14 Data-Driven Precision Agriculture

Opportunities and Challenges

Wenxuan Guo, Song Cui, Jessica Torrion, and Nithya Rajan

CONTENTS

14.1 INTRODUCTION

Agriculture is facing the greatest challenge of feeding more than 9 billion people by 2050 in a manner that advances economic development and a healthy environment. A 70% increase from 2006 in food production is required to meet this demand (World Resources Institute 2013). With limited resources of land and water, increased agricultural production is projected to come primarily from intensification on existing arable land (FAO 2011). Conventional farming practices treat an agricultural field uniformly despite the inherent variability in soil properties and crop growth conditions. Uniform management may result in over- or under-application of resources in specific locations within a field, which may have a negative impact on the environment and profitability (McKinion et al. 2001; Plant 2001). Sustainable agriculture is a viable means of meeting the food demand while balancing crop production and minimizing environmental impacts. Precision agriculture (PA) is a promising approach to attain sustainable agriculture.

PA is the management of soil and crops at subfield scale using information and technology for optimum profitability, sustainability, and protection of the environment (Robert et al. 1995, 1996; National Academy of Sciences 1997). Also called site-specific management, PA involves farming practices to apply the right amount of right resource at the right place, at the right time, and in the right manner (Khosla 2010; Robert et al. 1995). For PA to be applicable, as indicated in the definition, significant within-field spatial variability must exist in soil properties and crop growth; such variability can be identified and measured and information from these measurements can be used to improve crop production and environment (Miller et al. 1988). In order to determine the variability, intensive soil and plant samples may be collected and analyzed in the laboratory, but the costs of sampling and analyses will potentially exceed the benefit from the site-specific management (Swinton and Lowenberg-DeBoer 1998). Thus, efficient methods for accurately measuring within-field variability in soil properties and plant growth are important for PA (Bullock and Bullock 2000). The use of advanced technologies, such as the Global Positioning System (GPS), Geographic Information System (GIS), yield monitors, and remote sensing, enables efficient quantification of spatial variability in soil properties, crop growth, and crop yield within fields. With variable rate technologies (VRT) together with precision guidance systems, inputs (i.e., fertilizers, water, pesticides, seeds) can be applied according to the spatial variability in soil properties and crop yield potential.

The main components of PA include data collection, data interpretation and analysis, and implementation of management at an appropriate scale and time (National Academy of Sciences 1997). As technology advances, numerous tools and sensors have been developed to collect spatial data and information, such as yield monitors, satellites, combine-mounted crop scanners, handheld sensors, and on-the-go soil property sensors. Massive data and information can be collected within a relatively short period of time. It remains a challenge to appropriately collect, analyze, and interpret the data, as well as applying the derived information and knowledge to site-specific management. The objective of this chapter is to review the opportunities and challenges of PA technologies used to collect, analyze, and apply spatial and temporal data and information for optimized agricultural production.

14.2 TECHNOLOGIES AND DATA COLLECTION

14.2.1 Yield Monitoring and Yield Data

Yield monitoring and mapping using combine-mounted yield monitors, often recommended as a first step, has become one of the most widely adopted PA technologies (Figure 14.1). Yield monitors include sensors to measure the mass or volume of product flow, ground speed, moisture, and header position. Yield is derived as a product of these parameters being sensed. The combine ground speed and cut width are used to determine the harvest area per unit time. A submeter accurate differential GPS (DGPS) receiver is usually installed together with a yield monitor to record the geographic position (latitude and longitude), time, and elevation of each yield data point.

The yield monitor plays two roles in PA. On the one hand, yield monitors generate spatially dense data with relatively low cost, allowing characterization of the spatial and temporal yield variability (Dobermann et al. 2003; Pierce and Nowak 1999). Based on the yield variability, site-specific

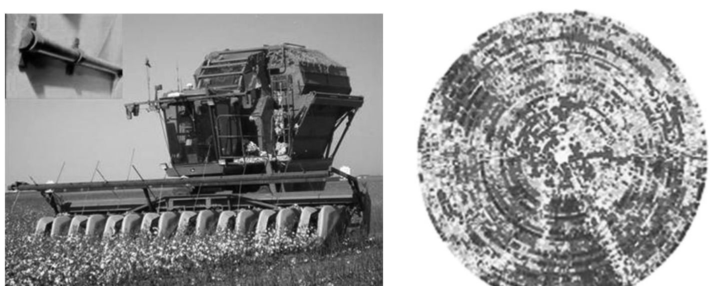

FIGURE 14.1 Cotton yield monitor mounted on a harvester and the yield map generated from the yield data.

management can be implemented to match the yield potential at different locations of the field. On the other hand, yield monitor data can also be utilized to evaluate the effectiveness or the response of PA practices. Producers have used yield data and maps for crop moisture monitoring, documenting yields, field experiments, drainage tiles installation, new crop lease negotiation, dividing crop production, and bottom-line considerations (Griffin 2009). In addition, yield maps can be used to guide field scouting, design soil sampling schemes, and calculate nutrient requirements for variable rate fertilizer applications.

Users need to be cautious when applying yield data and yield maps for site-specific management. Yield data contains various systematic and random errors. The sources of errors can be classified into four categories: sensor errors, errors due to operating conditions, operator errors, and yield mapping errors (Thylen et al. 1997). The main errors involved in yield mapping are unknown crop cutting width, grain lag time, GPS error, grain mixing through combine components, combine grain losses, and calibration (Blackmore and Marshall 1996). Postharvest processing is needed to remove these errors, especially when yield data is compared to other data layers within the decision support system. Various techniques have been applied to remove errors related to unrealistic cycle distance, moisture, combine speed surge, wrong cutting widths, overlapping yield points, and so forth. A software tool called Yield Editor (Sudduth and Drummond 2007) was developed and widely adopted to simplify the process of applying filtering techniques for yield data outlier detection and removal. However, when using the tool, users must make sure that the parameter settings are appropriate, because these parameters may be specific to each field. The user's experience and knowledge should be incorporated in such data cleaning.

Another challenge is interpreting yield variability and the underlying causes of such variability in yield maps. Yield variability may be caused by many factors, including spatial variability in soil type, landscape position, crop history, soil physical and chemical properties, and nutrient variability (Wibawa et al. 1993). Interactions among biotic (plant genotype, soil fauna, pests, and diseases) and abiotic factors (soil physical, chemical, moisture characteristics, and climatic conditions) influence spatial yield variability. Effects of crop stress, pests, and diseases on crop yield are temporal factors that could explain up to 50% of crop yield variability across years and sites (Machado et al. 2002). As a result, yield maps tend to vary from year to year, which makes it more difficult for producers to make decisions on site-specific management. A single-year yield map is useful for interpretation of possible causes of yield variation but may be of limited value for more strategic and long-term site-specific management. With multiple years of yield data, repeating patterns and their more stable natural causes may be separated from random variation in each year, providing a basis for spatially varying yield goals or other site-specific management practices (Dobermann et al. 2003). For example, Blackmore (2000) proposed a method to use multiple years of yield maps to classify a field into three categories: high-yielding and stable, low-yielding and stable, and unstable. Based on the spatial and temporal trends of the yield maps, the economic significance of these areas can be assessed and site-specific management can be implemented based on the economic return of each area. In summary, yield maps alone cannot give clear guidelines for site-specific management unless the sources of variation are identified, especially those related to soil physical and chemical properties and seasonal weather conditions.

14.2.2 Remote Sensing

Remote sensing data from various platforms, including satellites, aircrafts, unmanned aerial systems, and field vehicles, has been used as an important source for obtaining spatial information of soil and crops for PA. Remote sensing technology is a non-destructive method that can systematically collect information about agricultural fields over a large geographical area. Remote sensing data can reveal unbiased information about areas that are sometimes inaccessible to humans (Liaghat and Balasundram 2010). Remote sensing has been applied to evaluate crop growth (Clevers and Van Leeuwen 1996; Moran et al. 1997; Sakamoto et al. 2005), leaf area index (Delegido et al.

2013; Papadavid et al. 2013; Zhao et al. 2012), chlorophyll content (Daughtry et al. 2000; Gitelson and Merzlyak 1996; Zheng and Moskal 2009) ground cover (Maas 1998; Rajan and Maas 2009; Rajan et al. 2014), pest and disease infestation (Curran et al. 2000; Kelly and Guo 2007; Prabhakar et al. 2012; Qin and Zhang 2005), and soil physical and chemical properties (Bausch et al. 2004).

For many years, remote sensing relied on spectral reflectance data in the visible and infrared wavelengths of the electromagnetic spectrum for the quantitative and qualitative analysis of soil and crop characteristics. In recent years, remote sensing technology has advanced beyond the commonly used visible and infrared sensors and includes hyperspectral sensing, thermal sensing, and light detection and ranging (lidar). Hyperspectral remote sensing in narrow bandwidths has allowed the application of remote sensing in identifying specific biophysical and biochemical characteristics of crops (Pacheco et al. 2001; Thenkabail et al. 2013; Zhang et al. 2003). Lidar holds promise for many PA applications (Eitel et al. 2014; Tang et al. 2014). For example, Andujar et al. (2013) successfully used lidar for weed detection and discrimination in a maize (*Zea mays* L.) cropping system. Remote sensing in thermal wavelengths is being used to detect soil moisture status and water stress in plants. This offers the opportunity of assessing crop water requirement at a subfield scale for variable rate irrigation management (Torrion et al. 2014).

The application of unmanned aerial vehicles (UAVs) has tremendous potential for acquiring aerial images because of the low operational cost, high temporal and spatial resolutions, easy-to-use controlling system, and high flexibility in image acquisition planning. The majority of UAVs currently in use are designed to operate at low altitudes (less than 150 m) and adverse weather conditions. For UAV-based image acquisition systems, clouds and other kinds of atmospheric interference are not as influential as they are to satellite and piloted airplanes. Lower flying height also means much higher spatial resolution, enabling finer-scale interpretation of soil and crop characteristics from images. However, UAV-based image processing and interpretation at the producer level still require development of analyzing and processing tools before they can be easily adopted by farmers. Additionally, short flying duration, lack of stability, limited communication distance, scarcity of affordable lightweight camera systems, and public skepticism and misinformation greatly challenge the application of UAVs in PA.

The latest advancements in remote sensing data collection for precision management are geared at measuring individual plant characteristics in real time (Mulla 2013). Sensors mounted on field vehicles and stationary sensors are becoming valuable tools for measuring soil and crop characteristics. The considerable interest in collecting this data at high spatial, spectral, and temporal frequencies is posing new challenges in processing remote sensing data for PA applications because it requires sophisticated computing and analyzing capabilities for processing massive amount of data.

14.2.3 Soil Sampling and Soil Data

The purpose of soil sampling in PA is to assess soil fertility and determine the amount of fertilizer required to produce certain yield goals at different locations in the field. In conventional agriculture, soil samples are collected, the average of a certain nutrient is calculated, and the amount of fertilizer for the whole field is based on this average. In PA, soil samples are collected with geographic information attached to each soil sample using a GPS receiver. The amounts of fertilizer applied vary according to the needs at specific locations. There are two main methods of soil sampling in PA: grid sampling and directed sampling. Grid sampling is an ideal approach if there is no prior knowledge of the fertility variability within the field. Soil test results from a well-designed grid sampling scheme provide an accurate base nutrient map for long-term management. However, a great number of samples are required, which can be very expensive and time-consuming. Directed soil sampling, on the other hand, requires prior knowledge of the field characteristics that may be limiting crop yield. A background layer containing subfield regions with different characteristics can be constructed for directed soil sampling. The background layer can be a yield map, soil types, cropping systems, remote sensing imagery, and so forth. A standard test result usually includes available

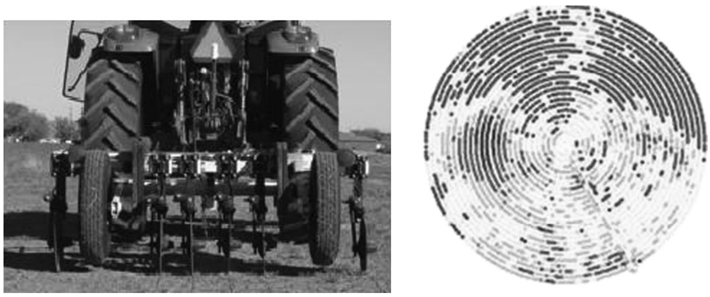

FIGURE 14.2 Veris soil EC mapping system (left: Veris Technologies) and an apparent soil electrical conductivity map.

phosphorous (P), exchangeable potassium (K), calcium (Ca), magnesium (Mg), cation exchange capacity (CEC), pH, and so forth. Some laboratories may also test for organic matter, nitrate, salinity, sulfate, certain micronutrients, and heavy metals (Foth and Ellis 1988).

Soil sampling and soil laboratory analysis can be labor-intensive, costly, and time-consuming. As a result, usually only a limited number of soil samples are collected and analyzed. The sparse spatial distribution of soil test data, such as the levels of macronutrients and micronutrients, limits the scale of site-specific management. Efficient methods for accurately measuring within-field variations in soil physical and chemical properties are critical for PA (Bullock and Bullock 2000). Various on-the-go soil sensors have been or are being developed to measure soil texture, organic matter, moisture content, salinity, bulk density, topsoil depth, pH, nitrate, CEC, and so forth (Adamchuk et al. 2004). The measurement of soil apparent electrical conductivity (EC_a) has been widely adopted and increasingly used in PA (Clay et al. 2001; Corwin et al. 2003; Johnson et al. 2001). There are two types of sensors commercially available: contact and noncontact. One commonly used contact type is the Veris mapping system (Figure 14.2) (Veris Technologies, Salina, KS) that consists of a Wenner array (coulters) and records EC_a by electrical resistivity at a shallow depth (0–30 cm) and a deep depth (0–90 cm) simultaneously. Noncontact EC_a sensors are typically composed of a transmitter and a receiver coil. They measure EC_a without contacting the soil surface via electromagnetic induction. Examples of these types of sensors are EM38 (Geonics Limited, Mississauga, Ontario, Canada) and GEM-2 (Geophex, Raleigh, NC).

The purpose of the field-scale EC_a survey used in site-specific management is to establish the within-field variability in soil properties influencing the variability in crop yield (Figure 14.2). However, directly relating crop yield and EC_a measurement has resulted in inconsistent results due to the fact that EC_a measurements are affected by multiple factors. Spatial variability in soil electrical conductivity is related to such factors as texture, organic matter, CEC, landscape positions, salinity, subsoil characteristics, soil water content, and depth to claypan (Clay et al. 2001; Rhoades 1993). To use EC_a in PA, it is necessary to understand the factors that most significantly influence the EC_a measurement. Simple statistics or wavelet analysis can be used to determine the dominant factors influencing EC_a measurement (Corwin and Lesch 2003). Although temporal variability exists, the relative spatial pattern of EC_a distribution within a field is considerably stable (Clay et al. 2001). As a result, an EC_a map provides useful spatial information to identify potential areas in need of improved irrigation, drainage, fertilizer, and pest management (Corwin and Lesch 2003).

14.2.4 Topographic Data and Digital Terrain Analysis

Topographic and hydrological attributes, including primary and secondary attributes, have been widely used in PA, especially in site-specific management of seed, irrigation water, and fertilizer (Iqbal et al. 2005). A digital elevation model (DEM), a three-dimensional (3-D) representation of a terrain's surface, is the most commonly used format to represent the elevation. A DEM is

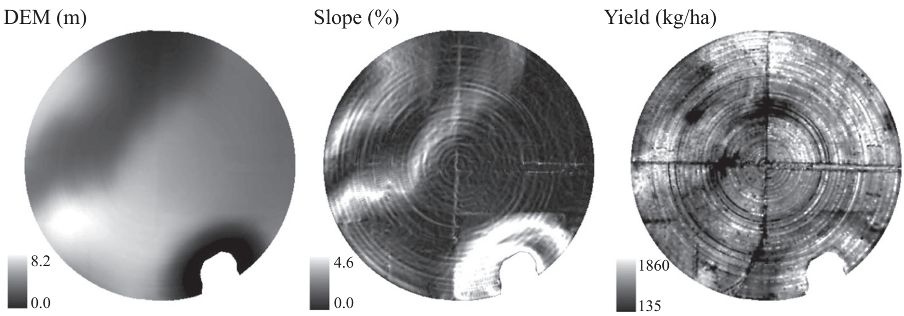

FIGURE 14.3 Digital elevation model (DEM), slope, and yield map of a field in Texas.

usually derived from point elevation data through various interpolation methods. After the DEM is obtained, primary and secondary terrain attributes can be derived (Figure 14.3). Primary attributes include slope, aspect, plan and profile curvature, flow path length, and upslope contributing area (Wilson and Gallant 2000). Elevation, slope, and curvature have a direct effect on infiltration and runoff through their influence on surface and subsurface water flow. Water tends to move downslope causing lower positions to receive water from adjacent higher positions (Kaspar et al. 2003). Secondary attributes are computed from two or more primary attributes. Secondary attributes mainly include topographic wetness index, sediment transport index, stream power index, flow direction, and flow length. These attributes may affect soil characteristics, distribution and abundance of soil water, susceptibility of landscapes to erosion by water, and the distribution and condition of plant growth (Wilson and Gallant 2000). In addition, topography influences the redistribution of soil particles, organic matter, and soil nutrients. Topography has been reported to be related with the yield of various crops, such as corn (Kaspar et al. 2003; Kravchenko and Bullock 2000), wheat (*Triticum aestivum* L.) (Sinai et al. 1981), cotton (*Gossypium hirsutum* L.) (Bronson et al. 2003; Guo et al. 2012; Iqbal et al. 2005; Li et al. 2001), sorghum (*Sorghum bicolor* L.) (Machado et al. 2002), soybean (*Glycine max* L.) (Kravchenko and Bullock 2000), and dry pea (*Pisum sativum* L.) (Mahler et al. 1979). In PA applications, seed, fertilizer, and irrigation water rates can be varied according to topographic properties. For example, consistently low-yielding areas (on summit or steep slope) would receive lower seed, fertilizer, and water rates to reduce the waste of recourse and improve overall profit.

In principle, any data that contains elevation information can be used as source for digital terrain analysis. However, for PA applications, highly accurate elevation data is required. Lidar can be used to effectively survey elevation of a field (Galzki et al. 2011). However, this is usually difficult to set up and may be cost-prohibitive. Automated guidance equipments with real-time kinematic (RTK) GPS receivers have gained rapid and widespread adoption. Real-time kinematic is a carrier-phase-based survey method of determining relative positions between receivers simultaneously tracking the same satellites (Sickle 2001). An RTK GPS receiver has a nominal positioning accuracy of 1 cm in horizontal directions and 2 cm in vertical directions (Tamura et al. 2002). With RTK auto-steering systems, dense and accurate elevation data are collected during any field operations at no additional cost. However, the user needs to pay close attention to the quality of raw elevation data. RTK GPS receivers may sometimes lose correction, resulting in low accuracy in elevation. These errors must be filtered out before conducting digital terrain analysis.

14.2.5 Temporal Variability

Many studies have documented spatial variability of soil and crop characteristics and PA management recommendations are provided. The time factor or temporal variability is less documented and mostly ignored. In reality, spatial variability of soil and crop characteristics is dynamic within each growing season and between growing seasons. Temporal variability occurs both intraseasonally and interseasonally. Soil water content, nitrogen (N) status, and climatic parameters change day to day within a season, whereas crop yield and weed infestation patterns vary from season to season (Zhang et al. 2003). Temporal variability in crop yield or some soil characteristics at the within-field scale is often larger in magnitude than spatial variability. This will increase the risk of economically and environmentally inappropriate actions if PA practices are solely based on spatial information (Whelan and McBratney 2000).

Incorporating the temporal variability and applying crop inputs and management at the right time is a key component of PA. Management considerations such as scouting for pests and diseases, application of chemical control, fertilizer application, and irrigation scheduling is known to be more effective and responsive when applied at a specific growth stage (Darby and Lauer 2000; Kranz et al. 2008; Specht et al. 1986; Torrion et al. 2011; Wise et al. 2011). When a crop input is suboptimal, particularly supplemental irrigation water, prioritization of such application can increase water use efficiency (WUE). Under extreme water supply deficit, applying the limited irrigation water to only the productive portion of the field was reported to be economical instead of delivering irrigation to the entire field by variable irrigation (Nair et al. 2012). The economic return of this approach can be attributed to a reduction of fuel and man-hours while increasing yield on the more productive part of a field.

Implementing timely irrigation in a variable rate manner requires intensive temporal data about soil moisture and plant conditions at different parts of the field. Recent advances in sensor and wireless radio frequency (RF) technologies along with the Internet offer great opportunities for development and application of sensor systems for agriculture (Pierce and Elliott 2008). Vellidis et al. (2008) discussed linking soil moisture sensors to RF identification (RFID) tags. Data transmitted to a local receiver monitors variable water needs of crops within fields. Camilli et al. (2007) proposed to apply wireless sensor networks (WSNs) consisting of sensor nodes to continuously measure soil moisture, temperature, solar radiation, and other environmental factors. For example, the granular matrix seems to be preferred due to affordability, life of operation, and ease of installation and maintenance (Irmak et al. 2014). The relatively low cost of the devices allows the installation of a dense population of nodes to adequately represent the variability present in the environment of the field. The real-time information obtained by the sensors from the fields can provide a solid base for farmers to adjust strategies at any time (Hwang et al. 2010).

Another way of handling temporal variability is through the use of crop models to predict critical stages for optimal timing of crop inputs and management. Commonly, these growth stages are simulated with accumulation of growing degree units and often integrated with environmental factors, such as stress, soil types, and microenvironment, as well as crop variety and management. Predicting crop stages requires historical and current weather data. To simplify the access of crop models to farmers, many models or tools are posted on the internet. One such example is the Nebraska *SoyWater* program (http://hprcc3.unl.edu/soywater) that integrates the SoySim model (Setiyono et al. 2010). In this model, daily crop water use is calculated and used for the soil-water balance to determine the daily irrigation requirement. When the occurrence of crop stages is known, it assists researchers, agronomists, and farmers to have a lead time planning for precise application of farm inputs (Torrion et al. 2011).

14.3 DATA INTEGRATION AND ANALYSIS

14.3.1 Data Aggregation and Integration

Geospatial data of different sources and formats at different scales, such as soil physical and chemical properties, crop development, remote sensing images, yield data, and as-applied planting data, are often stored in pieces without systematic integration. This limits data access and data use efficiency, potentially resulting in inadequate management decisions based on incomplete information.

The GIS and spatial databases play a key role in organizing and integrate different layers of data and information. A GIS is a computer-based system for capturing, storing, analyzing, and managing data and associated attributes that are spatially referenced to the earth. One of the advantages of GIS is its capability of overlaying different information and data and relating them in the same spatial context. For example, yield data may be compared with other data layers, such as the soil test data, landscape position, remote sensing crop canopy, soil electrical conductivity, and perhaps previous years' yield data. This may reveal some information as to why yields are high in one location and low in other locations. Based on the relationship of these different layers, better decisions can be made on site-specific application of fertilizers, water, and other crop inputs.

One of the challenges facing PA is data incompatibility. Agronomic data generated from one manufacturer's device often do not match devices from other manufacturers. A prescription map applied on one system cannot be directly transferred to another. Special software and often proprietary software programs have to be used to conduct such transitions. This is not only inconvenient, but also wastes time and increases the cost of operations. Various efforts have been pursued to solve this issue. In 2012, a nonprofit organization, AgGateway, formed a Standardized Precision Ag Data Exchange (SPADE) project to improve data sharing and interoperability (www.aggateway.org). This project is collaboration among agricultural suppliers of hardware, software, inputs, services, implements, and vehicles. Its goals are to establish a framework of standards to simplify data share and exchange among advisors, suppliers, and other partners who provide services with different system components. To further increase data transparency and integration, the Open Ag Data Alliance (OADA) is building open application programming interfaces (APIs) to allow different hardware and software systems to communicate automatically through secure cloud services. In addition, OADA is building guidelines regarding data privacy and use standards to ensure compliance with OADA principles. Many agribusiness companies have agreed or planned to adopt such standards. However, when it comes to implementing the standards, few companies are willing to abandon their proprietary data format. As a result, the process can take a long time.

Systematic approaches on data integration are greatly needed to store, analyze, and provide suggestions for site-specific management services. Just like an electronic health record (EHR), the digital data of a field can be shared so that any crop advisor can diagnose a field and provide service based on all its previous records. Databases with these shared data and information can help make recommendations on guiding site-specific management. Establishing such databases requires collaboration across different government agencies, research institutions, agribusiness companies, and farmers. The outcome database will not only be in technical formats for academic communities, but also as accessible summaries that crop advisors, policymakers, and producers can use to guide agronomic decisions and create managerial guidelines. An agribusiness, SST Software, is adopting such a systematic approach to build a database that is compatible across different hardware and software industries. They provide services including data storage and data analysis as well as management recommendations.

14.3.2 Spatial Analysis and Geostatistics

Tobler (1970) called this the first law of geography: "everything is related to everything else, but near things are more related than distant things." Agronomic data, such as soil clay content, soil nutrient status, electrical conductivity, and yield, is spatially distributed data that is more related close together than far apart. Classical statistical analysis assumes that data is independent. Geostatistics has been widely adopted in PA because it provides a collection of statistical methods to analyze spatially dependent data. Geostatistics treats a spatial variable as a random variable. For each point of a population, x, there is a series of values for a property, $Z(x)$, and the observed value, $z(x)$, is drawn at random according to a probability distribution function. The series of random variables is a random process and the actual value of Z observed is one number of realizations of that process (Oliver 2010).

Spatial patterns are usually described using the semivariogram, which measures the average dissimilarity between data separated by a distance h. It is calculated as

$$\gamma(h) = \frac{1}{2N}\sum_{i=1}^{N}[Z(x_{\mathrm{i}}) - Z(x_{\mathrm{i}} + h)]^2$$

where, x_{i} is a data location, h is a lag distance between samples, $Z(x_{\mathrm{i}})$ is the data value at the location x, and N is the number of data pairs at distance h. The semivariogram is usually modeled using several functions, which then fit the semivariogram data. Semivariogram models are described with three parameters, range, sill, and nugget (Figure 14.4). Range is the distance over which data are correlated. This means when separation distances are greater than the range, sampled points are no longer spatially correlated (i.e., random). Sill refers to the semivariance at which range is reached. Nugget represents the semivariance at zero separation distance (lag = 0), usually due to errors in sampling, measurement, or other unexplained sources of variance (Isaaks and Srivastava 1989).

There are two main applications of geostatistics in PA: modeling spatial dependence and predicting variable values at unsampled locations. A spatial description combined with good knowledge about the phenomenon can improve understanding the underlying physical mechanisms controlling the spatial patterns (Goovaerts 1998). Evaluating the spatial dependence of soil or plant variables helps to minimize the number of samples to make appropriate estimation of variable distribution without significant loss of information (Vachaud et al. 1985). The characterization of spatial structure has also been used to filter and remove errors in yield monitor data (Ping and Dobermann 2005) and serves as a tool to assess the response of error filtering (Simbahan et al. 2004). In these studies, yield data not spatially correlated within small regions are considered outliers and removed. In addition, the spatial structure of yield data improved after outlier removal. The analysis of spatial variation can also be used to improve analysis of treatment effects in experimental design and analysis. Most variety evaluation trials are analyzed using classical analysis of variance without adequately accounting for spatial variability. Stroup et al. (1994) found the analysis of breeding trials is improved by removing the spatial variance in the data.

The second application of geostatistics in PA is the spatial prediction of a variable at unsampled locations. Kriging is a generic term for a set of generalized least-squares regression algorithms for spatial prediction (Goovaerts 1998; Webster and Oliver 2001). The most commonly used kriging methods are the ordinary kriging and cokriging. Ordinary kriging unbiasedly estimates an unknown value as a linear combination of neighboring observations using weights applied to each observation. These weights are chosen to minimize the estimation error. Ordinary cokriging is an extension of the ordinary kriging method to incorporate additional information of another variable.

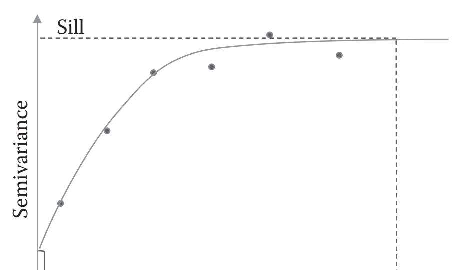

FIGURE 14.4 Example of a conceptual semivariogram (dots are observed semivariance).

For example, the sparsely distributed soil test P levels can be better estimated by taking into account secondary information originating from other correlated information, such as EC_a.

There are several challenges encountering geostatistics application in PA. Many professionals working in PA do not have a strong geostatistical background. More training opportunities or university courses need to be readily available for them to improve their knowledge and skills in geostatistics. Geostatistical software programs need to provide better documentation and explanation for each analysis method, instead of only providing a black-box-based final result. The time factor in PA is also not being handled adequately. Fundamental research is required to develop geostatistics to analyze the integrated dynamics of spatial and temporal variability (Schueller 2010).

14.3.3 Decision Support and Management Zones

The purpose of collecting information and data is to have site-specific management. Dividing the field into subfields or management zones is a natural first step. Management zones are field areas possessing approximately homogenous attributes in landscape and soil condition and are used as the smallest units for site-specific management, such as variable rate fertilizer application (Doerge 1999). Creation of management zones provides a convenient means of capturing the spatial distribution of yield-influencing factors in a season. Two criteria may be used to evaluate the appropriateness of management zones: (1) yield differences between zones should be substantially greater than those within zones, and (2) the major factors that influence yield within a zone must be approximately homogeneous (Plant et al. 1999).

Yield maps have often been used to delineate management zone delineation. For example, Kitchen et al. (1995) used a classified corn yield map from the previous year to determine yield potential zones for variable rate application of N fertilizer. The application of yield maps to identify zones is challenged by spatial and temporal variation in yield because it is affected by many interacting factors (Huggins and Alderfer 1995). Soil physical and chemical properties have been extensively used to identify management zones, including soil texture, bare soil brightness, and apparent electrical conductivity. Mzuku et al. (2005) delineated management zones based on bare soil aerial imagery, a farmer's perception of field topography, and past crop and management practices. Georeferenced apparent electrical conductivity in a field provides spatial information about soil salinity, texture, and water content, enabling producers to identify zones for particular management practices (Bullock and Bullock 2000; Clay et al. 2001; Corwin and Lesch 2003).

Many of the previous studies about the delineation of management zones did not integrate crop yield and soil properties. According to Lund et al. (2000), producers are unwilling to adjust inputs on different yielding areas until they have some evidence about the underlying soil properties. Any management zones produced without considering crop yield would be of little value (Doerge 1999). Delineation of management zones using yield maps along with soil properties can be used to directly associate crop production with soil properties so that management zones are more meaningful and interpretable. However, because yield varies from year to year due to complex factors, potential management zones are likely to be different from year to year. Taking temporal stability into account allows better management of weather and climatic risk (McBratney et al. 2005). Potentially stable management zones in a field provide important decision support for crop producers to apply inputs such as water, nitrogen, and chemicals in a site-specific manner.

Producers or crop advisors must identify the purpose of site-specific management before proceeding to collecting data and preparing for management zones delineation. Factors such as terrain elevation, soil physical properties, and soil nutrient levels have the most direct impact on yield, and hence should be included for zone delineation. Data with a stable temporal pattern can be the most cost efficient in applying management zones for site-specific management. Topography, EC_a, soil physical properties, or multi-year yield data along with supplemental information (e.g., soil electrical conductivity or elevation) is recommended for zone delineation in order to identify consistent yield patterns (Ortiz et al. 2011). For many in-season management practices, such as application of growth regulator and defoliants, a high resolution remote sensing image is a perfect choice for management zone delineation. A general guideline is provided in Table 14.1 to associate the agricultural inputs suitable for site-specific management and the suggested data for management zone delineation.

Management zone delineation is usually a multivariate clustering procedure to divide the field into different zones based on the input variables. These clustering methods used in agriculture include hierarchical agglomerative clustering, fuzzy clustering, hierarchical divisive clustering, and Kohonen self-organizing feature maps (Tiwari and Misra 2011). Fridgen et al. (2004) developed a software program, Management Zone Analyst (MZA), to automatically analyze data and output clusters using the fuzzy c-means theory. Fuzzy c-means applies a weighting exponent to control the degree of membership sharing between clusters (Bezdek 1981). This software program not only classifies the data set into zones, but also suggests the best number of zones that should be created.

Management zones are delineated using historical data. Therefore, the user needs to assess the performance of management zones or management strategies for a field during the growing season and after the season, depending on the management purpose. A grower can make a historical comparison to yield or profitability attained with a previous variable rate or uniform rate input strategy. Another method will be to directly compare the value of two management zone strategies using multiple side-by-side comparisons using yield monitor data (Doerge 1999). Depending on the comparison results, it may be necessary to adjust the management zones by combining or splitting for the next crop or season (USDA-NRCS 2010).

With the increasing applications of real time sensors, management zones may eventually become obsolete for some variable rate applications and management. One such sensor, for example, is the GreenSeeker (Trimble Navigation, Sunnyvale, CA) mounted on the sprayer. It uses an active light source to measure spectral reflectance from the crop canopy to calculate normalized difference vegetation index (NDVI) for determining the amount of N required at different locations of a field. As the sprayer is traveling and recording this data, a prescription is created to apply N in the right amount needed in that particular portion of the field. Another product, the WeedSeeker spray system, applies a similar principle to sense the presence of weeds and triggers spray nozzles to deliver a precise amount of chemical to spray the weeds. Maleki et al. (2008) demonstrated the feasibility

TABLE 14.1
Agricultural Inputs Suitable for Variable Rate Application and Data Used to Delineate Management Zones for these Inputs

Input Suitable for Variable Rate Application	Data Used to Delineate Management Zones
Seeding rate	Soil EC_a, topography, soil survey, historic yield data, soil organic matter
Nematicides	Soil EC_a, topography
Lime	Soil EC_a, grid or zone-sampled soil pH, buffer pH, soil survey, topography, bare soil imagery, historic yield data
Nitrogen	Soil texture, soil organic matter, soil color imagery, crop spectral reflectance using sensors like GreenSeeker or Crop Circle, yield data
Other nutrients	Soil EC_a, crop spectral reflectance, CEC, soil survey
Irrigation	Soil EC_a, soil color imagery, yield data, canopy imagery, soil survey, farmer's knowledge
Plant growth regulators, defoliants	Plant spectral reflectance of visible and infrared bands from in-season aerial or high-resolution satellite imagery

Source: Adapted with permission from Ortiz, B.V. et al., Management zones II–basic steps for delineation. Precision Agriculture Series–Timely Information. Agriculture, Natural Resources & Forestry. Alabama Cooperative Extension System, 2011.

and effectiveness of on-the-go variable-rate P fertilizer application using a visible (VIS) and near-infrared (NIR) soil sensor for on-the-go measurement of soil P. Variable rate irrigation has been also proved feasible without using management zones. For example, Kim et al. (2008) proposed a real-time variable rate irrigation system consisting of a WSN, software for real-time information, and an irrigation control system. The system site-specifically operates individual sprinklers to apply a specified amount of water based on the real-time information collected by the sensors distributed across the field.

14.4 BIG DATA CHALLENGE

Big data refers to data sets that cannot be processed, analyzed, and managed using traditional analyzing algorithms or tools due to its large volume, structural variety, and high data accessing and retrieving velocity (Fan and Bifet 2012; Zikopoulos et al. 2011). Data preprocessing and advanced multivariate data analysis techniques become crucial in analyzing high-dimensional data and building accurate analysis pipeline for applications in PA. The accumulated massive amount of agricultural data will require robust information technology (IT) infrastructure and complicated data-analyzing algorithms to develop weather forecast models, yield prediction algorithms, and offer decision supporting tools to farmers. For example, John Deere's FarmSight and Pioneer's Field360 are among the tools that provide detailed management prescriptions based on the multiple sources of data collected including those by farmers from their farm equipment.

14.4.1 Data Analytics Using Machine Learning

Machine learning is a scientific and engineering discipline based on programming computers to optimize a performance criterion using example data or previous experience (Alpaydin 2004). Machine learning serves as a proxy for handling complex data gathered across large spatial and temporal scales. The kernel machine method developed over the last decade is a powerful machine-learning technique that has found wide applications in biological and agricultural studies (Cui et al. 2014a,b; Mirik et al. 2014a,b). To date, the most popular machine-learning techniques widely used by researchers and industry teams include support vector machine (SVM), relevance vector machine (RVM), and artificial neural networks (ANN).

We focus our discussion on SVM and RVM in this section because of their superior performances and great adaptability. SVM is a popular kernel machine method used in bioinformatic and multispectral and hyperspectral image analysis (Bazi and Melgani 2006; Huang et al. 2002; Melgani and Bruzzone 2004). In a typical machine-learning paradigm, a variable of interest is treated as one dimension in a multidimensional input space, termed a feature. Take remote sensing data analysis for example; each spectral band could be treated as a feature and the corresponding reflectance value is the feature value. Each pixel (sample) associated with a specific geospatial location will be represented by its reflectance values (feature values) of different spectral bands. Figure 14.5 illustrates a simple scenario of a two-dimensional–two-class classification problem using SVM. The greatest advantage of SVM classification is that once the classification model has been constructed, only the samples that lie on the margin (called support vectors [SVs]) determine the position of the decision boundary, which provides the maximum margin that separates those two classes with minimized measures of errors.

The major trend of machine learning has led its way into the application of Gaussian processes. RVM based on probabilistic predictions introduced by Tipping (2001) is regarded as a successful alternative to SVM. RVM utilizes fewer kernel functions and generates fewer relevance vectors (RVs) compared with the kernel functions and SV generated by SVM (Cui et al. 2014b; Tipping 2001). Therefore, the classification processes could be finished much faster than SVM. In terms of classification accuracy, performances of RVM are similar to SVM (Demir and Ertürk 2007; Psorakis et al. 2010).

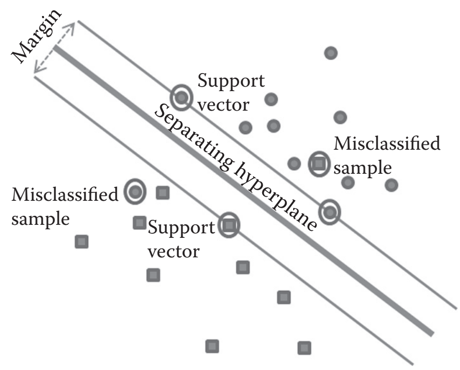

FIGURE 14.5 Linear support vector machine with two-dimensional input space and two classes.

14.4.2 Application of Big Data and Algorithms in Agriculture

A large amount of agronomic studies have applied data mining approaches on big data sets (Mucherino et al. 2009). Different machine-learning methods have been widely used in detecting biotic stresses in agricultural production, including weed encroachment (Ahmed et al. 2012; Longchamps et al. 2010; Lopez-Granados et al. 2008; Mirik et al. 2014a,b; Nieuwenhuizen et al. 2007), drought stress (Behmann et al. 2014), herbicide injury (Zhao et al. 2014), and plant diseases (Berdugo et al. 2014; Bock et al. 2010; Mahlein et al. 2012a,b; Rumpf et al. 2010). Gianpiero et al. (2014) conducted a study using discrete optimization procedures for automatic error correction on a large-scale data set in Italy and showed satisfactory results. Ananthara et al. (2013) proposed an improved crop yield prediction model called CRY, which used the bee hive clustering approach based on a big crop yield database (Crop Knowledge Base) in India, and compared CRY with other popular big-data analyzing algorithms. Likewise, another study conducted by Ramesh and Vardhan (2013) compared various data mining techniques such as K-Means, K-Nearest Neighbor, ANN, and SVM for crop yield prediction on a large data set obtained from 1965 to 2009 in the East Godavari District of Andhra Pradesh in India. Most recently, Holman et al. (2014) conducted a study using both Gaussian process models and ANN to estimate reference evapotranspiration (ET) for irrigation management in the Texas High Plains. The entire data set consists of daily reference ET and other critical climatic variable data over a period of 10 years (2001–2010) across 15 different ET network stations and five national weather service weather stations in the Texas High Plains. The results indicated significant predicting accuracy of reference ET using sophisticated machine-learning algorithms that could be used for irrigation management. In another study, SVM-based classification and regression approaches were both applied to surface water quality monitoring and analysis based on a data set collected from 1500 water samples representing 10 different sites over 10 years (Singh et al. 2011). The results indicated adequacy of constructed models and great predictive capacities.

14.5 SUMMARY AND OUTLOOK

PA has been increasingly dependent on collecting, analyzing, and utilizing data and information for optimal management practices. Various sensors along with platforms have been developed to efficiently and effectively collect spatial data and information of soil properties, crop growth status, crop yield, and environmental factors. With the increasing capabilities of hardware and software of these sensors, extremely large volumes of data and information are being collected, archived, and processed. Effectively and efficiently storing, processing, and analyzing big data for site-specific application and management is becoming one of the greatest challenges facing PA. Additionally, big data acquisition can also cause data ownership controversial issues. Who owns the data? How will the data be used? Proprietary data sharing rules and the law and unintended infringement are premature. However, with the advancement of computational science and innovation of PA, new management tools and data sharing platforms will be developed for the agricultural community to share big data, while protecting data privacy.

The spatial and temporal scales of data and information in PA are continuing to increase thanks to technological advancements. We are managing fields at a finer scale, eventually plant-by-plant, or even at subplant levels (i.e., managing roots or leaves), incorporating data and information from advanced sensors. Field management at this scale requires intensive data and information input. Decision support systems incorporating information from various sensors enable continuous monitoring the conditions of soil and crop growth, allowing farmers to adjust site-specific management at any time.

An integrated and systematic approach of data management and field operations is becoming a trend in PA. We expect more cloud-based data processing to increase cost effectiveness, agility, productivity, and scalability. At the same time, systematic approaches on data integration is becoming urgent and requires collaboration across different government agencies, research institutions, agribusiness companies, and farmers. New methods of data analytics such as machine learning are becoming popular. Computer technologies continue to evolve. Moore's law continues to predict the future storage and computational capabilities of computers. This eventually allows instantaneous decision support based on big data and information and on-the-fly site-specific management and application. High-speed wireless Internet allows virtually all devices and equipment connected for integrated information transfer, on-the-go data processing, and real-time decision support, leading to much more efficient farming operations.

It is also noteworthy that despite technological advancements in computer science, engineering, sensing technologies, and advanced analytics tools, the science involved in PA is still lagging behind the requirements of the technologies. The advancement of technology has to be integrated with agronomic background science to realize the full potential of PA. More interdisciplinary and innovative studies are needed to enhance the understanding of the interactions among crop growth, soil, and environmental factors. Based on this better understanding, more effective management of the soil and crops can further improve agricultural production while protecting the environment.

REFERENCES

Adamchuk, V.I., J.W. Hummel, M.T. Morgan, and S.K. Upadhyaya. 2004. On-the-go soil sensors for precision agriculture. *Computers and Electronics in Agriculture* 44:71–91.

Ahmed, F., H.A. Al-Mamun, H.A.S.M. Bari, E. Hossain, and P. Kwan. 2012. Classification of crops and weeds from digital images: A support vector machine approach. *Crop Protection* 40:98–104.

Alpaydin, E. 2004. *Introduction to Machine Learning (Adaptive Computation and Machine Learning)*, Cambridge, MA: MIT Press.

Ananthara, M.G., T. Arunkumar, and R. Hemavathy. 2013. CRY–An improved crop yield prediction model using bee hive clustering approach for agricultural data sets. *IEEE International Conference on Pattern Recognition, Informatics and Medical Engineering*, 473–478.

Andujar, D., A. Escola, J.R. Rosell-Polo, C. Fernandez-Quintanilla, and J. Dorado. 2013. Potential of a terrestrial LiDAR-based system to characterize weed vegetation in maize crops. *Computers and Electronics in Agriculture* 92:11–15.

Bausch, W.C., K. Diker, R. Khosla, and J.F. Paris. 2004. Estimating corn nitrogen status using ground-based and satellite multispectral data. In *Optical Science and Technology, the SPIE 49th Annual Meeting*, 489–498. International Society for Optics and Photonics.

Bazi, Y., and F. Melgani. 2006. Toward an optimal SVM classification system for hyperspectral remote sensing images. *IEEE Transactions of Geosciences and Remote Sensing* 44:3374–3385.

Behmann, J., J. Steinrücken, and L. Plümer. 2014. Detection of early plant stress responses in hyperspectral images. *ISPRS Journal of Photogrammetry and Remote Sensing* 93:98–111.

Berdugo, C., R. Zito, S. Paulus, and A.K. Mahlein. 2014. Fusion of sensor data for the detection and differentiation of plant diseases in cucumber. *Plant Pathology* doi: 10.1111/ppa.12219.

Bezdek, J.C. 1981. *Pattern Recognition with Fuzzy Objective Function Algorithms*. New York: Plenum Press.
Blackmore, B.S., and C.J. Marshall. 1996. Yield mapping: Errors and algorithms. In *Precision Agriculture*, ed. P.C. Robert, R.H. Rust, and W.E. Larson, 403–415. Madison, WI: ASA, CSSA, and SSSA.
Blackmore, B.S. 2000. The interpretation of trends from multiple yield maps. *Computer and Electronics in Agriculture* 26:37–51.
Bock, C.H., G.H. Poole, P.E. Parker, and T.R. Gottwald. 2010. Plant disease severity estimated visually, by digital photography and image analysis, and by hyperspectral imaging. *Critical Reviews in Plant Sciences* 29:59–107.
Bronson, K.F., J.W. Keeling, J.D. Booker, T.T. Chua, T.A. Wheeler, R.K. Boman, and R.J. Lascano. 2003. Influence of landscape position, soil series, and phosphorus fertilizer on cotton lint yield. *Agronomy Journal* 95:949–957.
Bullock, D.S., and D.G. Bullock. 2000. Economic optimality of input application rates in precision farming. *Precision Agriculture* 2:71–101.
Camilli, A., C.E. Cugnasca, A.M. Saraiva, A.R. Hirakawa, and P.L.P. Correa. 2007. From wireless sensors to field mapping: Anatomy of an application for precision agriculture. *Computers and Electronics in Agriculture* 58:25–36.
Clay, D.E., J. Change, D.D. Malo, C.G. Carlson, C. Reese, S.A. Clay, M. Ellsbury, and B. Berg. 2001. Factors influencing spatial variability of soil apparent electrical conductivity. *Communications in Soil Science and Plant Analysis* 32:2993–3008.
Clevers, J.G.P.W., and H.J.C. Van Leeuwen. 1996. Combined use of optical and microwave remote sensing data for crop growth monitoring. *Remote Sensing of Environment* 561:42–51.
Corwin, D.L., and S.M. Lesch. 2003. Application of soil electrical conductivity to precision agriculture: Theory, principles, and guidelines. *Agronomy Journal* 95:455–471.
Corwin, D.L., S.M. Lesch, P.J. Shouse, R. Soppe, and J.E. Ayars. 2003. Identifying soil properties that influence cotton yield using soil sampling directed by apparent soil electrical conductivity. *Agronomy Journal* 95:352–364.
Cui, S., E. Youn, J. Lee, and S.J. Maas. 2014a. An improved systematic approach to predicting transcription factor target genes using support vector machine. *PLoS ONE* 9(4):e94519. doi:10.1371/journal.pone.0094519.
Cui, S., N. Rajan, S.J. Maas, and E. Youn. 2014b. An automated soil line identification method using relevance vector machine. *Remote Sensing Letters* 5:175–184.
Curran, P.J., P.M. Atkinson, G.M. Foody, and E.J. Milton. 2000. Linking remote sensing, land cover and disease. *Advances in Parasitology* 47:37–80.
Darby, H., and J. Lauer. 2000. Plant physiology: Critical stages in the life of a corn plant. http://corn.agronomy.wisc.edu/Management/pdfs/CriticalStages.pdf (accessed October 1, 2014).
Daughtry, C.S.T., C.L. Walthall, M.S. Kim, E.B. De Colstoun, and J.E. McMurtrey III. 2000. Estimating corn leaf chlorophyll concentration from leaf and canopy reflectance. *Remote Sensing of Environment* 74:229–239.
Delegido, J., J. Verrelst, C.M. Meza, J.P. Rivera, L. Alonso, and J. Moreno. 2013. A red-edge spectral index for remote sensing estimation of green LAI over agroecosystems. *European Journal of Agronomy* 46:42–52.
Demir, B., and S. Ertürk. 2007. Hyperspectral image classification using relevance vector machines. *IEEE Geoscience and Remote Sensing Letters* 4:586–590.
Dobermann, A., J.L. Ping, V.I. Adamchuk, G.C. Simbahan, and R.B. Ferguson. 2003. Classification of crop yield variability in irrigated production fields. *Agronomy Journal* 95:1105–1120.
Doerge, T.A. 1999. Site-specific management guidelines: Management zone concepts. http://www.ipni.net/publication/ssmg.nsf/0/C0D052F04A53E0BF852579E500761AE3/$FILE/SSMG-02.pdf (accessed October 1, 2014).
Eitel, J.U., T.S. Magney, L.A. Vierling, T.T. Brown, and D.R. Huggins. 2014. LiDAR based biomass and crop nitrogen estimates for rapid, non-destructive assessment of wheat nitrogen status. *Field Crops Research* 159:21–32.
Fan, W., and A. Bifet. 2012. Mining big data: Current status, and forecast to the future. *ACM SIGKDD Explorations Newsletter* 14:1–5.
FAO. 2011. The state of the world's land and water resources for food and agriculture (SOLAW)–Managing systems at risk. Rome: Food and Agriculture Organization of the United Nations.
Foth, H.D., and B.G. Ellis. 1988. *Soil Fertility*. New York: Wiley.
Fridgen, J.J., N.R. Kitchen, K.A. Sudduth, S.T. Drummond, W.J. Wiebold, and C.W. Fraisse. 2004. Management Zone Analyst (MZA): Software for subfield management zone delineation. *Agronomy Journal* 96:100–108.

Galzki, J.C., A.S. Birr, and D.J. Mulla. 2011. Identifying critical agricultural areas with three-meter LiDAR elevation data for precision conservation. *Journal of Soil and Water Conservation Society* 66:423–430.

Gianpiero, B., B. Renato, and R. Alessandra. 2014. Balancing of agricultural census data using discrete optimization. *Optimization Letters* 8:1553–1565.

Gitelson, A.A., and M.N. Merzlyak. 1996. Signature analysis of leaf reflectance spectra: Algorithm development for remote sensing of chlorophyll. *Journal of Plant Physiology* 148:494–500.

Goovaerts, P. 1998. Geostatistical tools for characterizing the spatial variability of microbiological and physico-chemical soil properties. *Biology and Fertility of Soils* 27:315–334.

Griffin, T.W. 2009. Farmers' use of yield monitors. http://www.uaex.edu/publications/pdf/FSA-36.pdf (accessed October 1, 2014).

Guo, W., S.J. Maas, and K.F. Bronson. 2012. Relationship between cotton yield and soil electrical conductivity, topography, and Landsat imagery. *Precision Agriculture* 13:678–692.

Holman, D., M. Sridharan, P. Gowda et al. 2014. Gaussian process models for reference ET estimation from alternative meteorological data sources. *Journal of Hydrology* 517:28–35.

Huang, C., L.S. Davis, and J.R.G. Townshend. 2002. An assessment of support vector machines for land cover classification. *International Journal of Remote Sensing* 23:725–749.

Huggins, D.R., and R.D. Alderfer. 1995. Yield variability within a long-term corn management study: Implications for precision farming. In *Site-Specific Management for Agricultural Systems*, ed. P.C. Robert, R.H. Rust, and W.E. Larson, 417–426. Madison, WI: ASA, CSSA, SSSA.

Hwang, J., C. Shin, and H. Yoe. 2010. Study on an agricultural environment monitoring server system using wireless sensor networks. *Sensors* 10:11189–11211.

Iqbal, J., J.J. Read, A.J. Thomasson, and J.N. Jenkins. 2005. Relationships between soil–landscape and dryland cotton lint yield. *Soil Science Society America Journal* 69:872–882.

Irmak, S., J.O. Payero, B. VandeWalle, J. Rees, and G. Zoubek. 2014. Principles and operational characteristics of watermark granular matrix sensors to measure soil water status and its practical applications for irrigation management in various soil textures. University of Nebraska Extension EC783.

Isaaks, E., and R. Srivastava. 1989. *An Introduction to Applied Geostatistics*. New York: Oxford University Press.

Johnson, C.K., J.W. Doran, H.R. Duke, B.J. Wienhold, K.M. Eskridge, and J.F. Shanahan. 2001. Field scale electrical conductivity mapping for delineating soil conditions. *Soil Science Society America Journal* 65:1829–1837.

Kaspar, T.C., T.S. Colvin, D.B. Jaynes et al. 2003. Relationship between six years of corn yields and terrain attributes. *Precision Agriculture* 4:87–101.

Kelly, M., and R. Guo. 2007. Integrated agricultural pest management through remote sensing and spatial analyses. In *General Concepts in Integrated Pest and Disease Management*. ed. A. Ciancio and K.G. Mukerji, 191–207. Netherlands: Springer.

Khosla, R. 2010. Precision agriculture: Challenges and opportunities in a flat world. *Proceedings of the 19th World Congress of Soil Science, Soil Solutions for a Changing World*, 1–6 August 2010, Brisbane, Australia.

Kim, Y., R.G. Evans, and W.M. Iversen. 2008. Remote sensing and control of an irrigation system using a distributed wireless sensor network. *IEEE Transactions on Instrumentation and Measurement* 57:1379–1387.

Kitchen, N.R., D.F. Hughes, K.A. Sudduth, and S.J. Birrel. 1995. Comparison of variable rate to single rate nitrogen fertilizer applications: Corn production and residual soil NO_3-N. In *Proceedings of Site-Specific Management for Agricultural Systems*, ed. P.C. Robert, 427–439. Madison, WI: ASA, CSSA, SSSA.

Kranz, W.L., S. Irmak, S.J. van Donk, C.D. Yonts, and D.L. Martin. 2008. Irrigation management for corn. NebGuide G1850. University of Nebraska Extension.

Kravchenko, A.N., and D.G. Bullock. 2000. Correlation of corn and soybean grain yield with topography and soil properties. *Agronomy Journal* 92:75–83.

Li, H., R.J. Lascano, J. Booker, T. Wilson, and K.F. Bronson. 2001. Cotton lint yield variability in a heterogeneous soil on a landscape-scale. *Soil & Tillage Research* 58:245–258.

Liaghat, S., and S.K. Balasundram. 2010. A review: The role of remote sensing in precision agriculture. *American Journal of Agricultural and Biological Sciences* 5:50–55.

Longchamps, L., B. Panneton, G. Samson, G. Leroux, and R. Thériault. 2010. Discrimination of corn, grasses and dicot weeds by their UV-induced fluorescence spectral signature. *Precision Agriculture* 11:181–197.

Lopez-Granados, F., J.M. Pena-Barragan, M. Jurado-Exposita, M. Francisco-Fernandez, R. Cao, and A. Alosno-Betanzos. 2008. Multispectral classification of grass weeds and wheat (*Triticum durum*) using linear and nonparametric functional discriminant analysis and neural networks. *Weed Research* 48:28–37.

Lund, E.D., C.D. Christy, and P.E. Drummond. 2000. Using yield and soil electrical conductivity (EC) maps to derive crop production performance information. *Proceedings of the 5th International Conference on Precision Agriculture*, Bloomington, MN, July 16–19, 2000.

Maas, S.J. 1998. Estimating cotton ground cover from remotely sensed scene reflectance. *Agronomy Journal* 90:384–388.

Machado, S., E.D. Bynum, Jr., T.L. Archer et al. 2002. Spatial and temporal variability of sorghum grain yield: Influence of soil, water, pests, and disease relationships. *Precision Agriculture* 3:389–406.

Mahlein, A.K., E.C. Oerke, U. Steiner, and H.W. Dehne. 2012a. Recent advances in sensing plant diseases for precision crop protection. *European Journal of Plant Pathology* 133:197–209.

Mahlein, A.K., U. Steiner, C. Hillnhütter, H.W. Dehne, and E.C. Oerke. 2012b. Hyperspectral imaging for small-scale analysis of symptoms caused by different sugar beet diseases. *Plant Methods* 8:3, doi:10.1186/1746-4811-8-3.

Mahler, R.L., D.F. Benzdicek, and R.E. Witters. 1979. Influence of slope position on nitrogen fixation and yield of dry peas. *Agronomy Journal* 71:348–354.

Maleki, M.R., A.M. Mouazen, B.D. Ketelaere, H. Ramon, and J.D. Baerdemaeker. 2008. On-the-go variable-rate phosphorus fertilisation based on a visible and near-infrared soil sensor. *Biosystems Engineering* 99:35–46.

McBratney, A., B. Whelan, T. Ancev, and J. Bouma. 2005. Future directions of precision agriculture. *Precision Agriculture* 6:7–23.

McKinion, J.M., J.N. Jenkins, D. Akins et al. 2001. Analysis of a precision agriculture approach to cotton production. *Computers and Electronics in Agriculture* 32:213–228.

Melgani, F., and L. Bruzzone. 2004. Classification of hyperspectral remote sensing images with support vector machines. *IEEE Transactions of Geosciences and Remote Sensing* 42:1778–1790.

Miller, M.P., M.J. Singer, and D.R. Nielsen. 1988. Spatial variability of wheat yield and soil properties on complex hill. *Soil Science Society of America Journal* 52:1133–1141.

Mirik, M., R.J. Ansley, K. Steddom et al. 2014a. High spectral and spatial resolution hyperspectral imagery for quantifying Russian wheat aphid infestation in wheat using the constrained energy minimization classifier. *Journal of Applied Remote Sensing* 8(1):083661.

Mirik, M., Y. Emendack, A. Attia et al. 2014b. Detecting musk thistle (*Carduus nutans*) infection using a target recognition algorithm. *Advances in Remote Sensing* 3:95–105.

Moran, M.S., Y. Inoue, and E.M. Barnes. 1997. Opportunities and limitations for image-based remote sensing in precision crop management. *Remote Sensing of Environment* 61:319–346.

Mucherino, A., P. Papajorgji, and P.M. Pardalos. 2009. A survey of data mining techniques applied to agriculture. *Operational Research* 9:121–140.

Mulla, D.J. 2013. Twenty five years of remote sensing in precision agriculture: Key advances and remaining knowledge gaps. *Biosystems Engineering* 114:358–371.

Mzuku, M., R. Khosla, R. Reich, D. Inman, F. Smith, and L. MacDonald. 2005. Spatial variability of measured soil properties across site-specific management zones. *Soil Science Society of America Journal* 69:1572–1579.

Nair, S., S.J. Maas, C. Wang, and S. Mauget. 2012. Optimal field partitioning for center-pivot-irrigated cotton in the Texas High Plains. *Agronomy Journal* 105:124–133.

National Academy of Sciences. 1997. *Precision Agriculture in the 21st Century: Geospatial and Technologies in Crop Management*. Washington, DC: National Academy Press.

Nieuwenhuizen, A.T., L. Tang, J.W. Hofstee, J. Müller, and E.J. Van Henten. 2007. Colour based detection of volunteer potatoes as weeds in sugar beet fields using machine vision. *Precision Agriculture* 8:267–278.

Oliver, M.A. 2010. An overview of geostatistics and precision agriculture. In *Geostatistical Applications for Precision Agriculture*, ed. M.A. Oliver, 1–34. New York: Springer.

Ortiz, B.V., J. Shaw, J.P. Fulton, and A. Winstead. 2011. Management zones II–Basic steps for delineation. Precision Agriculture Series–Timely Information. Agriculture, Natural Resources & Forestry. Alabama Cooperative Extension System.

Pacheco, A., A. Bannari, J.C. Deguise, H. McNairn, and K. Staenz. 2001. Application of hyperspectral remote sensing for LAI estimation in precision farming. *Proceedings of 23rd Canadian Symposium on Remote Sensing*, 281–287.

Papadavid, G., D. Fasoula, M. Hadjimitsis, P.S. Perdikou, and D.G. Hadjimitsis. 2013. Image based remote sensing method for modeling black-eyed beans (Vigna unguiculata) Leaf Area Index (LAI) and Crop Height (CH) over Cyprus. *Central European Journal of Geosciences* 5:1–11.

Pierce, F.J., and P. Nowak. 1999. Aspects of precision agriculture. *Advances in Agronomy* 67:1–85.

Pierce, F.J., and T.V. Elliott. 2008. Regional and on-farm wireless sensor networks for agricultural systems in Eastern Washington. *Computers and Electronics in Agriculture* 61:32–43.

Ping, J.L., and A. Dobermann. 2005. Processing of yield map data. *Precision Agriculture* 6:193–212.

Plant, R.E., A. Mermer, G.S. Pettygrove et al. 1999. Factors underlying grain yield spatial variability in three irrigated wheat fields. *Transactions of the ASABE* 42:1187–1202.

Plant, R.E. 2001. Site-specific management: The application of information technology to crop production. *Computers and Electronics in Agriculture* 30:9–29.

Prabhakar, M., Y.G. Prasad, and M.N. Rao. 2012. Remote sensing of biotic stress in crop plants and its applications for pest management. In *Crop Stress and its Management: Perspectives and Strategies*, ed. B. Venkateswarlu, A.K. Shanker, C. Shanker, and M. Maheswari, 517–545. Netherlands: Springer.

Psorakis, I., T. Damoulas, and M.A. Girolami. 2010. Multiple relevance vector machines: Sparsity and accuracy. *IEEE Transaction on Neural Networks* 21:1588–1598.

Qin, Z., and M. Zhang. 2005. Detection of rice sheath blight for in-season disease management using multispectral remote sensing. *International Journal of Applied Earth Observation and Geoinformation* 7:115–128.

Rajan, N., and S.J. Maas. 2009. Mapping crop ground cover using airborne multispectral digital imagery. *Precision Agriculture* 10:304–318.

Rajan, N., N. Puppala, S. Maas, P. Payton, and R. Nuti. 2014. Aerial remote sensing of peanut ground cover. *Agronomy Journal* 106:1358–1364.

Ramesh, D., and B.V. Vardhan. 2013. Data mining techniques and applications to agricultural yield data. *International Journal of Advanced Research in Computer and Communication Engineering* 2:3477–3480.

Rhoades, J.D. 1993. Electrical conductivity methods for measuring and mapping soil salinity. *Advances in Agronomy* 49:201–251.

Robert, P.C., R.H. Rust, and W.E. Larson. 1995. *Site-Specific Management for Agricultural Systems*. Madison, WI: ASA, CSSA, and SSSA.

Robert, P.C., R.H. Rust, and W.E. Larson. 1996. *Precision Agriculture*. Madison, WI: ASA, CSSA, and SSSA.

Rumpf, T., A.K. Mahlein, U. Steiner, E.C. Oerke, H.W. Dehne, and L. Plümer. 2010. Early detection and classification of plant diseases with support vector machines based on hyperspectral reflectance. *Computers and Electronics in Agriculture* 74:91–99.

Sakamoto, T., M. Yokozawa, H. Toritani, M. Shibayama, N. Ishitsuka, and H. Ohno. 2005. A crop phenology detection method using time-series MODIS data. *Remote Sensing of Environment* 96:366–374.

Schueller, J.K. 2010. Geostatistics and precision agriculture: A way forward. In *Geostatistical Application for Precision Agriculture*, ed. M.A. Oliver, 305–312. New York: Springer.

Setiyono, T.D., K.G. Cassman, J.E. Specht et al. 2010. Simulation of soybean growth and yield in near-optimal growth conditions. *Field Crops Research* 119:161–174.

Sickle, J.V. 2001. *GPS for Land Surveyors*, 2nd Edition. New York: Taylor & Francis.

Simbahan, G.C., A. Dobermann, and J.L. Ping. 2004. Screening yield monitor data improves grain yield maps. *Agronomy Journal* 96:1091–1102.

Sinai, G., D.Z. Zaslavsky, and P. Golany. 1981. The effect of soil surface curvature on moisture and yield–Beer Sheba observation. *Soil Science* 132:367–375.

Singh, K.P., N. Basant, and S. Gupta. 2011. Support vector machines in water quality management. *Analytica Chimica Acta* 703:152–162.

Specht, J.E., J.H. Williams, and C.J. Weidenbenner. 1986. Differential responses of soybean genotypes subjected to a seasonal soil water gradient. *Crop Science* 26:922–934.

Stroup, W.W., P.S. Baenzier, and D.K. Mulitze. 1994. Removing spatial variation from wheat yield trials: A comparison of methods. *Crop Science* 34:62–66.

Sudduth, K.A., and S.T. Drummond. 2007. Yield Editor: Software for removing errors from crop yield maps. *Agronomy Journal* 99:1471–1482.

Swinton, S.M., and J. Lowenberg-DeBoer. 1998. Evaluating the profitability of site specific farming. *Journal of Production Agriculture* 11:439–446.

Tamura, Y., M. Matsui, L.G. Pagnini, R. Ishibashi, and A. Yoshida. 2002. Measurement of wind-induced response of buildings using RTK-GPS. *Journal of Wind Engineering and Industrial Aerodynamics* 90: 1783–1793.

Tang, H., M. Brolly, F. Zhao et al. 2014. Deriving and validating leaf area index (LAI) at multiple spatial scales through lidar remote sensing: A case study in Sierra National Forest, CA. *Remote Sensing of Environment* 143:131–141.

Thenkabail, P.S., I. Mariotto, M.K. Gumma, E.M. Middleton, D.R. Landis, and K.F. Huemmrich. 2013. Selection of hyperspectral narrowbands (HNBs) and composition of hyperspectral twoband vegetation indices (HVIs) for biophysical characterization and discrimination of crop types using field reflectance and Hyperion/EO-1 data. *IEEE Journal of Selected Topics in Applied Earth Observations and Remote Sensing* 6:427–439.

Thylen, L., P. Jurschik, and D.P.L. Murphy. 1997. Improving the quality of yield data. In *Precision Agriculture: Proceedings of the 1st European Conference on Precision Agriculture*, ed. J.V. Stafford, 743–750. Oxford, UK: BIOS Scientific Publishers Ltd.

Tipping, M.E. 2001. Sparse Bayesian learning and the relevance vector machine. *Journal of Machine Learning Research* 1:211–244.

Tiwari, M., and B. Misra. 2011. Application of cluster analysis in agriculture–A review article. *International Journal of Computer Applications* 36:43–47.

Tobler, W.R. 1970. A computer model simulation of urban growth in the Detroit region. *Economic Geography* 46:234–240.

Torrion, J.A., T.D. Setiyono, K.G. Cassman, and J.E. Specht. 2011. Soybean phenology simulation in the North-Central USA. *Agronomy Journal* 103:1661–1667.

Torrion, J.A., S.J. Maas, W. Guo, J.P. Bordovsky, and A.M. Cranmer. 2014. A three-dimensional index for characterizing crop water stress. *Remote Sensing* 6:4025–4042.

USDA-NRCS. 2010. Precision nutrient management planning. Agronomy Technical Note No. 3, February 2010.

Vachaud, G., A. Passerat de Silans, P. Balabanis, and M.Vauclin. 1985. Temporal stability of spatially measured soil water probability density function. *Soil Science Society of America Journal* 49:822–828.

Vellidis, G., M. Tucker, C. Perry, C. Kvien, and C. Bednarz. 2008. A real-time wireless smart sensor array for scheduling irrigation. *Computers and Electronics in Agriculture* 61:44–50.

Webster, R., and M.A. Oliver. 2001. *Geostatistics for Environmental Scientists*. New York: John Wiley & Sons, Ltd.

Whelan, B.M., and A.B. McBratney. 2000. The "null hypothesis" of precision agriculture management. *Precision Agriculture* 2:265–279.

Wibawa, W.D., D.L. Dludlu, L.J. Swenson, D.G. Hopkins, and W.C. Dahnke. 1993. Variable fertilizer application based on yield goal, soil fertility, and map unit. *Journal of Production Agriculture* 6:255–261.

Wilson, J.P., and J.C. Gallant. 2000. Digital terrain analysis. In *Terrain Analysis: Principles and Applications*, ed. J.P. Wilson, and J.C. Gallant, 1–28. New York: John Wiley & Sons, Ltd.

Wise, K., B. Johnson, C. Mansfield, and C. Krupke. 2011. Managing wheat by growth stage. Purdue Ext. 422. Purdue University, IN. https://www.extension.purdue.edu/extmedia/ID/ID-422.pdf (accessed October 1, 2014).

World Resources Institute. 2013. *Creating a Sustainable Food Future: Interim Findings*. Washington, DC: World Resources Institute.

Zhang, M., Z. Qin, X. Liu, and S.L. Ustin. 2003. Detection of stress in tomatoes induced by late blight disease in California, USA, using hyperspectral remote sensing. *International Journal of Applied Earth Observation and Geoinformation* 4:295–310.

Zhao, D., T. Yang, and S. An. 2012. Effects of crop residue cover resulting from tillage practices on LAI estimation of wheat canopies using remote sensing. *International Journal of Applied Earth Observation and Geoinformation* 14:169–177.

Zhao, F., Y. Huang, Y. Guo, K.N. Reddy et al. 2014. Early detection of crop injury from glyphosate on soybean and cotton using plant leaf hyperspectral data. *Remote Sensing* 6:1538–1563.

Zheng, G., and L.M. Moskal. 2009. Retrieving leaf area index (LAI) using remote sensing: Theories, methods and sensors. *Sensors* 94:2719–2745.

Zikopoulos, P., C. Eaton, D. deRoos, T. Deutsch, and G. Lapis. 2011. *Understanding Big Data: Analytics for Enterprise Class Hadoop and Streaming Data*. New York: McGraw-Hill Companies, Inc.