Available online at www.sciencedirect.com
SCIENCE DIRECT®
ELSEVIER
Computers and Electronics in Agriculture 46 (2005) 153–179
Computers and electronics in agriculture
www.elsevier.com/locate/compag

Sensor-directed response surface sampling designs for characterizing spatial variation in soil properties

S.M. Lesch*

USDA-ARS, George E. Brown Jr., Salinity Laboratory, 450 W. Big Springs Road, Riverside, CA 92507, USA

Abstract

In many applied precision farming applications, remotely sensed survey data are collected specifically because these data correlate well with some soil property of interest. Additionally, a general model for the functional relationship between the soil property and the sensor data is often known a priori, but the exact parameter estimates associated with the model must still be determined via some type of site-specific sampling strategy. The main objective of this paper is to present an objective sampling and simplified modeling strategy for predicting soil property information from such spatially referenced sensor data. Some common types of spatial linear prediction models and linear geostatistical models are reviewed, and the assumptions needed to reduce these more complicated models to a spatially referenced, ordinary linear regression model (LR) are discussed. Next, a model-based sampling strategy for estimating an ordinary linear regression model in the spatial setting is described. This sampling strategy incorporates a traditional response surface design into an iterative, space-filling type algorithm for purposes of selecting sample site locations that are (i) nearly optimal with respect to matching the selected response surface design levels and (ii) physically separated far apart as possible to ensure the best chance that the independent error assumption is adequately met. This strategy can in principal be used to select a minimal number of optimal sample site locations that satisfy the residual independence assumptions in the ordinary model. A detailed case study of a salinity survey using electromagnetic induction (EM) and four-electrode sensor data is then presented. These case study results confirm that the sampling strategy was highly effective at ensuring efficient regression model parameter estimates and a reliable salinity prediction map. An additional simulation study confirmed the effectiveness of this model-based strategy over a more traditional simple random sampling strategy with respect to four regression model design criteria. Under the right conditions, this methodology should be applicable to many types of precision farming survey applications where

* Corresponding author.
E-mail address: slesch@ussl.ars.usda.gov.

0168-1699/$ – see front matter. Published by Elsevier B.V.
doi:10.1016/j.compag.2004.11.004

soil property/sensor data prediction models need to be fitted using only a limited number of soil samples.
Published by Elsevier B.V.

Keywords: Precision agriculture; Salinity; Spatial linear model; Variogram; Response surface design

1. Introduction

The collection of apparent soil electrical conductivity (EC_a) survey data for the purpose of characterizing various spatially referenced soil properties has received considerable attention in the soils literature in the last 20 years (Corwin and Lesch, 2003). Most of the original interest was directed towards the characterization of field scale soil salinity patterns (Rhoades et al., 1999; Hendrickx et al., 2002). However, EC_a survey data are being increasingly used in precision farming applications in an effort to obtain information on numerous soil properties. In practice, apparent soil conductivity survey data often correlate reasonably well with various soil properties (salinity, soil texture, soil water content, etc.) under different field conditions (Lesch and Corwin, 2003). Not surprisingly, EC_a data have therefore been used extensively in precision agriculture survey applications for characterizing the spatial variability of these properties.

The basic idea behind the theory of precision farming is to exploit the knowledge of the inherent spatial variability of the soil and crop condition(s) in a specific field or region to design better management practices (Larson and Robert, 1991). In turn, these better agricultural management practices should lead to higher crop yield and/or more optimal use of agrichemicals, water, time and labor, etc., and thereby improve sustainability through increased production and profit margin, decreased input requirements, and/or a reduction in detrimental environmental impacts.

In principal, many precision farming management strategies hold great promise. However, in practice, these same management strategies are greatly affected by both the availability and accuracy of the spatially referenced input soil properties. In situations where direct, exhaustive sampling must be employed to gather adequate spatial precision about one or more input soil properties, the theoretical gain in profit from a site-specific management strategy can be quickly offset by the extra cost incurred by the sampling effort. Hence, many promising management strategies are often not cost effective in practice, due to the need for exhaustive sampling of the necessary baseline input variable(s).

This sampling cost issue has led many researchers to aggressively pursue the idea of collecting secondary ground- or air-based remote sensing information as surrogate data, i.e., data that can be used to help infer the detailed spatial pattern(s) of the primary input property(ies) of interest. Survey EC_a data are perhaps the most common example of surrogate remote sensing data, but certainly not the only one. Other examples of remote sensing data are numerous, and include various types of imagery data, natural gamma ray measurements, time-domain electromagnetic induction (EM) and time-domain reflectometry, ground-penetrating radar, and direct-yield monitoring measurements, etc. Regardless of the type of data collected, the basic idea is the same: the sensor data are acquired to increase