DATA COLLECTION AND ANALYSIS METHODS FOR DATA FROM FIELD EXPERIMENTS

S. Shibusawa and **C. Haché**
Faculty of Agriculture, Tokyo University of Agriculture and Technology, Japan

Keywords: Field experiments, sampling methods, spatial and temporal variability, experimental design, precision agriculture, remote sensing, soil and crop sensors, multivariate analysis, geostatistics.

Contents

Summary

Field experiments are conducted to extract in-situ features of interest from complex agricultural phenomena. Attributes of data and information obtained from the field depend on instrumentation tools, data analysis methods and experimental designs. Currently researchers across the world have been developing precision agriculture, which in addition to getting averages and variances of both crop and soil parameters, also enhance description and understanding of the spatio-temporal variability using new developed technologies. In this chapter, a remote sensing approach is described focusing on the spatial variability of crop and soil in an experimental field, using spectroscopic techniques from visible to near-infrared light energy reflection. Sensors installed on airborne platforms collected images of an experimental field and the differences between tillage practices and between fertilizers treatments were confirmed. On-the-go soil sensors and crop sensors are also introduced for providing the data of variability of soil and crop parameters. A real-time soil spectrophotometer is one of the innovating tools to provide information about multiple underground soil parameters, such as moisture and soil organic matter content, as well as to supply correct location data. A prototype of mobile fruit-grading robot is also an attractive approach for creating field maps of yield and quality of pepper fruits during in-situ grading operation. Multivariate methods are available for the analysis of high dimensional data such as those obtained from hyper-spectral sensors. Techniques for smoothing, Kubelka-Munk transformation and multiplicative scatter correction are explained as spectral data treatments. Calibration models are also discussed, such as principal component analysis and partial least square regression, with regard to multi-collinearity and model accuracy. The

semi-variance analysis and kriging method is introduced as a mapping technique, and a case study shows that sample size clearly influences the kriging error, followed by a recommendation for appropriate sampling size.

1. Introduction

The main motivation for field experimentation is to produce information relevant to producers and/or to determine the effects of agricultural practices on the environment. In order to achieve these goals, it is imperative that the interrelationships among environmental conditions, biological processes, and management are well understood by the researcher. This need drives the development of new methods and devices for data collection in agriculture, in addition to the adoption of advanced analysis techniques. New devices and sensors are making it possible to collect vast amounts of new data covering, in some cases, whole fields and giving details of spatial and temporal variability. More advanced data analysis methods are helping to extract more information from the data, develop more accurate prediction models, and optimize simulations for decision support in agriculture.

Conventional tools and methods for data collection and analysis are not covered in this chapter. The focus is rather on state-of-the-art technology applied currently in field research and on analysis techniques that allow the inclusion of numerous variables resulting in better description of the agricultural phenomena of a whole field.

2. Data Collection

2.1. Conventional Data Collection

Traditionally, agronomic field research has applied replication, blocking and randomization in experimental design to avoid influences of spatial variability as errors or biases. Yet, conventional experimental designs are characterized by limitations (e.g., small plots, treatments oversimplification, and brief duration) and consequently may not represent a realistic cropping system. In field experiments effects and quantification of variation are measured through sampling. Sampling density depends on several factors (objectives, field variability, costs), and can range from one sample for several hectares to a more detail coverage of the field. Conventionally, samples are obtained for whole fields or parts of fields to provide average values. There are several commonly used sampling methods characterized by destructive sampling (Figure 1):

- Simple random: Locations are randomly selected, and may not capture the variation structure of the attributes of interest (Figure 1a).
- Stratified random: The field is divided into several areas according to its characteristics (e.g. topography), and sampling locations are selected randomly and then composite, reducing the influence of local heterogeneity (Figure 1b).
- Systematic (grid sampling): The field is divided in grids and samples are collected randomly within each cell and then composite (Figure 1c). Another approach is to position the center point on grid intersections, where samples are collected randomly within a 3 m radius (10 feet) and then composite (Figure 1d).
- Stratified-systematic: Each cell is further divided into smaller cells to try to

overcome the bias introduced by systematic sampling (Figure 1e).

- Judgmental: Sampling locations are decided based on observation of a specific problem (e.g., low yield) and is not statistically accurate (Figure 1f).

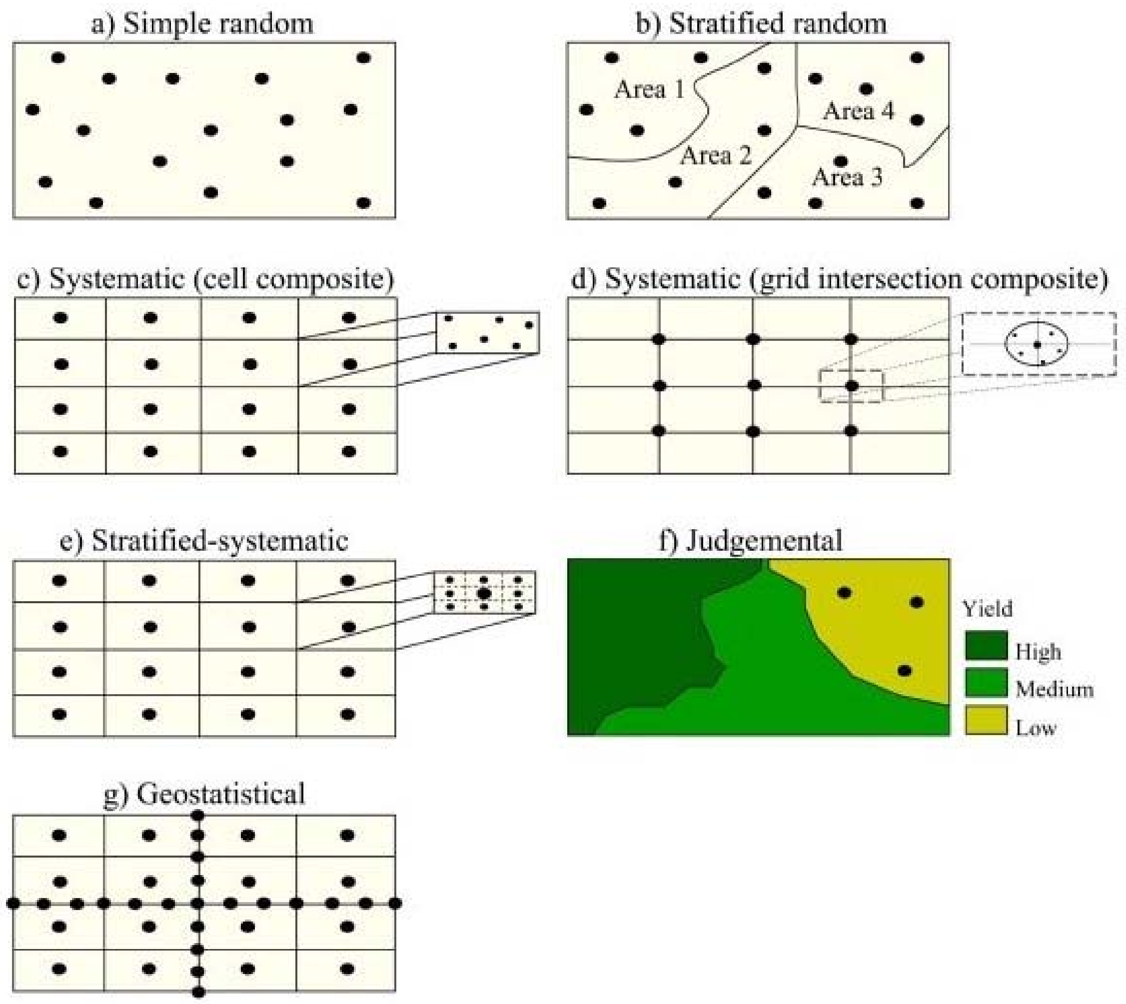

Figure 1: Sampling strategies.

Sample collection involves intensive labor and costs of laboratory analysis, imposing a limitation on the number of samples that can be collected to quantify the experimental error among treatments repetitions. Nevertheless, reducing the number of samples has direct implications on management since it can lead to incorrect decisions. The requirement for improved efficiency has increased the interest in conducting field experiments that take into account spatial variability and reproduce better scenarios for real farm.

2.2. Precision Agriculture

Recently, it has become possible to quantify within-field spatial variability because of the availability of technologies such as Global Positioning Systems (GPS) and Geographic Information Systems (GIS). The GPS enables collection of geo-referenced data, while the GIS allows spatial analysis and visualization of interpolated maps.

Application of GPS/GIS into agriculture has caused a revolution called precision agriculture (PA), where fields are managed at a detailed scale based on information and knowledge. The PA cycle covers all steps in crop management as presented in Figure 2.

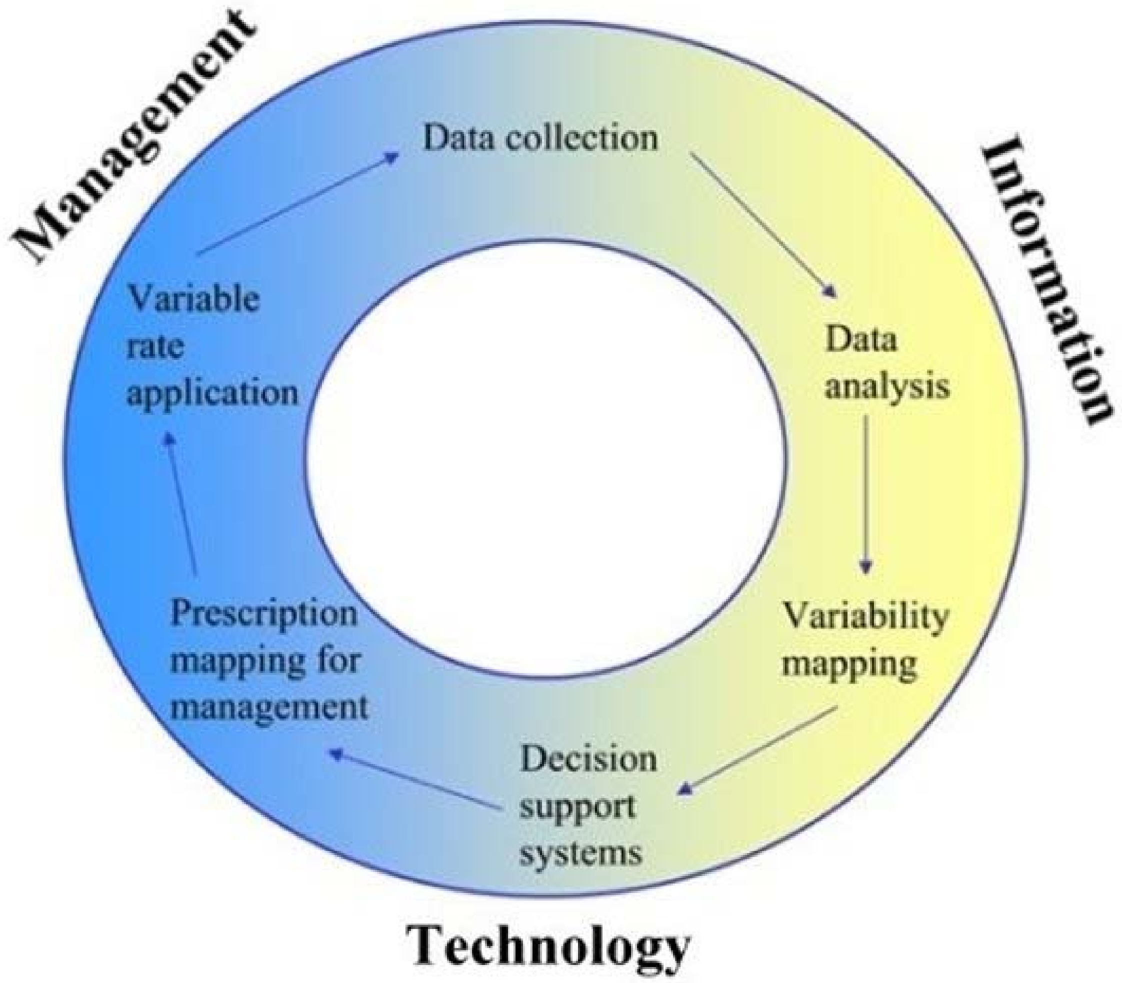

Figure 2: Precision agriculture cycle.

New technologies used in PA allow collection of large amounts of data. As a result, interest is now directed toward understanding spatial and temporal variability in agricultural systems, including their effects or constraints on production and relationships among multiple components and factors. Consequently, field experimentation is moving from small homogenous experimental areas to large and variable on-farm areas. This new concept allows farmers to integrate in the experimental process and to accept new successful practices. At on-farm level, experimental units have been single fields with uniform management without replication. However, knowledge of within-field variability leads us to divide a whole field into sub-unit areas according to soil or other variability.

New technologies and analysis methods have accordingly changed strategies for data sampling as shown Figure 1:

- Targeted or directed: Samples are collected according to statistically rigorous sampling designs. Evidence for a change in the value of a measured property observed from aerial images, yield or other maps is then used to determine supplemental locations.

- Geostatistical: Applied when the aim is to produce accurate interpolated kriged maps. When the semivariogram is known the distance between sample locations is equal to half the semivariogram range. But when the semivariogram is not known, the samples are collected as in systematic sampling and in several transects composed of adjacent sampled sites; then the semivariogram as well as the accurate sampling distance can be defined reducing the number of samples to be collected in the near future (Figure 1g). The concept of geostatistics is explained later in this chapter.

To understand spatial variability a large number of data is needed, which can require a lot of human effort for data acquisition. To address this fact, automatic mobile soil samplers have been developed. Continuous-sampling is another emerging solution, where every location in the field is measured by non-invasive techniques (e.g., electromagnetic induction, remote sensing), and there is no need for interpolation or soil sampling design. A considerable amount of effort has now focused on developing real-time sensors to aid in sampling schemes for PA. Nevertheless, sampling will continue to be useful for calibration purposes, as well as development of statistical methods suitable for the analysis of different soil and crop types.

What follows is an overview presenting new technologies developed to collect high amounts of data from field experiments. Remote sensing (RS) is applicable to both soil and crop data collection, and therefore a brief outline of RS is addressed first.

2.2.1. Remote Sensing

Remote sensing is the process of gathering information about an object without direct physical contact. Passive remote sensing systems use solar radiation as a source to collect information about objects, based on the principle that visible (VIS, 400-700 nm), near infra-red (NIR, 700-2500 nm) and mid infra-red (MIR, 3000-5000 nm) light is absorbed, transmitted or reflected, and that in the thermal infra-red region (TIR, 7500-14000 nm) heat is emitted. Passive systems can also obtain information of gamma rays emitted from the Earth surface with wide application in geological surveys. Active systems use either TIR or an artificial source of radiation as in the case of synthetic aperture radar. Currently, RS systems used in agriculture are mounted on space-, airborne-, and/or on ground-platforms. Spectra are collected using either multispectral or hyperspectral sensors, covering VIS, NIR, and/or TIR. Each band gives different information about the object under observation. Figure 3 shows an airborne multispectral image and its spectral bands, collected over a field in Japan where crop response to two types of tillage interacting with two types of nutrients is investigated. From these images it is clear that each band gives different information about treatments in the field.

While multispectral images use non contiguous or wide bands in the VIS and NIR regions, hyperspectral images contain from 10 to hundreds of contiguous bands. The dimensionality of hyperspectral images allows the identification of bands most responsive to specific target characteristics and the potential for the improvement of classification analysis (Figure 4). Hyperspectral remote sensing is also known as imaging spectroscopy since it combines imaging and spectroscopy in a single system.

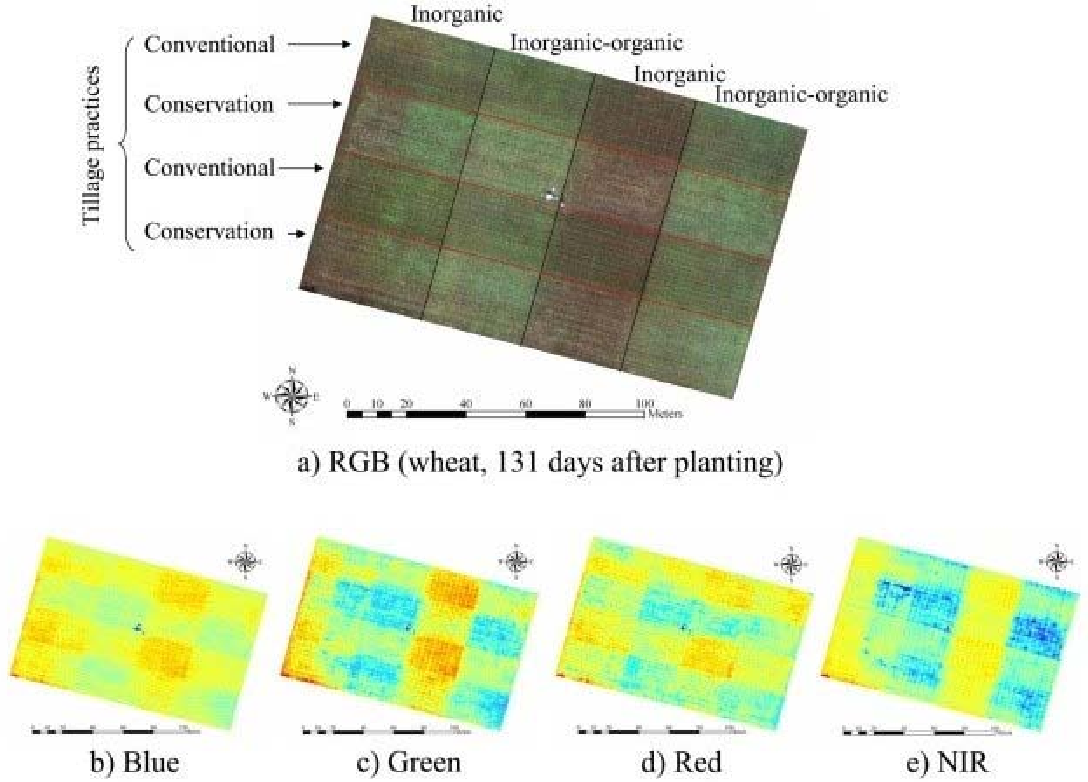

Figure 3: Multispectral airborne image and its spectral bands collected over an experimental field in Japan.

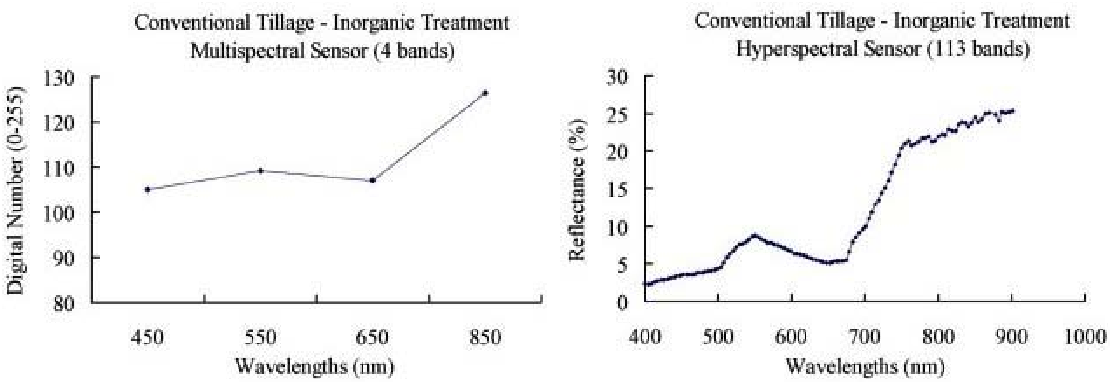

Figure 4: Data extracted from multispectral and hyperspectral images collected on the same day over the experimental field shown in Figure 3. Spectral curves correspond to conventional tillage-inorganic treatment.

Although reflectance of soils and crops are related to numerous parameters, agricultural RS has not yet been widely adopted by farmers. Since most targets are non-Lambertian and the atmosphere absorbs part of the light energy reflected in narrow bands, image pre-processing is needed to reduce sensor noise, correct geometric and optical distortions and georegister images, followed by calibration to reflectance for variable illumination if time series data are required. This process is skill demanding and time consuming, and therefore the interest in terrestrial vehicle-mounted RS systems has increased given that images and/or reflected spectra can be collected near the soil or canopy and therefore those effects are less significant. However, these measurements are still influenced by bi-directional reflectance due to inclination and cupping of individual leaves, and require data pre-treatment, correction, and on-ground calibration measurement to assure the quality of the data, especially for images. The final goal of vehicle-mounted-sensors is to develop real-time data acquisition tools and simultaneously perform variable rate management to meet site-specific requirements.

-
-
-

Bibliography

Dyke, G. V. (1998). *Comparative experiments with field crops*, 262 pp. Charles Griffin & Company Ltd. New York. USA. [This gives comprehensive information on planning field experiments].

Isaaks, E. Srivastava, R. (1989). *Applied Geostatistics* 561pp. Oxford University Press, USA. [This is an excellent book for self-tutoring in geostatistics].

Moran, M. S., Inoue, Y., Barnes, E. M. (1997). Opportunities and limitations for image-based remote sensing in precision crop management. *Remote Sensing of Environment*, **61**, 319-346. [This presents a comprehensive review of current satellite-based sensors].

Reyns, P., Missotten, B., Ramon, H., De Baerdemaeker, J. (2002). A review of combine sensors for precision farming. *Precision Agriculture* **3**, 169-182. [This is instructive on yield monitor sensors].

Webster, R., Oliver, M.A., 2001. *Geostatistics for Environmental Scientists*, 286 pp. John Wiley & Sons. [This is recommended for more information on sampling procedures and for a complete coverage of geostatistics].

Biographical Sketches

Dr. Sakae Shibusawa is a Professor of the Faculty of Agriculture at Tokyo University of Agriculture and Technology. His work is focused on agricultural engineering, precision farming and regional business models. His research interests include real-time soil sensing, foods and waste supply chain, and the organization of learning groups of farmers and companies for the community-based precision farming. In addition, he has been actively involved in the development of agricultural machines, such as deep rotary tillage, and in phytotechnolgy based on speaking-plant approaches. He is the editor and co-author of the book “Precision Agriculture” (in Japanese). This book describes the Japanese model for the comunity-based precision agriculture and how it relates to science, technology and business. It also explains how to organize a community for precision agriculture practices. He is also a co-author for the “Handbook of Precision Agriculture” (Ed. Ancha Srinivasan), where he describes the worldwide state-of-the-art of soil sensing technologies available for precision agriculture.

Carolina Haché is a research fellow under the supervision of Dr. Sakae Shibusawa at Tokyo University of Agriculture and Technology. Her research interests are precision agriculture, remote sensing, and the application of conservation practices in agriculture to reduce Carbon emissions. She has experience in collecting vast amounts of data from field experiments, and in applying multivariate techniques for the analysis of hyperspectral data from remote sensing sensors and from the Real-Time Soil Spectrophotometer.