Table 1). On the other hand, fitted values of the nugget, sill, and structural variance at the large scale were intermediate between the ones obtained at the small and medium scales (Fig. 5, Table 1). At the large scale, the nugget effect was probably caused by a significant intrafield variability that could not be resolved by the large sampling support size. At the medium scale, the semivariogram of all 21 fields combined had the smallest structural variance, largest nugget, and smallest sill values among the scales. The best semivariogram model at the medium scale was the exponential one, which is typical of properties showing overlapping scales of variation and when the zone of transition between scales is not defined (Burrough, 1983, 1993).

Semivariograms of data from the individual fields at the medium scale (data not shown) were fitted by the exponential model (data for only two bogs had linear model and for one field the model was spherical). The values of the fitted parameters, however, varied considerably among fields. Values for the nugget, structural variance, and range varied between 0.001 and 0.636 m, 0.269 and 0.999 m, and 13 and 233 m, respectively. Based on the similarity in the values of the range, data from fields were pooled into two groups. Twelve fields had short range (<70 m) whereas seven fields had longer range (>70 m). The group with short range (Srt_Rng) typically comprised fields that were planted before 1991 with Early Black and Stevens cultivars, while the group with long range (Lng_Rng) included fields planted since 1991 with any of the three cultivars. Both data sets were described well by the exponential model, but for the group of fields planted before 1991 (Srt_Rng) the value of the nugget/sill variance was smaller than for the group of younger fields (Lng_Rng), suggesting that the spatial structure of cranberry yield becomes more defined over time.

Medium- and small-scale yield variability was studied together by combining the data from both scales gathered in Field 31 (Fig. 1). When only 148 widely spaced data points were analyzed, the semivariogram parameters were intermediate between the Srt_Rng and the Lng_Rng groups (Fig. 6a, Table 2). However, when

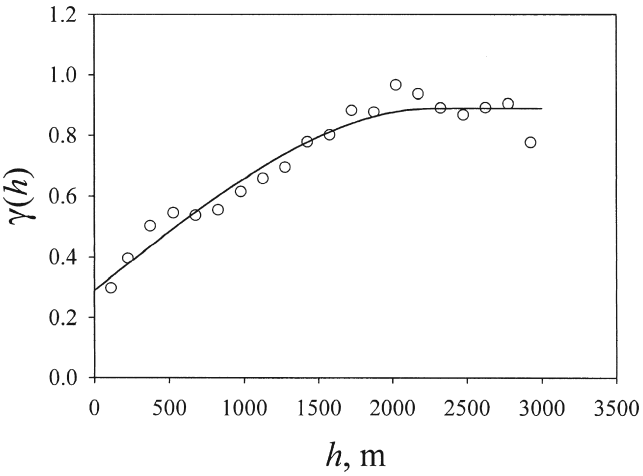


Fig. 5. Semivariogram of cranberry yield at the large scale.

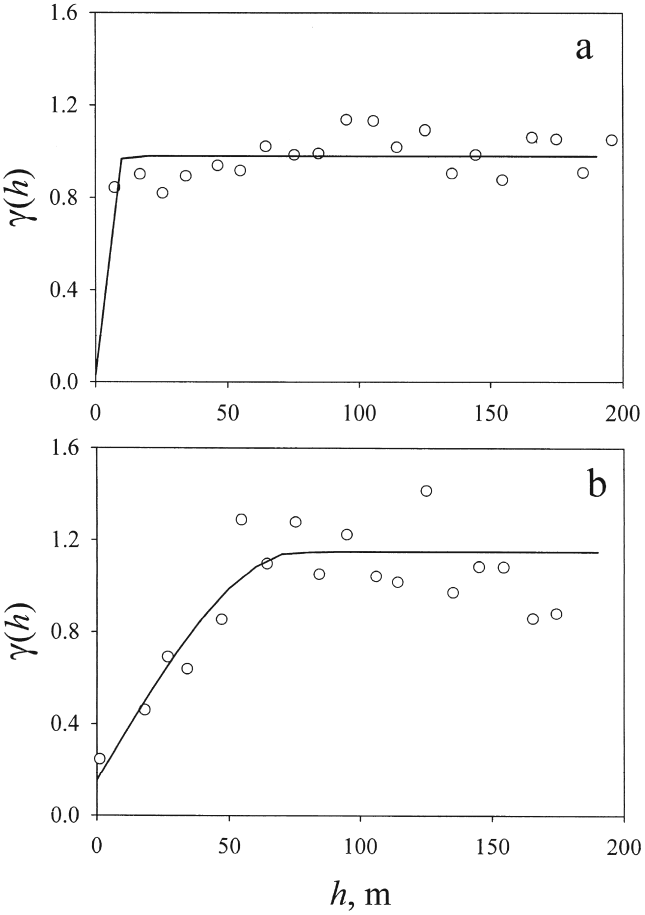


Fig. 6. Semivariograms of cranberry yield in Field 31 (see Fig. 1) sampled with (a) 148 wide spaced points and (b) for 723 wide and close sampled points combined.

additional closely spaced points were added to the data, increasing the total number of points to 723, the model that described the data best was the spherical one (Fig. 6b). Incorporation of the closely spaced points improved the fit of the semivariogram model at the short lag distances, but the fit at the longer lags was worse (Fig. 6b). Overall, the goodness-of-fit of the semivariogram model decreased (larger residual sum of squares) when the closely spaced points were added to the data (Table 2). Although the range did not change substantially with the addition of close spaced points, the nugget decreased and the structural variance increased because of improved information on short-range variability. The variation at scales smaller than 10 m accounted for about 36% of the total variability in the data, whereas 50% was attributed to variability at the medium scale, and the remaining 14% was undefined.

Structure Function

The increase in the values of semivariogram ranges with scale (and the fact that variability at the small and medium scales cannot be described with the same model or by the same set of parameters) suggest that yield varies at several scales simultaneously because of complex interaction of environmental factors acting at different scales. Multiscale phenomena are characterized by the deviation from linearity of the structure functions (Eq. [4]) (Davies et al., 1994). The general slope of the structure function is determined by the value of H, while deviations from linearity are determined by the values of C and α.

Yield data collected at the small and large scales shared multiscale properties (i.e., nonlinear structure functions), while intrafield variability (medium scale) was at a single scale when data from all fields were analyzed together (Fig. 7, Table 1). The fit of the structure function at the small scale showed the best fit for all the q values considered (Fig. 2, Table 1). At this scale, the values of H and C were twice as large, while α was 16% smaller than the respective values at the large scale. Relatively large values of H and C imply a more pronounced spatial correlation of the average absolute value of the increments and more patchiness in their spatial distribution, respectively. Relatively smaller values of α indicate that the distribution of yield contain extreme variation that is probably associated with Fairy Ring disease and with the related variability in vine density (short range phenomena) that resulted in samples having zero yield (Table 1). At the large scale, the structure function indicates a more uniform spatial distribution of cranberry yield with fewer but larger deviations from the mean increment value that could be related to differences in management practices.

At the medium scale, the structure function was nonlinear in the group of fields planted before 1991 (Srt_Rng) and linear in fields planted since that year (Lng_Rng) (Table 2, Fig. 8). The value of $H = 0.17$ that characterizes the single scale in the Lng_Rng group indicates a negative correlation in yield increments produced by short range variability. The latter was probably not resolved by the smallest lag distance used and resulted in relatively large nugget values at the medium scale (Table 2). At the time of sampling, most fields in the Lng_Rng group were in production for <12 yr; therefore, it is likely that the complex interaction of the factors determining yield has not been fully expressed

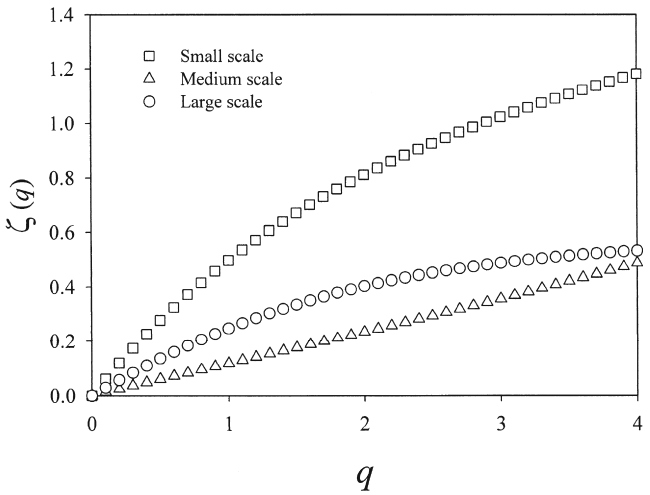


Fig. 7. Structure functions ζ(q) of cranberry yield at small, medium (all fields combined), and large scales.

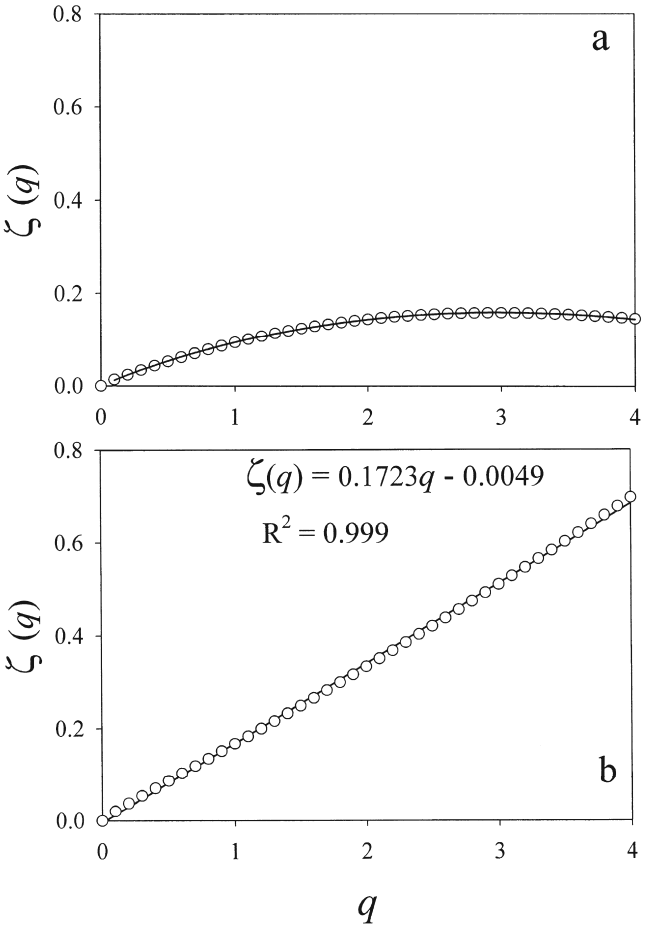


Fig. 8. Structure functions ζ(q) for the medium scale variability: (a) fields with short range (Srt_Rng) semivariograms combined, (b) fields with long range (Lng_Rng) semivariograms combined.

or some of those factors have not yet manifested themselves and so there is only one scale. The only older field included in the Lng_Rng group was a small healthy field with uniform drainage. On the other hand, multiple environmental factors may have developed over time to produce the observed multiscale spatial distribution of cranberry yield in fields of the Srt_Rng group. Two younger fields included in this group had considerable areas affected by Fairy Ring and Phytophthora Root Rot diseases, which probably contributed to their multiscale properties. Furthermore, the multiscale properties at the medium scale (Srt_Rng) were different from those at the small or large scales. The structure function of older fields had the smallest values of C and H and the largest values of α, implying that intrafield variation in these fields is relatively homogeneous (i.e., absence of patchiness) and without a strong spatial correlation. These properties are also expressed in the relatively high values of structural variance and short range in semivariograms (Table 2). The largest values of α indicates the presence of a few large deviations from the mean increment values.

The intrafield variability in the field with PRR disease was also multiscale (F31-1 and F31-2). Including the closely spaced samples to the points taken at the medium scale resulted in larger values of H (better defined spatial scaling) and C (more patchiness), whereas the overall number of increments deviating from the mean and the magnitude of the variation did not change (similar α values) (Table 2). The addition of many closely spaced points did also increase the minimal correlation coefficient of the $\zeta(q)$ estimates considerably (Table 2). The multifractal character of the structure function for this particular field was probably due to simultaneous variability of several soil properties and crop characteristic, namely severity of the PRR disease, soil water content, infiltration rate, and surface relative elevation (Pozdnyakova et al., 2002).

CONCLUSIONS

The complexity of the spatial variability of cranberry yield was manifested through the analysis of yield data collected with two support sizes and at three spatial scales. Our results suggest that intersampling distance is more important than support size in determining spatial structure of cranberry yield. Sampling schemes aimed to characterize intrafield variation should include sampling at distances of <5 m to capture short-range variation.

Field age is an important consideration when analyzing intrafield spatial variability of cranberry yield. In older fields, the spatial structure of yield is more homogeneous, containing a few areas or patches with very high yield fluctuations. In that sense, practices leading to precision agriculture are probably easier to implement in older than in younger fields. This research revealed that the most important and variable spatial structure in cranberry yield is at the field scale (defined as *medium scale* in this paper). Management of this variation either by segmented precision agriculture or by implementing long-term measures to bring the whole field to uniformity should inevitably increase overall yield and farmer's profitability. Thus, future research should focus on understanding relationships among management and yield spatial structure in perennial crops and the changes that occur over time.

The structure function is a promising approach to characterize spatial heterogeneity. Further applications of this technique in the context of precision agriculture are the analysis of remote sensing images and the simulation of field heterogeneity.

ACKNOWLEDGMENTS

This paper presents a research supported by Initiative for Future Agriculture and Food Systems financed by USDA and NASA, grant no. 2001-04782 (Enhanced management of agricultural perennial systems using GIS and remote sensing).

REFERENCES

Blackmore, S. 2000. The interpretation of trends from multiple yield maps. Comput. Electr. Agric. 26:37–51.

Brownie, C., D.T. Bowman, and J.W. Burton. 1993. Estimating spatial variation in analysis of data from yield trials: A comparison of methods. Agron. J. 85:1244–1253.

Burrough, P.A. 1983. Multiscale source of spatial variation in soil: I. The application of fractal concepts to nested levels of soil variation. J. Soil Sci. 34:577–597.

Burrough, P.A. 1993. Soil variability: A late 20th century view. Soils Fert. 56:529–562.

Caruso, F.L., and D.C. Ramsdell. (ed.). 1995. Compendium of blueberry and cranberry diseases, p. 1–78. *In* Cranberries. APS Press, St. Paul, MN.

Cassel, D.K., D.R. Upchurch, and S.H. Anderson. 1988. Using regionalized variables to estimate field variability of corn yield for four tillage regimes. Soil Sci. Soc. Am. J. 52:222–228.

Chang, L.-P. 1989. Fungi associated with fairy ring and dieback of cranberry (Vaccinium macrocarpon): History and pathogenicity. M.S. thesis. Rutgers Univ., New Brunswick, NJ.

Davies, A., A. Marshak, W. Wiscombe, and R. Cahalan. 1994. Multifractal characterizations of non-stationarity and intermittency in geophysical fields: Observed, retrieved, or simulated. J. Geophys. Res. 99:8055–8072.

Deutsch, C.V., and A.G. Journel. 1998. Geostatistical software library and user's guide. Oxford Univ. Press, New York.

Eghball, B., G.W. Hergert, G.W. Lesoing, and R.B. Ferguson. 1999. Fractal analysis of spatial and temporal variability. Geoderma 88: 349–362.

Eghball, B., J.S. Schepers, M. Negahban, and M.R. Schlemmer. 2003. Spatial and temporal variability of soil nitrate and corn yield: Multifractal analysis. Agron. J. 95:339–346.

Gajem, Y.M., A.W. Warrick, and D.E. Myers. 1981. Spatial dependence of physical properties of a Typic Torrifluvent soil. Soil Sci. Soc. Am. J. 45:709–715.

Gelhar, L. 1993. Stochastic subsurface hydrology. Prentice Hall, Englewood Cliffs, NJ.

Horner, I.J., and W.F. Wilcox. 1996. Spatial distribution of *Phytophthora cactorum* in New York apple orchard soils. Ecol. Epidemiol. 86:1122–1132.

Isaaks, E.H., and R.M. Srivastava. 1989. Applied geostatistics, Oxford Univ. Press, New York.

Kravchenko, A.N., C.W. Boast, and D. Bullock. 1999. Multifractal analysis of soil spatial variability. Agron. J. 91:1033–1041.

Kravchenko, A.N., D.G. Bullock, and C.W. Boast. 2000. Joint multifractal analysis of crop yield and terrain slope. Agron. J. 92: 1279–1290.

Liu, H.H., and F.J. Molz. 1997. Multifractal analyses of hydraulic conductivity distributions. Water Resour. Res. 33:2483–2488.

Morris, S.J. 1999. Spatial distribution of fungal and bacterial biomass in southern Ohio hardwood forest soils: Fine scale variability and microscale patterns. Soil Biol. Biochem. 31:1375–1386.

Neuman, S.P. 1994. Generalized scaling of permeabilities: Validation and effect of support scale. Geophys. Res. Lett. 21:349–352.
Novy, G.G., N. Vorsa, and K. Patten. 1996. Identifying genotypic heterogeneity in 'McFarlin' cranberry: A randomly-amplifies polymorphic DNA (RAPD) and phenotypic analysis. J. Am. Soc. Hortic. Sci. 121:210–215.
Novy, R.G., C. Kobak, J. Goffreda, and N. Vorsa. 1994. [*Vaccinium macrocarpon* (Ait.) Pursh] RAPDs identify varietal misclassification and regional divergence in cranberry. Theor. Appl. Genet. 88:1004–1010.
Oudemans, P.V., M.G. Hughes, and L. Pozdnyakova. 2000. Evaluating commercial cranberry beds for variability and yield using remote sensing techniques. p. 444–448. *In* Proc. of the Second Int. Conf. Geospatial Information in Agriculture and Forestry, Lake Buena Vista, FL. 10–12 Jan. 2000. ERIM International, Ann Arbor, MI.
Oudemans, P.V., L. Pozdnyakova, M.G. Hughes, and F. Rahman. 2002. GIS and remote sensing for detecting yield loss in cranberry culture. J. Nematol. 34:207–212.
Pozdnyakova, L., P.V. Oudemans, M.G. Hughes, and D. Gimenez. 2002. Estimation of spatial and spectral properties of Phytophthora Root Rot and its effects on cranberry yield. Comput. Electr. Agric. 37:57–70.
Robertson, G.P. 2000. Geostatistics for the environmental sciences. Version 5. Gamma Design Software, Plainwell, MI.
Robertson, G.P., and D.W. Freckman. 1995. The spatial distribution of nematode trophic groups across a cultivated ecosystem. Ecology 76:1425–1432.
Royston, J.P. 1982. An extension of Shapiro and Wilk's W test for normality to large samples. Appl. Stat. 31:115–124.
Schmitt, F., S. Lovejoy, and D. Schertzer. 1995. Multifractal analysis of the Greenland ice-core project climate data. Geophys. Res. Lett. 22:1689–1692.
Seuront, L., F. Schmitt, Y. Lagadeuc, D. Schertzer, and S. Lovejoy. 1999. Universal multifractal analysis as a tool to characterize multiscale intermittent patterns: Example of phytoplankton distribution in turbulent coastal waters. J. Plankton Res. 21:877–922.
Tennekoon, L., and M.C. Boufadel. 2003. Multifractal anisotropic scaling of the hydraulic conductivity. Water Resour. Res. 39 doi: 10.1029/2002WR001645.
Timmer, L.W., S.E. Zitko, H.A. Sandler, and J.H. Graham. 1989. Seasonal and spatial analysis of populations of Phytophthora parasitica in citrus orchards in Florida. Plant Dis. 73:810–813.
Turechek, W.W., and L.V. Madden. 1999. Spatial pattern analysis and sequential sampling for the incidence of leaf spot on strawberry in Ohio. Plant Dis. 83:992–1000.
Zhang, R., P. Shouse, S. Yates, and A. Kravchenko. 1997. Applications of geostatistics in soil science. Trends Soil Sci. 2:95–103.
Zuckerman, B.M., K.J. Rochefort, and G.B. Rounsville. 1968. Control of fairy ring disease of cultivated cranberry. Plant Dis. Rep. 52: 87–88.

ELSEVIER

Computers and Electronics in Agriculture 37 (2002) 57–70

Computers and electronics in agriculture

www.elsevier.com/locate/compag

Estimation of spatial and spectral properties of phytophthora root rot and its effects on cranberry yield

Larisa Pozdnyakova [a,*], Peter V. Oudemans [a], Marilyn G. Hughes [b], Daniel Giménez [c,1]

[a] *The Philip E. Marucii Center for Blueberry and Cranberry Research and Extension, 125A Lake Oswego Road, Chatsworth, NJ 08019 2006, USA*

[b] *Grant F. Walton Center for Remote Sensing and Spatial Analyses (CRSSA), New Brunswick, NJ, USA*

[c] *Department of Environmental Sciences, Rutgers, The State University of New Jersey, 14 College Farm Road, New Brunswick, NJ 08901-8551, USA*

Abstract

Current agricultural practices are aimed at maximizing productivity while minimizing the area of cultivated land. This is especially important in cranberry production because strict federal guidelines curtail development of new cranberry acreage on wetlands. A major component of this research is focused on the chronic effects of phytophthora root rot (PRR) because of the difficulties in detection and the significant impact on yields. PRR causes a reduction in root mass, which results in reduced canopy biomass and alters the spectral reflectance characteristics of the canopy. Detection of acute cases of PRR using color-infrared (CIR) aerial photography is straightforward from apparent bare soil on May images; however, the level of detectable chronic infection is unknown. The objectives of this study are to investigate the relationships between soil characteristics, spectral properties of the crop surface, and the severity of *Phytophthora* effects on cranberries. Soil, pathogen, and crop data were entered in a GIS and the relationships among the factors were studied using geostatistical methods and surface maps of the relevant GIS layers. These maps were then compared and incorporated with the data derived from remotely sensed images (CIR aerial photographs—May, 2001 and July, 2001). The spatial pattern of stressed vegetation was fairly consistent through 5 years and corresponded to spread of PRR chronic injury and low yield. The disease develops in surface depressions with low infiltration rates, which have high soil water content

* Corresponding author. Fax: +1-609-726-1593

E-mail address: larisa@rci.rutgers.edu (L. Pozdnyakova).

[1] Tel.: +1-732-932-9477; fax: +1-732-932-8644.

PII: S0168-1699(02)00119-9

during July–August. The results suggest that early-season (May) CIR images have more predictive power for the yield and vine density, whereas late-season (July) images are more correlated with PRR and soil infiltration rate.

Keywords: Remote sensing; Geostatistics; Cranberry yield; Phytophtora root rot

1. Introduction

Well-managed cranberry beds can yield over 500 barrels of berries per acre (56 t/ha) providing gross income of about $5000.00–$30 000.00/acre (depending on the market price). However, yields vary considerably within a bed and the average yield of a bed is generally much lower. This variability is the result of various factors including disease, weed, and insect pressures as well as variations in soil properties, water table, and water availability. As a first step towards intensification of cranberry production it is important to measure the spatial distribution of the yield and investigate the physical and biological factors influencing yield within cranberry beds. Accurate estimation of the spatial distribution of soil and crop properties within the agricultural fields requires dense sampling, which is costly and time consuming.

Color-infrared (CIR) aerial imagery can be used to identify and quantify areas with stressed vegetation and help target crop protection strategies (Everitt et al., 1999). CIR imagery interpretation is a commonly used approach for a variety of applications. Imagery is relatively inexpensive and easily available. It is relatively inexpensive and easily available. Quantification of the spectral properties of the CIR imagery and subsequent correlation with ground data, such as crop yield or yield components and soil properties, is necessary for different crop types. The ground-based data are usually collected from georeferenced points and entered into a Geographic Information Systems (GIS) where the relationships between imagery and ground data can be examined. Correlations between spectral data and plant/soil characteristics have been reported for wheat (Maas, 1991), sugar beet (Clevers, 1997), grape (Johnson et al., 1998), and cranberry (Hughes et al., 1998) fields. The modeling of the spectral and ground data was usually conducted using statistical methods of simple and multiple regression.

Geostatistical methods, such as kriging and cokriging, are commonly used estimation techniques to generate maps (raster layers) from georeferenced point data (vector layers). Kriging has been applied to quantify spatial variability of various soil and crop properties. For example, Tabor et al. (1984) and Tabor et al. (1985) used variograms and kriging to analyze and correlate the spatial variability of nitrates in cotton petioles and soil. Pozdnyakova and Zhang (1999) showed that there was a large improvement in the estimation accuracy of soil salinity when the extensive data on electrical resistivity was incorporated into analysis through cokriging. Zhang et al. (1992) utilized kriging and cokriging to improve the estimation of soil texture using the correlations with spectral properties. Geostatis-

tical analysis for agricultural plant diseases is a growing area of research and results from some studies were recently published (Ristaino et al., 1993; Nelson et al., 1999; Wu et al., 2001).

Remotely sensed spectral characteristics of the crop fields are usually available in image form, providing complete raster data coverage for the area of interest. Remotely sensed data correlate with crop/soil characteristics sampled at points on the ground. Thus, geostatistical methods coupled with imagery provide a powerful tool for characterizing spatial distributions of crop yield and its influencing factors to optimize agricultural production (Dungan, 1998). However, despite the power of geostatistical methods to characterize spatial distributions of soil and crop properties, these methods are rarely applied to study spatial structure of remotely sensed images (Carr and Myers, 1984; Stein et al., 1999) or crop disease characteristics measured as ground-based vector data (Nelson et al., 1999). In this study, kriging and multiple regression methods together with CIR photography are utilized to estimate spatial distributions of cranberry yield, phytophthora root rot (PRR), and vine density in cranberry fields in New Jersey.

2. Materials and methods

2.1. Ground data acquisition

Based on a thorough evaluation of the CIR photographs, five cranberry fields located in southern New Jersey are chosen for the current study. The annual yield data and historical information for the fields are provided by the grower. On all these fields cranberries were planted in the fall of 1992 following removal of blueberries. Problem areas are evident on black-and-white representation of CIR imagery from 23 May 1997 (Fig. 1a). The white areas across the fields resulted from the exposed sand, which was applied to improve drainage over the areas of constant cranberry dieback. Fig. 1b shows CIR of the same fields taken on 5 July 1999.

To estimate the spatial variability of crop and soil properties, 212 geo-referenced locations (Fig. 1c) were set on bed N3 in August 1999 using a differential ground positioning system (DGPS). Ground elevation was measured using a laser level at each location and the data were converted to represent relative difference in elevation between points, i.e. the minimum value of elevation in the data set was subtracted from all the data. Soil infiltration rate was measured for each point on the soil surface in September 1999 with two ring infiltrometers (Klute, 1986). The soil water content was determined gravimetrically in soil samples taken simultaneously at the same locations. Berries were counted at each location using $30.5 \times 30.5\text{-cm}^2$ frame (two replications per location) and approximated into tons per hectare for subsequent analysis. The PRR and vine density ratings were estimated visually around every sampled point (~ 4 m^2 area) using a simple rating scheme. Unity values were assigned for the dense vines and zero value was attributed if cranberry canopy was somewhat sparse in vine density estimation. Unity value was assigned if dead vines were present at the location and zero values were assigned for healthy

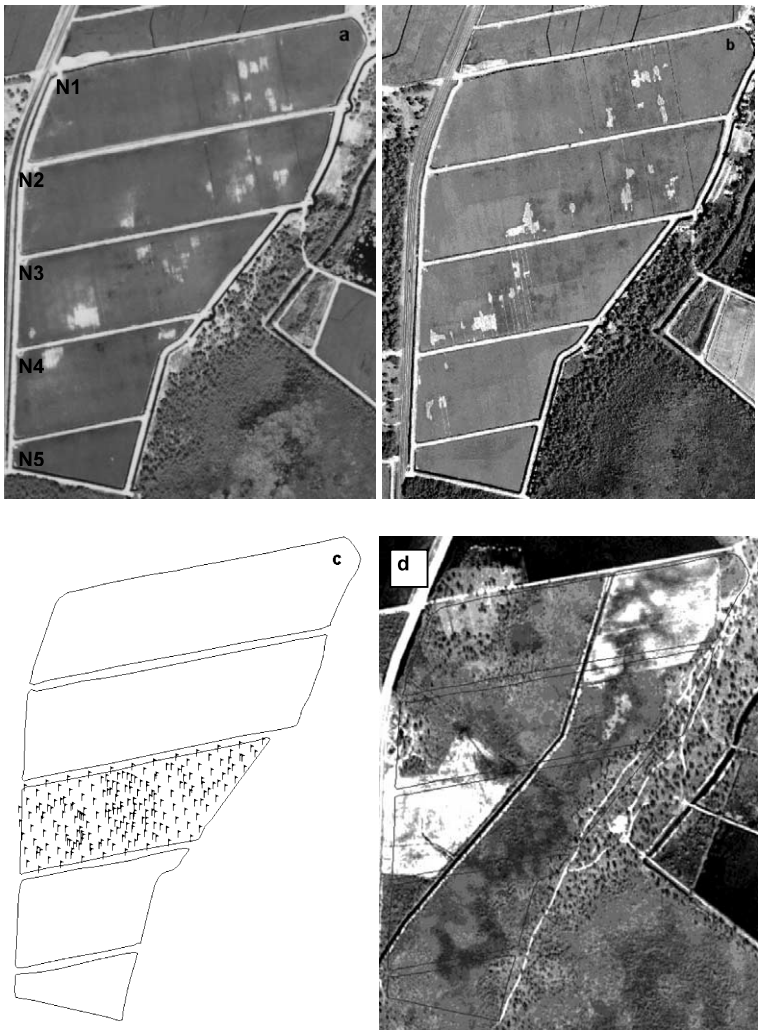


Fig. 1. Studied fields in New Jersey: Black-and-White representations of CIR at (a) 23 May, 1997 and (b) 5 July, 1999; (c) field outlines with posted sampled locations; (d) and Black-and-White aerial photography

vines in PRR evaluation. PRR was isolated from the cranberry roots taken at the sampling locations on 18 October 1999 and also through the growing season of 2000 (five isolation dates). Data suggest that percent of infestation on the root is a very dynamic characteristic and highly infected areas move within the area prone to disease, sometimes not causing a dieback of vines at a first visit. Thus, the area prone to PRR is better outlined through simple scouting of the PRR symptoms, i.e. dead vines.

2.2. Spectral data processing

The spectral data for the test areas were derived from CIR aerial photographs taken on 21 May 1999 and 5 July 1999. The 1:10 000 imagery was collected by a commercial contractor (Air Photographics, Inc.). This same contractor scanned the imagery from transparent positives at a resolution of 30 μm. Images were geo-referenced using DGPS-obtained coordinates of the visual markers around the fields (field boundaries, pump houses, etc.) and second-order polynomial transformation in IMAGINE 8.4 (ERDAS Inc., 1999). Values for spectral characteristics, i.e. Near-Infrared, Red, and Green bands, represent integer data for light reflectance at those bands on the scale from 1 to 256. No correction was made for incident radiation due to lack of the ground calibration platforms at the time of the flight. However, the extent of the bed chosen for further analysis is small (3.33 ha) and it is located near the center of the photo frame, which suggests uniform incident radiation and makes such correction not critical.

Spectral characteristics were extracted from the geo-referenced May and July images at the pixels (0.3×0.3 m^2), which spatially correspond to GPS-referenced sample locations for ground data. Since the DGPS unit used in this study supplies coordinates for the sample location with 0.5-m error, the effort was made to further process spectral data to improve co-location of aerial and ground data. Spectral data for both images were spatially averaged over 4, 9, 16, 25, and 100 pixels using a 'moving average' technique in ArcView 3.2 (ESRI, Inc., 1999). A square sampling window is set to cover several pixels (i.e. 3×3 for 9 pixels), the spectral data from pixels 'under the window' are averaged, and a value is assigned to the center of the area being averaged. At the next step, the window moves 1 pixel to the east and the process is repeated until all the data in the row are recalculated. Next row is processed in a similar fashion. This technique results in an image with the retained 0.3-m resolution, but spectral values under the each pixel incorporate the correlations for spectral data in the May image with yield and vine density is 50% higher in the May image than in July. A normalized difference vegetative index (NDVI) was derived from the spectral data extracted from the images of the fields at 21 May and 5 July 1999 for the same 212 locations (Huete, 1988):

$$\mathrm{NDVI} = (\mathrm{NIR} - \mathrm{Red})/(\mathrm{NIR} + \mathrm{Red}) \quad (1)$$

where NIR, Red, and Blue is spectral reflectance of the corresponding bands. NDVI indices for all the pixels within an image were calculated using the software package IMAGINE 8.4 (ERDAS Inc., Atlanta, GA). Results were saved as image files with a single layer.

2.3. Data analysis

The crop and image data were analyzed together with soil properties, such as elevation, water content, and infiltration rates, measured in August–September 1999 to test for the spatial relationships among the variables. The data were analyzed using descriptive statistics and correlation analyses (STATISTICA Statsoft Inc.,

1993) and geostatistical methods of variography, kriging, and cross-validation (GS+ Gamma Design Software, 1995; GSLIB Stanford University, 1997).

2.4. Geostatistical methods

The ordinary kriging estimator, $Z^*(x_0)$, of an non-sampled site is a linear sum of weighted observations within a neighborhood:

$$Z^*(x_0) = \sum_{i=1}^{n} \lambda_i Z(x_i) \tag{2}$$

where $Z^*(x_0)$ is the estimate of Z at x_0, λ_i is the weight assigned to the ith observation, and n is the number of observations within the neighborhood. The weighing factors of λ_i's are determined based on a variogram of Z.

The sample variogram is calculated from experimental data with

$$\gamma(h) = \frac{1}{2N(h)} \sum_{(i,j)h_{ij}=h} (Z_i - Z_j)^2 \tag{3}$$

where Z_i and Z_j are values of sampled variable at i and j location, respectively, separated by h lag distance, and N is a number of such pairs. The variogram (γ) is then plotted vs. lag distance and can be fitted with different models (Isaaks and Srivastava, 1989).

Cross-validation was used to evaluate the accuracy of the variogram models and kriging estimations. In this procedure, every known point was estimated using neighboring values, but not the point itself. The summary statistics from the cross validation includes mean error, mean squared error, mean kriging variance, reduced kriging variance, correlation between estimates and error, and correlation between estimated and actual values (Isaaks and Srivastava, 1989). Mean error, mean squared error, correlation between estimates and error, and mean kriging variance should be as small as possible. Reduced kriging variance and correlation between estimated and actual values should be as close to unity as possible. After a trial and error process of the cross validation, a variogram model with the best summary statistics was chosen. The parameters of this model, such as range, sill, and nugget, describe the spatial distribution of a soil/crop characteristic within a field.

3. Results and discussion

Areas with low yield, disease, and stressed vegetation in 1999 are clearly outlined even on the 1997 CIR imagery (Fig. 1a). The most affected area seems to be located over an old canal that can be detected on 1951 aerial photography (Fig. 1d). In 1993 the canal was filled with soil, the fields were reshaped and leveled, and cranberries were planted on five beds (Fig. 1a–c). Data indicate that the healthy field (N5) yielded denser crop after 4–5 years of production (Fig. 2a). Yield distribution within

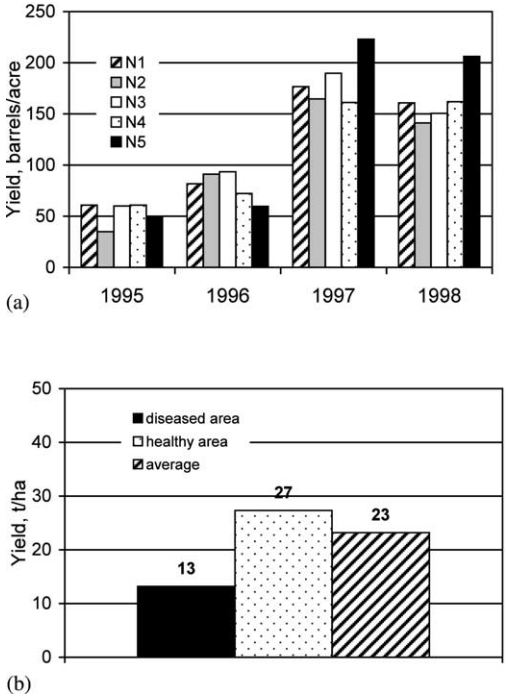


Fig. 2. Cranberry yield in study area: (a) average annual yield from the fields and (b) estimated 1999 yield from the diseased and healthy areas on N3 field. Fields were planted in 1992 and first harvested in 1995.

N3 field suggests considerable reduction of yield on the areas affected by PRR (Fig. 2b and Fig. 3b).

Table 1 lists the descriptive statistics of the data collected in the N3 field at the GPS referenced points and spectral data derived for the same locations from the CIR aerial photography for May and July images (averaged over 9 pixels). All of the variables, with the exception of infiltration rate, were normally distributed.

Correlation coefficients between selected variables are shown in Table 2. The vine density and disease data were kriged to the sampled locations and estimated values rather than 'unity-or-zero' observed values were used for regression analyses. The correlations for spectral data derived from the July image with yield or vine density are about 50% less than those for spectral data from the May image. This is likely due to weed growth, which is difficult to distinguish from cranberry tissue using the CIR photography. Correlation coefficient between PRR symptoms and NIR reaches -0.36 for July image. Correlation between the spectral properties and soil infiltration rate is higher for the July image reaching 0.78 for the Red band and -0.67 for NDVI. Correlation between water content of soil surface and NDVI is increased to 0.34 for the July image. Thus, the spring images of cranberry fields provide more information on future crop data than the late summer images, which tend to highlight more soil water conditions. This can be explained by specifics of

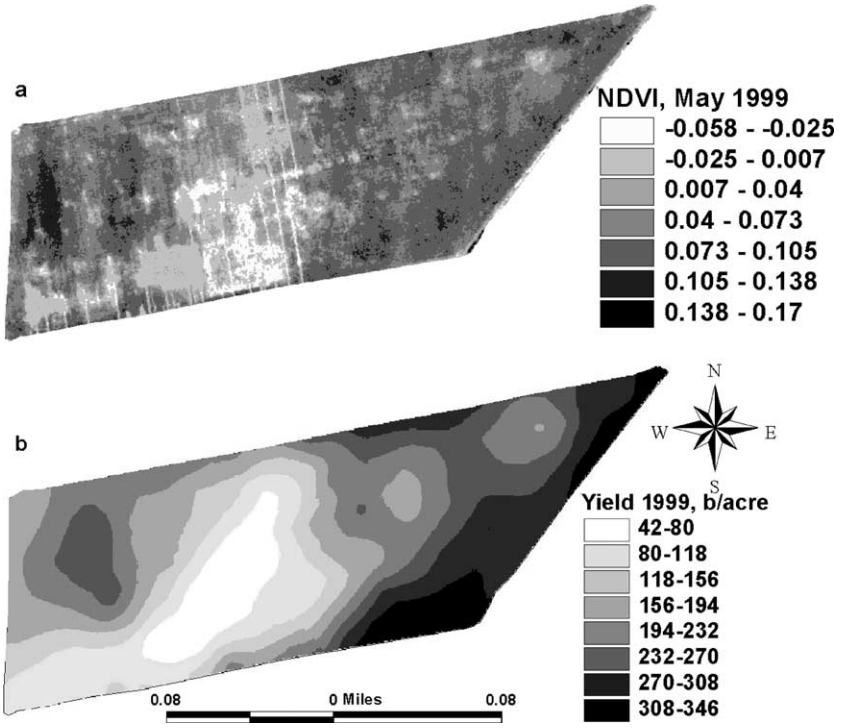


Fig. 3. NDVI image derived from 21 May CIR (a) and cranberry yield map (b). Vertical strips are reflections of underground drainage pipes. (Illustrations in colour on line in Science Direct.)

Table 1
Summary statistics for the experimental data

Variable	*N*	Mean	Median	SD[a]	Skewness	Kurtosis
Yield (t/ha)	216	20.24	19.25	15.95	0.51	−0.42
Soil water content (g/g)	166	0.18	0.17	0.11	0.60	0.55
Elevation (m)	214	0.05	0.05	0.02	0.91	4.28
Infiltration rate (cm/min)	166	0.59	0.34	0.75	2.90	9.26
NIR (May)	212	179.75	175.61	13.99	2.61	6.75
Red (May)	212	162.92	153.97	21.21	1.86	2.90
Green (May)	212	160.27	150.87	21.11	1.92	3.05
NDVI (May)	212	0.05	0.06	0.03	−0.51	−0.91
NIR (July)	212	192.86	194.50	14.70	−0.79	3.39
Red (July)	212	115.41	111.50	29.26	2.30	6.39
Green (July)	212	168.84	166.86	16.49	2.02	5.31
NDVI (July)	212	0.26	0.27	0.08	−1.19	3.72

[a] SD: standard deviation.

Table 2
Correlation matrix for the studied properties

	Disease	Vine density	Elevation	Infiltration	Water content	NIR	Red	Green	NDVI
						May/July			
Yield	−0.12	0.48*	0.34*	−0.20*	0.12	−0.38*/0.17*	−0.59*/−0.20*	−0.58*/−0.08	0.65*/0.26*
PRR symptoms	1	−0.48*	−0.40*	−0.31*	0.14	−0.27*/−0.36*	−0.09/−0.10	−0.12/−0.19*	−0.17*/−0.07
Vine density		1	0.27*	−0.13	0.08	−0.24*/0.15*	−0.45*/−0.24*	−0.42*/−0.13	0.55*/0.36*
Elevation			1	0.35*	−0.26*	0.16*/0.40*	−0.09/0.19*	−0.07/0.27*	0.37*/−0.02
Infiltration				1	−0.38*	0.67*/0.61*	0.55*/0.78*	0.56*/0.76*	−0.19*/−0.67*
Water content					1	−0.34*/−0.26*	−0.26*/−0.37*	−0.26*/−0.32*	0.06/0.34*
May/July									
NIR						(0.55*)	0.88*/0.57*	0.89*/0.72*	−0.42*/−0.63*
Red						(0.24*)	(0.63*)	1.00*/0.96*	−0.80*/−0.94*
Green						(0.72*)	(0.54*)	(0.55*)	−0.78*/−0.85*
NDVI						(−0.63*)	(−0.62*)	(−0.61*)	(0.39*)

For spectral data correlation coefficients of property in the left column with May (on first row)/July (on second row) imagery provided for each property. Values in parentheses are correlation coefficients between spectral values of May and July images. *, marked correlations are significant at $P < 0.05$.

cranberry growth and management. In spring the fields are saturated with water and bear no weeds, therefore, the signature of healthy cranberry vines appears clear on a relatively uniform saturated soil background. In July, the soil water content was more variable due mostly to differences in soil infiltration characteristics while, at the same time, the cranberry canopy signature is masked by weed growth. In July–August the areas with low infiltration are wetter and prone to *Phytophthora* infection. Fig. 4a is a gray image of the Red band and shows areas of exposed sand with no cranberry vines in black. Wet areas with stressed cranberries expressed in dark gray (Fig. 4a) and those areas generally correspond to the areas with low infiltration rate (Fig. 4b).

Multiple regression equations were developed for the relationships among variables. After a trial-and-error process of multiple regression we considered that yield can be best estimated ($r = 0.65$) from the single linear regression with NDVI derived from May image smoothed over 3×3 pixels (Fig. 3a). The infiltration rate is best estimated ($r = 0.79$) from Red band values extracted from the July image with 3×3-pixel smoothing (Fig. 4a). The effect of PRR can be evaluated ($r = 0.53$) with three variables, i.e. elevation, infiltration rate, and NDVI derived from the July

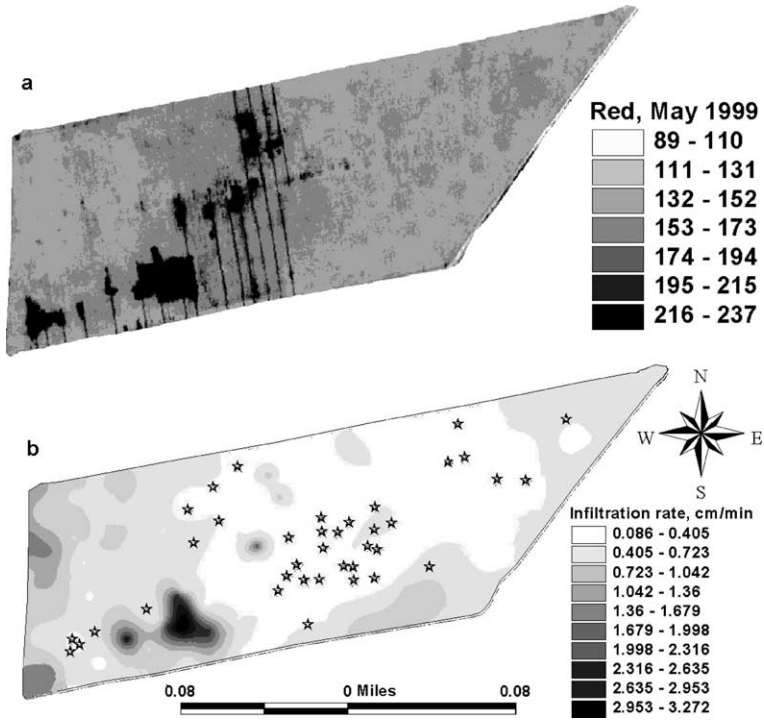


Fig. 4. Black-and-White representation of the Red Band from 5 July CIR (a) and map of infiltration rate with points of PRR symptoms appearance (b). (Illustrations in colour on line in Science Direct.)

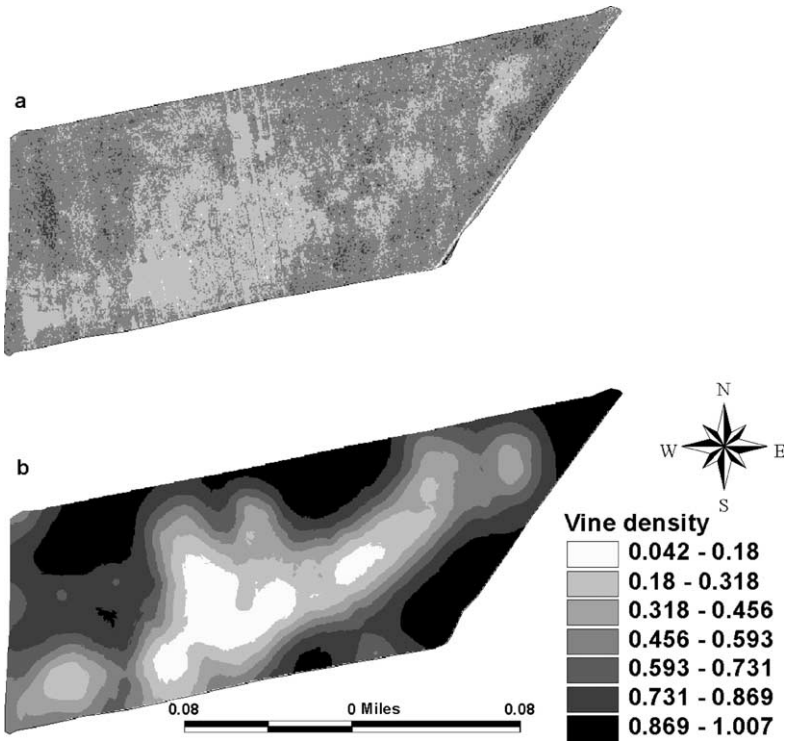


Fig. 5. Image calculated from Red and Green Bands of 21 May CIR with the multiple linear regression equation selected as having best correlation with vine density (a) and map of vine density (b). (Illustrations in colour on line in Science Direct.)

image. Since the infiltration rate is well correlated with the Red band from the July image, the time-consuming measurements of infiltration rate in the field can be eliminated or the number of sampled locations can be significantly reduced. The multiple linear regression analysis was conducted for the vine density and spectral properties. The best correlation ($r = 0.65$) was found with Red and Green bands sampled from 3×3-pixel smoothed May image (Fig. 5a). Fig. 5a shows a gray image of a single layer calculated from Red and Green bands in each pixel using derived best multiple regression equation, which correspond to the map of vine density (Fig. 5b).

In order to develop maps (Fig. 3b, Fig. 4b, Fig. 5b) from soil and crop properties measured at 212 point locations those data were subjected to geostatistical analyses. Calculations of omnidirectional variograms were conducted using various lag-distances until the spatial structures were clear for all measured properties. The calculation of an omnidirectional variograms does not assume that the spatial continuity is the same in all directions (Isaaks and Srivastava, 1989), but this approach seemed useful for our data set with a relatively limited number of sampling

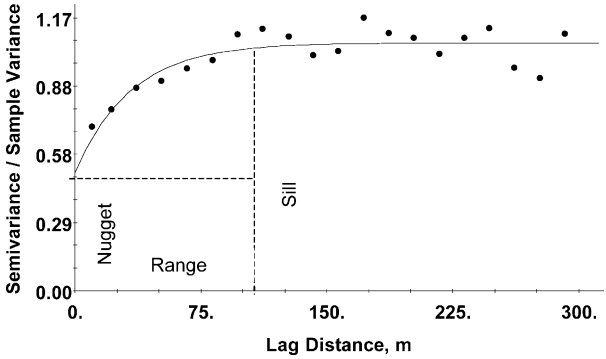


Fig. 6. Omnidirectional variogram for cranberry yield. Separation distance is 10 m with 5 m lag tolerance.

points (2 1 2) and the fact that the field extended further in East–West than in North–South direction (Fig. 3). The data generally obeyed strong spatial auto-correlations that were well described with either spherical or exponential variogram models for omnidirectional calculations (Fig. 6). The directional variograms generally showed the same spherical or exponential patterns for the studied properties, if they included enough data points to construct a valid variogram.

The parameters for omnidirectional variogram model for each property were chosen through cross validation and included range, sill, and nugget (Fig. 6). Range is a distance at which the exponential or spherical variogram models reach a plateau. Range characterizes the extent of auto-correlation in the data, i.e. samples separated by the distance less then a range are related to each other, whereas samples at the lag-distance further apart than a range are statistically independent. The auto-correlation extended within the ranges of 90 m for yield, 54 m for disease, 68 m for vine density, 44–120 m for various spectral properties, 50 m for elevation, 90 m for water content, and 10 m for infiltration rate. The soil and crop data were kriged using the obtained variogram for each property to estimate values at the non-sampled locations and to develop surface maps (Fig. 3b, Fig. 4b, Fig. 5b). Thus, the geostatistical technique of kriging was successfully applied to study relationships between sparse ground-sampled data in vector format and the spectral information from the imagery in raster format.

4. Conclusions

PRR decreases the vine density and considerably diminishes cranberry yield. The spatial patterns of stressed cranberry vegetation were consistent through 5 years as was evident from the analysis of historical CIR aerial photography of the fields. The disease appeared in surface depressions with low infiltration rates, which were subjected to excess water accumulation later in the growing season. This suggests that the groundwater table or soil texture may also contribute to disease development. The spatial analyses and monitoring of these soil properties as well

as water content, soil–water potential, soil resistance to penetration, and humus content, will be conducted in the future. While the areas with chronic *Phytophthora* infestation were fairly stable from year to year and could be outlined by the low infiltration rate, the areas with acute disease (from root isolation data) were moving within the areas prone to infestation. Acute disease outbreaks most often resulted in partial or complete dieback of the vines, leaving areas of bare soil, evident from early-season imagery. Thus, the areas subjected to chronic PRR can be outlined on CIR photography by analyzing the spectral reflectance in the Red band from later summer CIR photographs, which also correlates with the soil infiltration rate. The effect of chronic and acute diseases on the yield and vine density is better estimated from the NDVI from spring images.

The quantification of qualitative 'unity-or-zero' ground data on vine density and *Phytophthora* appearance by kriging provided relatively accurate surface maps of these properties, which spatially match the features found on CIR photographs. The study suggests that scouting in combination with geostatistics, and timely CIR photography can provide essential information on spatial distributions of vine density, yield, and PRR within the cranberry fields. The CIR photography can be utilized in cranberry production to estimate and predict yield for each field within the area, to study the disease spatial distributions on the fields for the previous years, to evaluate areas with low infiltration prone to disease, and to predict yield distributions within a field from the early-season aerial photography. The work is being done to develop a system that will allow cranberry growers to access data and maps from this and other station's research on the Internet and farmers are being trained to use ArcExplorer 1.1 (ERSI Inc., 1999) to view spatial information.

Acknowledgements

The authors thank Pine Island Cranberries, Inc. and William Haines for providing unlimited access to cranberry beds, equipment, and harvest information. The authors also thank Lawrence Dapsis, Joseph Daverna, and Lynn DiBiase of Ocean Spray Cranberries, Inc. and Joan Davenport of Washington State University, for providing access to data on cranberry yield. Funding for this project was provided through USDA-NRI, USDA-CSREES, The New Jersey Blueberry and Cranberry Research Council, and Ocean Spray Cranberries, Inc.

References

Carr, J.R., Myers, D.E., 1984. Applications of the theory of regionalized random variables to the spatial analysis of Landsat data. IEEE Computer Society Proceedings of the Pecora, 9 Spatial Technologies for Remote Sensing Today and Tomorrow, IEE Computer Society Press, Silver Spring, MD, pp. 55–61.

Clevers, J.G.P.W., 1997. A simplified approach for yield prediction of sugar beet based on optical remote sensing. Remote Sensing and Environment 62, 171–179.

Dungan, J., 1998. Spatial prediction of vegetation quantities using ground and image data. International Journal of Remote Sensing 19 (2), 267–285.

Everitt, J.H., Escobar, D.E., Appel, D.N., Riggs, W.G., Davis, M.R., 1999. Using airborne digital imagery for detecting oak wilt disease. Plant Disease 83, 502–505.

Huete, A.R., 1988. A soil-adjusted vegetation index (SAVI). Remote Sensing of Environment 25, 295–309.

Hughes, M.G., Oudemans, P.V., Davenport, J.R., Ayres, K., Airola, T.M., Lee, A., 1998. Evaluating commercial cranberry beds for variability and yield using remote sensing techniques. Precision Agriculture. Proceedings of the Fourth International Conference. American Society of Agronomy, St. Paul, MN, Part B, pp. 1493–1500.

Isaaks, E.H., Srivastava, R.M., 1989. An Introduction to Applied Geostatistics. Oxford University Press, New York, Oxford.

Johnson, L., Lobitz, B., Bosch, D., Wiechers, S., Williams, D., Skinner, P., 1998. Of pixels and palates: can geospatial technologies help produce a better wine? First International Conference on Spatial Information in Agriculture and Forestry. Lake Buena Vista, FL.

Klute, A. (Ed.), 1986. Methods of Soil Analysis, Part 1. Physical and mineralogical methods. Second Edition. ASA, SSSA, Madison, Wisconsin.

Maas, S.J., 1991. Within-season calibration of modeled wheat growth using remote sensing and field sampling. Agronomy Journal 85, 669–672.

Nelson, M.R., Orum, T.V., Jaime-Garcia, R., Nadee, A., 1999. Applications of geographic information systems and geostatistics in plant disease epidemiology and management. Plant Disease 83, 308–319.

Pozdnyakova, L.A., Zhang, R., 1999. Estimation of spatial variability of soil salinity using geostatistical methods. Fourth International Conference on Precision Agriculture, St. Paul, MN, Part A: pp. 79–90.

Ristaino, J.B., Larkin, R.P., Campbell, C.L., 1993. Spatial and temporal dynamics of Phytophthora epidemics in commercial bell pepper fields. Phytopatology 83, 1312–1320.

Stein, A., van der Meer, F., Gorte, B. (Eds.), 1999. Spatial statistics for remote sensing. Remote sensing and digital image processing. Kluwer Academic Publishers, Dordrecht/Boston/London.

Tabor, J.A., Warrick, A.W., Pennington, D.A., Myers, D.E., 1984. Spatial variability of nitrate in irrigated cotton: I. Petioles. Soil Science Society of American Journal 48, 602–607.

Tabor, J.A., Warrick, A.W., Myers, D.E., Pennington, D.A., 1985. Spatial variability of nitrate in irrigated cotton: II. Soil nitrate and correlated variables. Soil Science Society of American Journal 49, 390–394.

Wu, B.M., van Bruggen, A.H.C., Subbarao, K.V., Pennings, G.G.H., 2001. Spatial analysis of lettice downy mildew using geostatistics and geographic information systems. Phytopathology 91, 134–142.

Zhang, R., Warrick, A.W., Myers, D.E., 1992. Improvement of the prediction of soil particle size fractions using spectral properties. Geoderma 52, 223–234.

Precision Agriculture, 1, 153–165 (1999)

Geostatistical Analyses of Soil Salinity in a Large Field

LARISA POZDNYAKOVA AND RENDUO ZHANG
Department of Renewable Resources, University of Wyoming, Laramie, WY 82071-3354

Abstract. Estimating spatial variability of soil salinity is an important issue in precision agriculture. Geostatistical methods provide a means to study the heterogeneous nature of spatial distributions of soil salinity. In this study, geostatistical methods, kriging and cokriging, were applied to estimate sodium adsorption ratio (SAR) in a 3375 ha agricultural field. In cokriging, more easily measured data of electrical conductivity (EC) were incorporated to improve the estimation of SAR. The estimated spatial distributions of SAR using the geostatistical methods with various reduced data sets were compared with the extensive salinity measurements in the large field. The results suggest that sampling cost can be dramatically reduced and estimation can be significantly improved using cokriging. Compared with the kriging results using total SAR data, cokriging with reduced data sets of SAR improves the estimations greatly by reducing mean squared error and kriging variance up to 70% and increasing correlation of estimates and measurements about 60%. The sampling costs for SAR estimation can be reduced approximately by 80% using extensive EC data together with a small portion of SAR data in cokriging.

Keywords: soil salinity, spatial variability, geostatistical analyses

Introduction

Precision agriculture practices in arid and semiarid areas require periodic information on soil salinity. Excessive soil salinity may cause crop loss and, eventually, land degradation (Lesch *et al*., 1992). Soil salinity problems appear worldwide during irrigation practices (Ballantyne, 1963; Bettenay *et al*., 1964; Greenlee *et al*., 1968; Halvorson and Rhoades, 1976). High saline areas often distribute randomly within non-saline and low saline irrigation fields, therefore, receive the same inputs of tillage, water, fertilizer, and seed as non-saline soils, even though no crop is produced. Soluble salts affect the productivity of soils in two principal ways: changing the osmotic potential of soil solution and increasing the content of exchangeable sodium, which produces in most soils an unfavorable physical conditions (Richards *et al*., 1956). The attempts are made to remediate high saline areas by biological methods and by site-specific irrigation (Szabolcs, 1989; Ismail *et al*., 1991; Malcolm, 1991; Mankin *et al*., 1997). Excess salinity of soil solution can be corrected by leaching with water of good quality. The removal of excess exchangeable sodium requires an application of gypsum amendment. Therefore, development and maintenance of reclamation and irrigation projects on saline soils require accurate and updated information about spatial distributions of soil soluble salts, especially the exchangeable sodium.

Assessing soil salinity is complicated by the nature of its spatial and temporal variability (Rhoades and Ingvalson, 1971). Conducting soil salinity measurements at high sampling density is costly and time-consuming. Fortunately, it is possible to use quick in-situ methods of electrical conductivity (EC), which is related to soil salinity, to evaluate salinity. The relationship between EC and soil salinity is complicated by other factors influencing measured conductivity in the field, such as soil texture, water content, and bulk density (Rhoades *et al*., 1976; Banton *et al*., 1997). Thus, in situ measurements of electrical conductivity require calibration for a certain field (case) to be suitable to monitor and map soil salinity. Such calibration usually is conducted using common statistical methods of correlation and regression (Halvorson and Rhoades, 1976; Chang *et al*., 1983; Rhoades *et al*., 1989b).

Geostatistical methods, kriging and cokriging, are becoming commonly used estimation techniques to generate soil maps. Kriging has been applied to quantify variability of various spatial variables in soil science. For example, Tabor *et al*. (1984, 1985) used variograms and kriging to determine the spatial variability of nitrates in cotton petioles and analyzed spatial variability of soil nitrate and correlated variables. Istok and Cooper (1988), Cooper and Istok (1988 a,b) applied kriging to study groundwater contamination. Yates *et al*. (1993) used geostatistics in the description of salt-affected soils. Samra and Gill (1993) used kriged results to assess variations of pH and sodium adsorption ratio associated with tree growth on a sodium-contaminated soil. Using disjunctive kriging, Yates *et al*. (1986 a,b) presented spatial distributions and corresponding conditional probability maps of soil electrical conductivity.

Whereas kriging is used to evaluate the spatial distribution of a variable based on sampled data of the variable itself, cokriging is applied to improve estimation of undersampled variables using abundant data of other variables. Yates and Warrick (1987) applied cokriging to estimate soil water content with the auxiliary data of bare soil surface temperature and sand content. Using kriging and cokriging, Zhang *et al*. (1995) estimated trace elements in soils and plants. Zhang *et al*. (1992a, 1997) used cokriging with pseudo-cross-variograms to estimate spatial distributions of soil chemicals.

Geostatistical methods can be powerful tools for characterizing large-scale spatial distributions of soil properties for precision agriculture. In this study, kriging and cokriging were utilized to estimate spatial distributions of soil salinity and sampling strategies. We used cokriging and electrical conductivity data to improve the estimation of sodium adsorption ratio (SAR). The estimated results were compared with extended salinity measurements in a large agricultural landscape.

Materials and methods

To determine the spatial distribution of salinity, a 3375 ha area of irrigated farmland within South-Fork Kings River Watershed in central California was investigated in summer of 1987 (Rhoades *et al*., 1988). The study area is located in

northern Kings County, east of the Kings River, and in the LeMoore, Hanford, Guernsey, and Stratford quadrangles. The site was selected because irrigated agriculture predominates and a cross-section of agricultural crops and management practices is well represented within the area. Well-managed large corporate farming units, small family farms, and abandoned parcels of land occur randomly within the area. The area also contained a wide distribution of soil types with texture ranging from loam sand to clay (Rhoades *et al*., 1989a). The water table in the area was located at a depth ranging from less than one to greater than three meters below the soil surface. The whole area was divided with a 200 × 200 m grid. Within each grid one soil sample was taken at random resulting in total 898 soil samples. To evaluate salinity, sodium adsorption ratio (SAR) was measured in samples following conventional methods (U.S. Salinity Laboratory Staff, 1954; Munk, 1992). In-situ electrical conductivity (EC) data were collected at the same locations using the four-electrode technique (Rhoades *et al*., 1990). Statistical and geostatistical analyses of soil salinity were conducted using all 898 data points of EC and various reduced data sets randomly selected from the total 898 points of SAR. The data were analyzed using descriptive statistics (STATISTICA, Statsoft Inc., 1993) and geostatistical methods of kriging and cokriging (GSLIB, Stanford University, 1997).

The ordinary kriging estimator, $Z^*(x_0)$, of an unsampled site is a linear sum of weighted observations within a neighborhood:

$$Z^*(x_0) = \sum_{i=1}^{n} \lambda_i Z(x_i) \qquad [1]$$

where $Z^*(x_0)$ is the estimate of Z at x_0, λ_i is the weight assigned to the ith observation, and n is the number of observations within the neighborhood. The weighing factors of λ_i's are determined based on a variogram of Z. Cokriging is a method for estimating one or more variables of interest using data from several variables by incorporating not only spatial distribution, but also intervariable correlation. A cokriging estimate is a weighted average of available data with weights chosen so that the estimate is unbiased and has minimum variance, analogous to ordinary kriging (Myers, 1982, 1984).

Cross-validation is used to evaluate the accuracy of the variogram and cross-variogram models for kriging and cokriging. In this procedure, every known point is estimated using the values at the neighborhood around it, but not itself. The summary statistics from the cross validation includes mean error, mean squared error, mean kriging variance, reduced kriging variance, correlation between estimates and error, and correlation between estimated and actual values (Isaaks and Srivastava, 1989). Mean error, mean squared error, correlation between estimates and error, and mean kriging variance should be as small as possible. Reduced kriging variance and correlation between estimated and actual values should be as close to unit as possible. After trial and error process of the cross validation a variogram (cross-variogram) model with the best summary statistics is chosen.

The following methods are used to compare results estimated with kriging and cokriging. Relative improvement, or relative reduction of estimation accuracy, is

defined by

$$R_E = 100\%(|MSE_R| - |MSE_E|)/|MSE_R| \quad [2]$$

where MSE_E and MSE_R are the mean squared errors for evaluated and reference methods, respectively. For example, comparison between the estimation accuracy of cokriging using a reduced data set of SAR and the EC data versus the accuracy of kriging using the total SAR data is obtained using the mean squared error of cokriging (MSE_E) and the mean squared error of kriging (MSE_R). If R_E is positive, estimation accuracy for the evaluated method is better than the reference method. If R_E is negative, the evaluated method is worse than the reference method. The relative improvement or reduction of accuracy based on the mean kriging variance is defined in the same way by replacing MSE_E and MSE_R in Eq. 2 with respective values of the mean kriging variance. The relative improvement or reduction of correlation between estimated and actual values is defined by

$$R_r = 100\%(|1 - r_R| - |1 - r_E|)/|1 - r_R| \quad [3]$$

where r_E and r_R are the correlation coefficients between estimated and actual values for evaluated and reference methods, respectively.

Results

Table 1 lists the descriptive statistics of the raw and log-transformed EC and SAR data, including mean, median, coefficient of variation (CV), skewness, and kurtosis. The descriptive statistics of the data suggest that both electrical conductivity and sodium adsorption ratio are lognormally distributed variables. Therefore, the log-transformed data of SAR and EC were used for geostatisical analyses. This transformation not only modified variable distribution to normal, but also improved the correlation coefficient (r) between EC and SAR from 0.61 to 0.82. To study advantages of cokriging and sampling strategies, 11 reduced data sets of SAR with 50, 75, 100, 200, 300, 400, 500, 600, 700, 800, and 850 data points were randomly selected from the original data set. Locations of the original data and some of the randomly selected data sets are shown in Fig. 1. All of these reduced data sets are normally distributed after log-transformation and have a correlation coefficient (r) with EC about 0.8.

Table 1. Summary statistics for experimental data.

Variable	N	Mean	Median	CV	Skewness	Kurtosis
EC (dS m^{-1})	898	6.675	3.560	1.37	3.318	15.564
log(EC)	898	0.527	0.551	0.98	0.142	−0.846
SAR	898	22.447	7.710	2.35	8.397	96.081
log(SAR)	898	0.916	0.887	0.66	0.188	−0.375

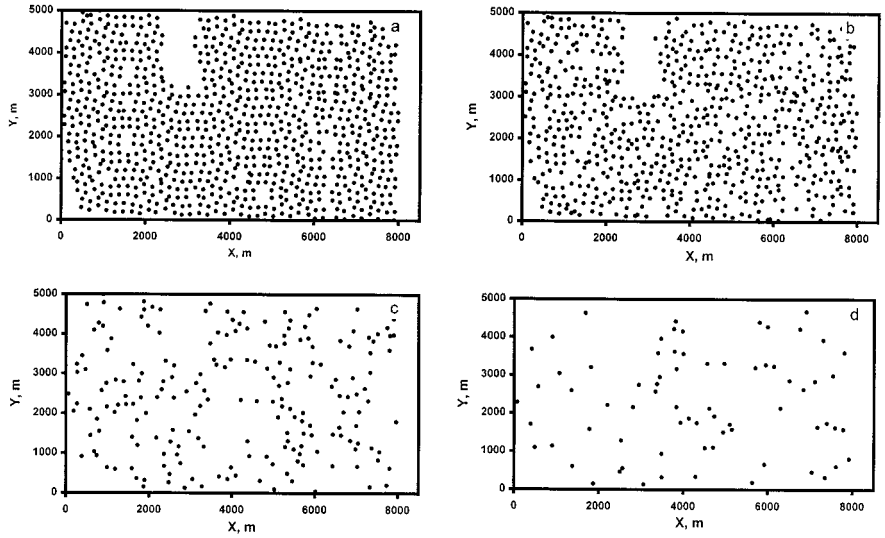


Figure 1. Spatial location of the (a) total 898 and randomly selected (b) 700, (c) 200, and (d) 100 data points of SAR.

Variograms and cross-variograms were calculated for the data sets. Preliminary calculations of variograms in different directions showed that all variograms were isotropic. Therefore, the omnidirectional variograms were used for analyses. Sample variograms of all the data sets show that the spatial structures of SAR and EC can be well characterized by the spherical model:

$$\gamma(h) = \begin{cases} C_0 + C_1\left[1.5(h/a) - 0.5(h/a)^3\right] & h \le a \\ C_0 + C_1 & h > a \end{cases} \quad [4]$$

where h is the lag distance, C_0 is the nugget, $C_0 + C_1$ is the sill, and a is the range. The parameters of variogram models for each data set were chosen through cross validation. The parameters of the models for the different data sets of SAR were very similar (Table 2 and Fig. 2). Therefore, an average spherical model with range, nugget, and sill equal to 700, 0.14, and 0.33, respectively, was used for all the data sets of SAR. These parameters were obtained based on the cross validation for the 100-point data set of SAR (Table 2 and Fig. 2b) and used for each data set of SAR in kriging and cokriging. The summary statistics of cross validation using this variogram model with the different reduced data sets are listed in Table 3.

Cokriging was conducted for each reduced data set of SAR based on the three variograms: the variogram for the reduced data set of SAR, a variogram for total EC data, and a variogram of the sum of SAR and EC (EC + SAR) at the common locations (Zhang *et al.*, 1997). The variogram of the EC data was also spherical with nugget, sill, and range equal to 0.14, 0.26, and 700, respectively (Table 2). Using the same variogram model for the different reduced data sets of SAR and

Table 2. Parameters of variogram models.

Variable, number of data points	Variogram			Variogram of sum EC + SAR		
	Nugget	Sill	Range (m)	Nugget	Sill	Range (m)
SAR, 100	0.14	0.33	700	0.57	1.1	950
SAR, 200	0.16	0.35	700	0.58	1.1	850
SAR, 300	0.14	0.33	800	0.57	1.1	950
SAR, 400	0.15	0.33	700	0.57	1.1	1000
SAR, 500	0.14	0.32	700	0.57	1.1	950
SAR, 600	0.15	0.32	500	0.57	1.1	950
SAR, 700	0.14	0.34	600	0.57	1.1	950
SAR, 800	0.14	0.33	700	0.56	1.1	900
SAR, 850	0.14	0.33	700	0.57	1.1	950
SAR, 898	0.14	0.33	700	0.57	1.1	950
EC, 898	0.14	0.26	700			

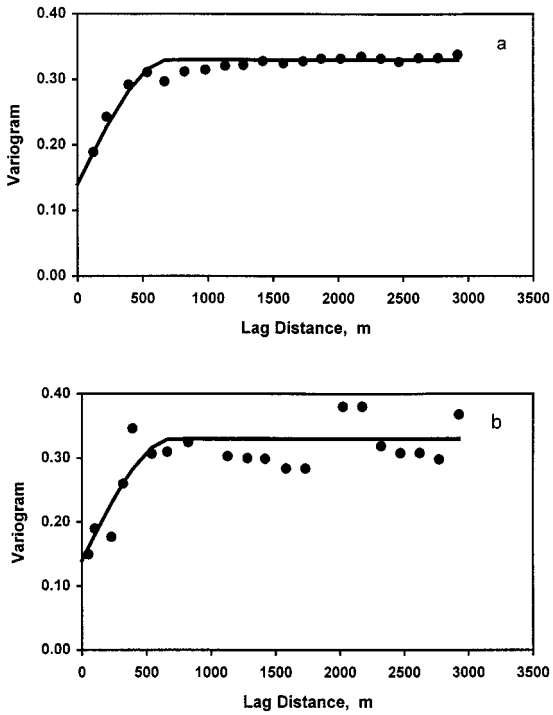


Figure 2. Sample variogram (circles) and model (solid line) of the (a) total 898 and (b) 100 randomly selected SAR data.

Table 3. Summary statistics of kriging for different data sets.

Variable, number of data points	Statistics*					
	ME	MSE	MKV	RKV	r_e	r_a
SAR, 50	0.039	0.349	0.331	1.055	0.115	0.208
SAR, 75	0.030	0.364	0.318	1.173	0.116	0.234
SAR, 100	0.010	0.303	0.328	0.903	0.139	0.318
SAR, 200	0.008	0.371	0.291	1.269	0.135	0.383
SAR, 300	−0.021	0.316	0.307	1.042	0.286	0.460
SAR, 400	0.000	0.286	0.271	1.059	0.180	0.442
SAR, 500	−0.010	0.259	0.261	0.990	0.149	0.490
SAR, 600	−0.005	0.287	0.253	1.138	0.151	0.48
SAR, 700	−0.001	0.275	0.247	1.114	0.117	0.532
SAR, 800	0.000	0.263	0.242	1.085	0.095	0.539
SAR, 850	0.000	0.259	0.237	1.095	0.072	0.550
SAR, 898	−0.001	0.257	0.235	1.096	0.060	0.549
EC, 898	0.000	0.189	0.206	0.918	0.022	0.538

*ME is mean error. MSE is mean squared error. MKV is mean kriging variance. RKV is reduced kriging variance, which is MKV divided by MSE. r_e is correlation coefficient between estimated value and error. r_a is correlation coefficient between estimated and actual values.

the model of EC, variograms of EC + SAR for each data set were selected through cross validation. The variogram models of EC + SAR for the reduced data sets were also spherical and have slightly variable parameters (Table 2). Therefore, a spherical variogram model of EC + SAR with nugget, sill, and range equal to 0.57, 1.1, and 950, respectively, was applied for conducting cokriging with each reduced data set of SAR and the whole data of EC. The cross-variogram was calculated using (Zhang *et al*., 1992b):

$$\gamma_{12} = 0.5(\gamma_{12}^{+} - \gamma_{11} - \gamma_{22}) \qquad [5]$$

where γ_{12} is the cross-variogram of EC and SAR, γ_{11} is the variogram for SAR, γ_{22} is the variogram for EC, and γ_{12}^{+} is the variogram for EC + SAR. The cross-variogram of EC and SAR satisfies the Cauchy-Schwartz inequality (Myers, 1982; Zhang *et al*., 1997). This condition guarantees that the cokriging coefficient matrix is positive definite and the variance of the estimated variables is positive. The summary statistics for cokriging results with the different reduced data sets of SAR are shown in Table 4.

The summary statistics for kriging (Table 3) and cokriging (Tables 4) show that cokriging provides much better estimation results for SAR than kriging. On average, for kriging of the reduced data sets the mean squared error is 0.276, the mean kriging variance is 0.255, the correlation coefficient of estimation and error is 0.124, and the correlation coefficient of estimated and actual values is 0.399 (Table 3). For cokriging, the same average statistics are 0.100, 0.099, 0.042, and 0.770, respectively (Table 4). The accuracy of kriging and cokriging increases with increasing number of the used data points. The mean kriging variance decreases from 0.331 to 0.237 when the number of data points used for kriging increases from

Table 4. Summary statistics of cokriging for different data sets of SAR and total 898 EC data

Variable, number of data points	Statistics*					
	ME	MSE	MKV	RKV	r_e	r_a
SAR, 50	0.057	0.143	0.176	0.860	0.230	0.794
SAR, 75	0.055	0.125	0.155	0.837	0.116	0.826
SAR, 100	0.003	0.117	0.143	0.827	0.008	0.804
SAR, 200	0.009	0.153	0.119	1.334	0.010	0.802
SAR, 300	−0.007	0.123	0.115	1.091	0.009	0.818
SAR, 400	−0.002	0.097	0.098	1.102	0.008	0.848
SAR, 500	−0.003	0.080	0.088	0.963	0.038	0.872
SAR, 600	−0.002	0.094	0.080	1.214	0.018	0.861
SAR, 700	−0.002	0.090	0.075	1.210	0.008	0.874
SAR, 800	−0.001	0.090	0.070	1.303	0.028	0.869
SAR, 850	0.000	0.090	0.068	1.345	0.034	0.871

*Abbreviations as in Table 3.

50 to 850 (Table 3). For cokriging, this statistic decreases more quickly from 0.176 to 0.068 for the same range of data points (Table 4). The correlation between estimated and actual values only increases slightly with number of data points for cokriging, and shows relatively stable increase for kriging (Tables 3 and 4).

We calculated the relative improvement or reduction in estimation accuracy based on some kriging and cokriging statistics for the different reduced data sets of SAR (Eqs. 2 and 3). Kriging with all 898 SAR data is used as a reference method both for kriging and cokriging (Fig. 3, the dashed line). Whereas kriging conducted with any reduced data sets of SAR shows a reduction in accuracy (negative values) for all statistics compared with kriging using all SAR data, cokriging reveals stable improvement in prediction accuracy (Fig. 3a, b, c). Cokriging with only 50 SAR values and all 898 EC data gave even better estimation than kriging with 898 SAR data. The relative improvement in the mean squared error and mean kriging variance increases with the number of data points and shows a 40 to 70% improvement for cokriging of data sets with 100 or more SAR measurements and all 898 EC data compared with kriging of all SAR data (Fig. 3a, b). The improvement in correlation between estimated and actual values for cokriging is better than 60% for any reduced data set of SAR. The correlation only improves slightly with increasing number of SAR data (Fig. 3c). The vertical distance between the best fitting lines in Fig. 3 is used to show the improvement in accuracy of cokriging versus kriging for the same data sets of SAR. The improvement in estimation accuracy based on the mean cokriging variance or the mean squared error does not vary much with the number of data points and ranges from 60 to 80% (Fig. 3a, b). The improvement in estimation accuracy based on the correlation between estimated and actual values is up to 140% (Fig. 3c). In summary, all the statistics are highly improved by cokriging using the reduced data sets of SAR and all 898 EC data compared with kriging using all SAR data. The improvement increases slightly as the number of SAR measurements increases. Comparing cokriging results with kriging using the same reduced data sets of SAR, the improvement in mean

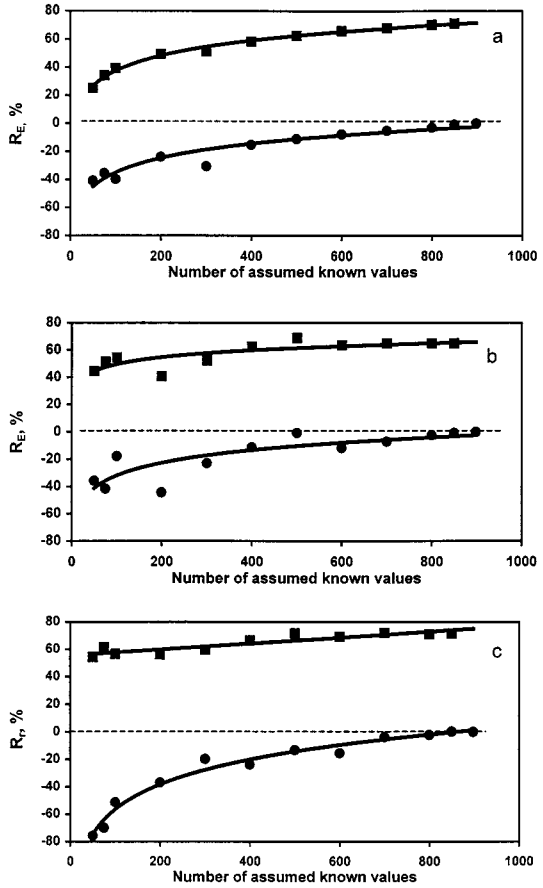


Figure 3. Relative improvement (+) or reduction (−) of estimation accuracy based on (a) mean kriging variance, (b) mean squared error, and (c) correlation of estimated and actual values for kriging (circles) and cokriging (squares) using randomly selected data sets of SAR, compared with kriging using the total SAR data.

squared error and mean kriging variance remains almost constant for any data set. Therefore, increasing the number of SAR measurements does not significantly influence the accuracy of estimation with cokriging.

Spatial patterns of SAR estimated by different mapping strategies are shown in Fig. 4. Although almost all the territory is saline with EC of saturated soil extract more than 2 dS m^{-2}, only small area is with high exchangeable sodium (Fig. 4a). The high SAR area is relatively well delineated by cokriging using 200 SAR data with the total EC data (Fig. 4b). The cokriging results adequately represent both the spatial patterns and range of SAR values. Kriging with 200 SAR data tends to overestimate the areas with SAR greater than 60 and spread the areas with SAR from 15 to 30 on the non-sodic area (Fig. 4c). Both kriging and cokriging methods using 100 SAR data failed to outline the high SAR area adequately (Fig. 4d). Thus,

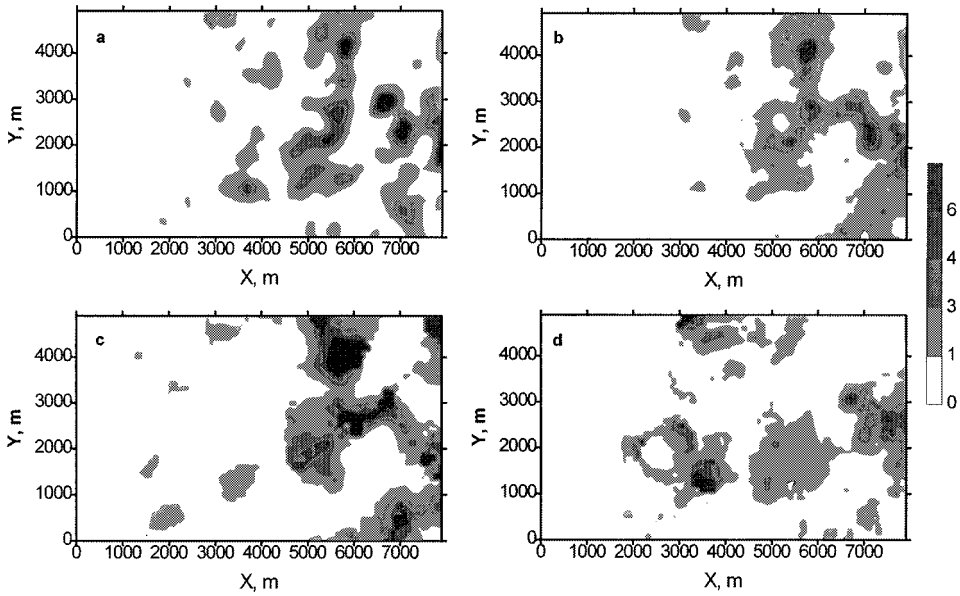


Figure 4. Contour maps of SAR estimated by (a) kriging with 898 SAR data, (b) cokriging with 200 SAR and 898 in-situ EC data, (c) kriging with 200 SAR data, and (d) cokriging with 100 SAR and 898 in-situ EC data.

cokriging using reduced data sets of SAR with 200 and more data points well characterizes the spatial distribution of SAR. Cokriging conducted using data sets of 100 and less SAR data, although still improves estimation accuracy, does not represent the spatial pattern of SAR adequately.

Cokriging with reduced data of SAR and extensive in-situ EC data has been shown to improve estimation of SAR significantly compared with kriging using only extensive SAR data. To conclude that cokriging is a better estimator than kriging, the sampling cost for the methods needs to be considered together with the estimation accuracy. In terms of sampling cost, electrical conductivity is very easy to obtain. Hundreds of the data points can be measured directly in the field during one day without actual sampling and analyzing soil samples in the laboratory. On the other hand, soil adsorption ratio analysis requires soil sampling and laboratory measurement of Ca, Mg and Na in each sample, therefore, is very costly and time consuming. Relative cost of in-situ EC measurements versus SAR analysis is approximately 1:60. The relative improvement in estimation accuracy for cokriging using reduced data sets for SAR and all EC data was compared with the relative sampling cost for the same data sets in Fig. 5. The sampling cost linearly increases with the number of data points. However, the improvement in estimation accuracy is almost constant (60%) for 200 and more data points of SAR. The sampling cost for 200 SAR and 898 EC data is about one forth of the sampling cost for 898 SAR data. Reducing SAR measurements to less than 200 points is not recommended, since it would be difficult to obtain adequate spatial pattern of SAR distribution.

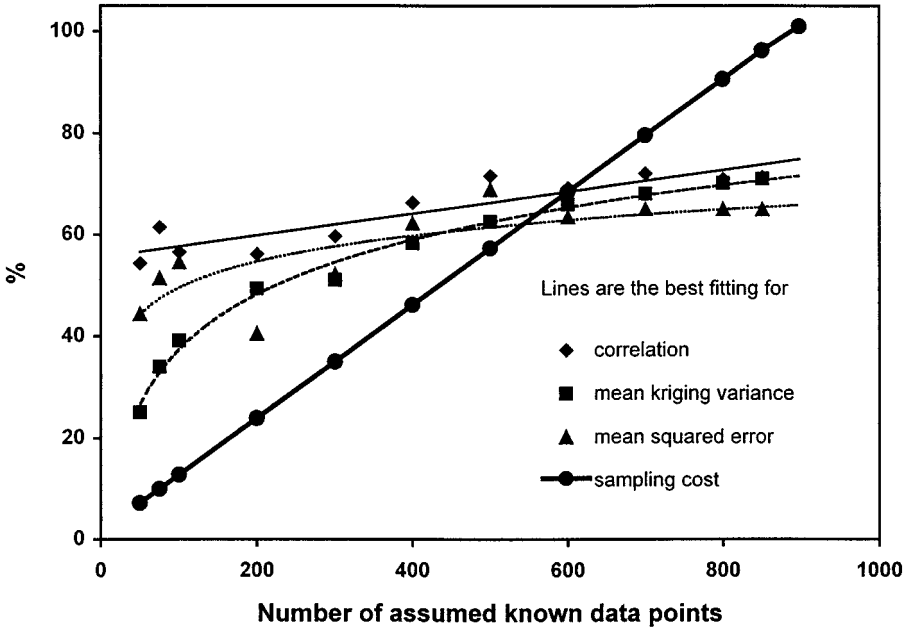


Figure 5. Relative sampling cost and relative improvement of prediction accuracy based on mean kriging variance, mean squared error, and correlation between estimated and actual values for cokriging using various reduced SAR data sets and 898 EC data compared with kriging using the 898 SAR data.

Therefore, the optimal sampling strategy for this example is to measure randomly about 200 locations for SAR and about 4 to 5 times as many locations for EC.

Conclusions

In this study, kriging and cokriging were used to estimate the soil salinity, including electrical conductivity (EC) and sodium adsorption ratio (SAR), over a 3375 ha area of irrigated farmland. Geostatistical analyses were conducted for the reduced data sets containing 6 to 90% of the original SAR data. The analyses show a great potential for reducing sampling costs and increasing prediction accuracy using cokriging. Using only 22% of the SAR data, cokriging improves the prediction of SAR significantly by incorporating the information of EC. Compared with the kriging results using all 898 SAR data, cokriging using 200 SAR and 898 EC data improved the mean squared error and mean kriging variance by 40 to 50%, and the correlation between estimated and actual values by 60%. In addition, the sampling cost was reduced up to 5 times. Cokriging is shown to be an accurate, yet economic, method for evaluation of spatial distributions of soil salinity in large fields.

References

A. K. Ballantyne, Recent accumulation of salts in the soils of southeastern Saskatchewan. *Can. J. Soil Sci.* **43**, 52 (1963).

O. Banton, M.-K. Seguin, and M.-A. Cimon, Mapping field-scale physical properties of soil with electrical resistivity. *Soil Sci. Soc. Am. J.* **61**, 1010 (1997).

E. Bettenay, A. V. Blackmore, and F. J. Hinston, Aspects of the hydrologic cycle and related salinity in the Belka Valley, Western Australia. *Aust. J. Soil Res.* **2**(2), 187 (1964).

S. Chang, T. G. Sommerfeldt, J. M. Carefoot, and G. B. Schaalje, Relationships of electrical conductivity with total dissolved salts and cation concentrations of sulfate-dominant soil extrats. *Can. J. Soil Sci.* **63**, 79 (1983).

R. M. Cooper, and J. D. Istok, Geostatistics applied to groundwater contamination. I: Methodology. *J. Env. Eng.* **144**, 270 (1988a).

R. M. Cooper, and J. D. Istok, Geostatistics applied to groundwater contaminants. II. Application. *J. Env. Eng.* **114**, 287 (1988b).

G. M. Greenlee, S. Pawluk, and W. E. Bowser, Occurrence of soil salinity in dry lands of southwestern Alberta. *Can. J. Soil Sci.* **48**, 65 (1968).

A. D. Halvorson, and J. D. Rhoades, Field mapping soil conductivity to delineate dryland saline seeps with four-electrode technique. *Soil Sci. Soc. Am. J.* **40**, 571 (1976).

E. H. Isaaks, and R. M. Srivastava, *An introduction to Applied Geostatitics*, (Oxford University Press, New York, 1989).

S. Ismail, R. Ahmad, and N. J. Davidson, Design and analysis of provenance trials in Pakistan, in: *Productive use of Saline Land*. ACIAR Proceeding No 42, Perth, Western Australia, 10–14 May 1991, pp. 38–45.

J. D. Istok, and R. M. Cooper, Geostatistics applied to groundwater pollution. III. Global estimates. *J. Env. Eng.* **114**, 915 (1988).

S. M. Lesch, J. D. Rhoades, L. J. Lund, and D. L. Corwin, Mapping soil salinity using calibrated electromagnetic measurements. *Soil Sci. Soc. Am. J.* **56**, 540 (1992).

C. V. Malcolm, The potential of halophytes for rehabilitation of degraded land, in: *Productive Use of Saline Land*, ACIAR Proceeding No 42, Perth, Western Australia, 10–14 May 1991, pp. 8–12.

K. R. Mankin, K. L. Ewing, M. D. Schrock, and G. J. Kluitenberg, Field measurement and mapping of soil salinity in saline seeps, in: *Proceeding for ASAE International Meeting*. No 97 3145, Minneapolis, MN, 10–14 August 1997.

L. P. Munk, Sodium adsorption ratio screening procedure for soils using pH and electrical conductivity. *Soil Sci. Soc. Am. J.* **56**, 1127 (1992).

D. E. Myers, Matrix formation of co-kriging. Math. Geology. **14**, 249 (1982).

D. E. Myers, Co-kriging—new developments, in: *Geostatistics for Natural Resources Characterization. Part 1*, edited by G. Verly *et al.* (D. Reidel Publishing Co., Hingham, MA, 1984), pp. 205–305.

J. D. Rhoades, D. L. Corwin, and P. J. Shouse, Use of instrumental and computer assisted techniques to assess soil salinity, in: *Proceedings of the International Symposium on Solonetz Soils*. Osijek, Yugoslavia, 1988, pp. 50–103.

J. D. Rhoades, and R. D. Ingvalson, Determining salinity in field soils with soil resistance measurements. *Soil Sci. Soc. Amer. Proc.* **35**, 54 (1971).

J. D. Rhoades, N. A. Manteghi, P. J. Shouse, and W. J. Alves, Estimating soil salinity from saturated soil-paste electrical conductivity. *Soil Sci. Soc. Am. J.* **53**, 428 (1989a).

J. D. Rhoades, N. A. Manteghi, P. J. Shouse, and W. J. Alves, Soil electrical conductivity and soil salinity: new formulations and calibrations. *Soil Sci. Soc. Am. J.* **53**, 433 (1989b).

J. D. Rhoades, P. A. C. Raats, R. J. Prather, Effects of liquid-phase electrical conductivity, water content, and surface conductivity on bulk soil electrical conductivity. *Soil Sci. Soc. Am. J.* **40**, 651 (1976).

J. D. Rhoades, P. J. Shouse, W. J. Alves, N. A. Manteghi, and S. M. Lesch, Determining soil salinity from soil electrical conductivity using different models and estimates. *Soil Sci. Soc. Am. J.* **54**, 46 (1990).

L. A. Richards, C. A. Bower, and M. Fireman, *Tests for Salinity and Sodium Status of Soil and of Irrigation Water*. USDA Circular No 982, (U. S. Government Printing Office, Washington, DC, 1956) pp. 1–19.

J. S. Samra, and H. S. Gill, Modeling of variation in a sodium-contaminated soil and associated tree growth. *Soil Sci*. **155**, 148 (1993).

I. Szabolcs, *Salt-affected soils* (CRC Press, Inc, Boca Raton, FL, 1989).

J. A. Tabor, A. W. Warrick, D. A. Pennington, and D. E. Myers. Spatial variability of nitrate in irrigated cotton: I. Petioles. *Soil Sci. Soc. Am. J*. **48**, 602 (1984).

J. A. Tabor, A. W. Warrick, D. E. Myers, D. A. Pennington. Spatial variability of nitrate in irrigated cotton: II. Soil nitrate and correlated variables. *Soil Sci. Soc. Am. J*. **49**, 390 (1985).

U.S. Salinity Laboratory Staff, *Diagnosis and Improvement of Saline and Alkaline Soils*. USDA Agricultural Handbook No, 60 (U.S. Government Printing Office, Washington, DC, 1954).

S. R. Yates, and A. W. Warrick, Estimating soil water content using cokriging. *Soil Sci. Soc. Am. J*. **51**, 23 (1987).

S. R. Yates, A. W. Warrick, and D. E. Myers, Disjunctive kriging. I. Overview of estimation and conditional probability. *Water Resour. Res*. **22**, 615 (1986a).

S. R. Yates, A. W. Warrick, and D. E. Myers, Disjunctive kriging. II. Examples. *Water Resour. Res*. **22** 615 (1986b).

S. R. Yates, R. Zhang, P. J. Shouse, and M. Th. van Genuchten. Use of geostatistics in the description of salt-affected soils, in: *Water Flow and Solute Transport in Soils*, edited by D. Russo and G. Dogan (Springer-Verlag, New York, 1993), pp. 283–304.

R. Zhang, D. E. Myers, and A. W. Warrick, Estimation of the spatial distribution of soil chemicals using pseudo-cross-variograms. *Soil Sci. Soc. Am. J*. **56**, 1444 (1992a).

R. Zhang, S. Rahman, G. F. Vance, and L. C. Munn, Geostatistical analyses of trace elements in soils and plants. *Soil Sci*. **159**, 383 (1995).

R. Zhang, P. Shouse, and S. Yates, Use of pseudo-cross-variograms and cokriging to improve estimates of soil solute concentration. *Soil Sci. Soc. Am. J*. **61**, 1342 (1997).

R. Zhang, A. W. Warrick, and D. E. Myers, Improvement of the prediction of soil particle size fractions using spectral properties. *Geoderma* **52**, 223 (1992b).