ZONE MAPPING APPLICATION FOR PRECISION-FARMING: A DECISION SUPPORT TOOL FOR VARIABLE RATE APPLICATION

Xiaodong Zhang, Craig Helgason, and George Seielstad

Northern Great Plains Center for People and the Environment
University of North Dakota
Grand Forks, North Dakota

Lijian Shi

Ocean Remote Sensing Institute
Ocean University of China
Qingdao, China

ABSTRACT

We have developed a web-based decision support tool, Zone Mapping Application for Precision Farming (ZoneMAP, http://zonemap.umac.org), which can automatically determine the optimal number of management zones and delineate them using satellite imagery and field survey data provided by users. Application rates, say for fertilizer, can be prescribed for each zone and downloaded in a variety of formats to ensure compatibility with GPS-enabled farming applicators. ZoneMAP is linked to Digital Northern Great Plains, a web-based application which hosts a rich archive of satellite imagery from Landsat, MODIS, and ASTER, as well as high resolution airborne imagery from AeroCAM and AgCam. ZoneMAP transparently handles projection conversion, grid resampling, and spatial subsetting for data from a variety of sources. We used an unsupervised clustering method, Fuzzy C mean (FCM), for classification. We tested two cluster estimation algorithms and found that the diagonal distance algorithm gives more consistent results than Mahalanobis distance. We also evaluated efficacy of ZoneMAP using real field data provided by end users. Management zones created by ZoneMAP mapped natural variation of the soil organic matter and other nutrients relatively well, and are consistent with zone maps created by the users. The results demonstrated that ZoneMAP can serve as an effective yet easy-to-use tool for those who want to practice precision agriculture.

Keywords: precision agriculture, FCM, management zone, zone map, remote sensing, variable rate application.

INTRODUCTION

A major goal of precision agriculture is to increase profitability, while reducing input costs and protecting the environment. Precision farming requires the ability to vary rates of application and precisely apply needed inputs, such as herbicide and fertilizer. In order to fully utilize the potential of variable rate technology (VRT), the variability of soil, yields and landscape characteristics need to be determined *a priori*.

Soil surveys are the most accurate method to get the physical and chemical attributes of soil and they are often used in the fertilizer recommendations for agronomic crops. But usually a large number of distributed samples are needed to achieve statistical significance, which can be cost prohibitive and time consuming. On the other hand, coarse soil surveys are seldom useful in determining management zones (Franzen et al., 2002).

Spatial variability in yields has been considered as another widely used factor for variable rate nutrient management (Johnson et al., 2003). The yield variation not only reflects within-field variation in soil productivity potential (Brock et al., 2005) but also provides an indication of the nutrient level for the following spring if litter is left on soil to decay. However, the ability to gather yield information at harvest requires advanced combines, which are not available to most producers. In addition, the decision rules and underlying concepts of creating management maps from yield data still require additional research and development (Brock et al., 2005).

The experiences of farmers are also important. They know qualitatively which areas of a field yield well and which areas are low in productivity. Fleming et al. (2000) evaluated the farmer-developed management zone maps and concluded that soil color from aerial photographs, topography, as well as the farmer’s past management experience are effective in developing variable rate application maps. By comparing and evaluating management zones developed from soil color and farmer experience with those developed from measurements of soil electric-conductivity, Fleming et al. (2004) showed that both methods were able to identify homogeneous sub-regions within a field.

Spatial imagery has been used for crop management since 1929, when aerial photography was used to map soil resources (Seelan et al., 2003). Remote sensing for precision agriculture is based on the relationships of surface spectral reflectance with various soil properties and crop characteristics (Moran et al., 1997). Many laboratory-based, machine-attached, hand-held or airborne spectrometers have been used. Multi- or hyper-spectral reflectance of soil in the visible and near-infrared (VNIR) spectral regions have been used to map soil organic matter (SOM) with promising results (Daniel et al., 2004; Wetterlind et al., 2008). Read et al. (2002) used a spectroradiometer to measure the leaf and canopy reflectance within 350 to 950 nm and found that the reflectance of red edge region (690-730 nm) is sensitive to Nitrogen (N) stress. Fleming et al. (2000) found that soil color from aerial photography is useful in developing variable rate application maps.

Satellite observations provide measurements of surface reflectance with 15 - 60 m resolution (e.g. SPOT, Landsat or ASTER) on a temporal scale of multiple visits during a growing season. Sullivan et al. (2005) used soil sampling data and IKONOS imagery to estimate the soil properties' variability and concluded that the high resolution multi-spectral data is a good soil-mapping tool. Seelan et al. (2003) compared a 9-m-wide N test strip where purposely no fertilizer was applied with adjacent strips on IKONOS imagery and could detect N deficiency. Bhatte et al. (1991) found that the soil organic matter (SOM) distribution estimated from Landsat images was strongly correlated with that determined from soil sampling. Salisbury and D'Aria (1992) reported that thermal infrared band ratios from the ASTER sensor (range 8-14 µm, resolution 90 m) could be used to differentiate soil properties such as particle size, soil moisture, and soil organic content.

Given steeply rising prices in chemicals and fuel, and increasing awareness of the need to preserve our natural environment, producers rely more and more on precision farming to reduce costs, both economical and environmental. Despite the potential advantage and many industrial efforts to develop various hardware and software tools, precision farming has yet to be adopted widely.

Fridgen et al. (2004) developed a Management Zone Analysis (MZA) tool for subfield management zone delineation. This Windows-based software is easy to use and effective in delineating management zones (2005). However, it places a stringent and onerous requirement on data preparation – all the input layers, vector or raster, have to be gridded into common grid cells. In addition, MZA and many other application tools fail to address a major issue that has prevented the wide adoption of precision agriculture: access to data.

We have developed and recently released a web-based decision support tool, Zone Mapping Application for Precision Farming (ZoneMAP), which not only can be used for classifying fields into zones but also has seamless access to a rich archive of remote sensing data spanning the past 30 years. By streamlining format conversion, reprojection, and gridding of data from various sources, ZoneMAP (http://zonemap.umac.org), provides users with a tool as well as data that are available at their finger tips.

Here we report development of the algorithms used in ZoneMAP for classification and automatic determination of optimal number of zones, describe the image database, and provide examples of ZoneMAP outputs for two farm fields and their evaluation.

CLASSIFICATION ALGORITHM FOR ZONEMAP

We chose fuzzy c-means (FCM) as the clustering algorithm for ZoneMAP. It is basically the same as that used by Fridgen et al. (2004) but slight difference in estimating the measure of similarity between an observation and cluster centers. Typically, measure of similarity can be estimated using Euclidean distance, diagonal distance, or Mahalanobis distance. Since the Euclidean distance algorithm requires variables to be of equal variances, which are rarely true in reality, we only implemented the latter two algorithms. For diagonal distance, instead of adjusting the estimate by the variance of the related variables, we reshape the input variables such that they all have a mean of zero and a unit variance. This process does not affect the classification results, but it is faster and therefore highly suitable for a web-based application, such as ZoneMAP.

Mahalanobis distance accounts for situations where input variables are statistically dependent with unequal variance. Since it relies on a variance-covariance matrix for weighting, which has to be calculated for all the input variables, it is more computationally intensive than estimating diagonal distance. Also we noted that with Mahalanobis distance the final classification may vary depending on the initial values assigned to the cluster centers, a problem also reported by Fridgen et al. (2004). However this variation was not encountered when diagonal distance was used. We do not know if this has anything to do with reshaping of input data.

ALGORITHM FOR OPTIMAL NUMBER OF ZONES

Determining the most appropriate number of zones is difficult in the interpretation of unsupervised classification. Fridgen et al. (2004) used the convergence of fuzziness performance index (FPI) and normalized classification entropy (NCE) to determine the optimal number of management zones. Theoretically, the best classification occurs when membership sharing (FPI) and the amount of class disorganization (NCE) is at a minimum with the least number of classes used. However, sometimes NCE and FPI do not converge and the optimal number of zones suggested by one parameter is significantly different from the one suggested by another (Brock et al., 2005).

Another method to evaluate classification success is to estimate how much within-cluster variability is reduced for a number (n) of clusters as compared with n-1 clusters. We have found that generally the percentage of total within-cluster variability with respect to the total initial variability decreases as the number of

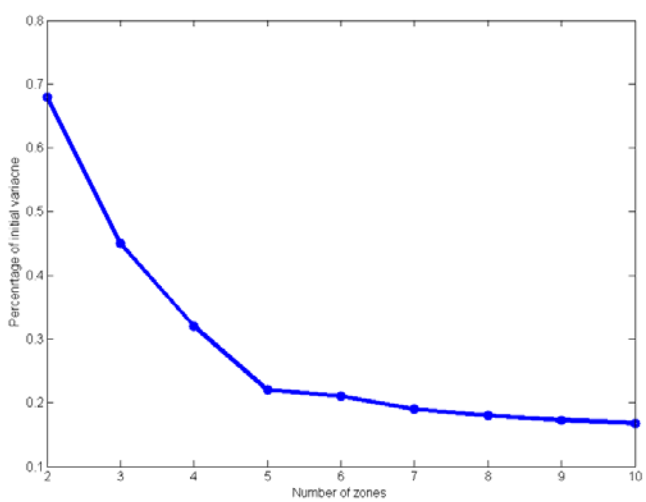

Fig. 1. The total within-cluster variability as a percentage of initial variance normally decreases with the number of zones.

clusters increases as shown in Figure 1. A similar trend for the variance reduction was found by Broker et al. (2005). We also found that typically the total within-cluster variance decreases rapidly initially and then approaches an asymptotic value slowly as the number of clusters continues to increase. The optimal number of zones is therefore decided as the number of clusters that reduces the variance significantly as compared to the initial variability, yet changes little when the number of zones is further increased. By trial and error, we came to two criteria that can capture this turning point in a relatively consistent manner: 1) overall reduction of variance is > 50%; and 2) consecutive reduction of variance is < 20% or the trend is broken, i.e., within-cluster variability increases instead of decreasing. For the case shown in Fig. 1, the optimal number of zones should be 5.

DESIGN OF ZONEMAP

ZoneMAP was designed for end users like farmers, ranchers, or extension specialists to practice precision agriculture; therefore the ease of use is important. Also important is the access to data, especially remote sensing observations which have been shown to be extremely effective in capturing field variability (e.g., Seelan et al., 2003; Sullivan et al., 2005).

Remote Sensing Imagery

We have collected a rich archive spanning more than 30 years of remote sensing imagery over the northern Great Plains including North and South Dakota, Minnesota, Montana, Wyoming, and Idaho. Data include medium resolution (20 – 250 m) multispectral images from satellite sensors of Landsat MSS, TM and ETM+, ASTER, and MODIS, surface relief from SRTM, and high resolution (1 – 2 m) images from AeroCam, a multispectral airborne camera that we developed and operate. These data are publicly available through Digital Northern Great Plains website (http://dngp.umac.org). ZoneMAP is internally linked to the database and has seamless access to this valuable digital resource.

To ensure consistency in temporal and spatial comparisons, all the satellite images have been atmospherically corrected. The final product is reflectance on the ground. Although the AeroCam sensor has been carefully calibrated at the NASA Ames Research Center to determine its spectral and radiometric characteristics, we did not perform the atmospheric correction for AeroCam images because at typical altitude of 2 – 3 km, the contribution to signals by the atmosphere is small. The final product for AeroCam is radiance at the aircraft.

Reflectance of canopy will change during a growing season as vegetation goes through stages of first growth, maturity, reproduction, and senescence. Vegetation indexes, such as Normalized Differential Vegetation Index (NDVI) estimated using reflectance measurements at the red and near-infrared (NIR) wavelengths, or Green NDVI estimated by replacing the red with the green, have been developed to track the vigor of plants and have been used widely for developing management zones (Metternicht, 2003; Moran et al., 1997). ZoneMAP will estimate NDVI and GNDVI on-the-fly if a user chooses to the vegetation index for classification.

Image Processing

For management zones to be representative various factors affecting the soil characteristics and potential productivity need to be considered. This often entails the use of data from different sources, of different ground sampling distances, and with different formats and projections. Before being combined for further cluster analysis, different data need to be projected onto a common grid, which often involves subsetting, reprojecting, and resampling procedures.

Typically, a remote sensing image covers a much bigger area than a farm field. Instead of processing the entire image, ZoneMAP automatically crops the image using an area of interest (AOI) defined by the user, which considerably enhances the overall performance. ZoneMAP also automatically reprojects and resamples different images to a common projection plane with an equal ground sampling distance determined by the user. We used the open source libraries GDAL and OGR to implement these procedures.

By automating these tedious yet critical image processing steps transparently to users, we expect the learning curve for using ZoneMAP will be greatly reduced.

Users Data

All users' data are saved in a secure online database so that across-season or multi-year comparisons of management zones can be performed to evaluate their consistency. Field measurements, such as yield or electrical-conductivity (EC) of soil, can be uploaded into the ZoneMAP database and used along with remote sensing imagery for delineating management zones. For each creation of a set of management zones, metadata is generated describing the procedure and datasets used such that the classification can be reproduced later.

Users of ZoneMAP can download their results in three formats, raster image, grid text, and shape file. For each format, there are multiple projections to choose from. In addition, users can input application rates for each zone to generate a variable rate application map.

TESTING

We tested the performance of ZoneMAP using data from two private farm fields. The two fields are for production and are not specifically designated for research. Despite uncertainties that may be associated with this data collection policy, we feel that it is important for us to evaluate the performance of ZoneMAP using real data by real users.

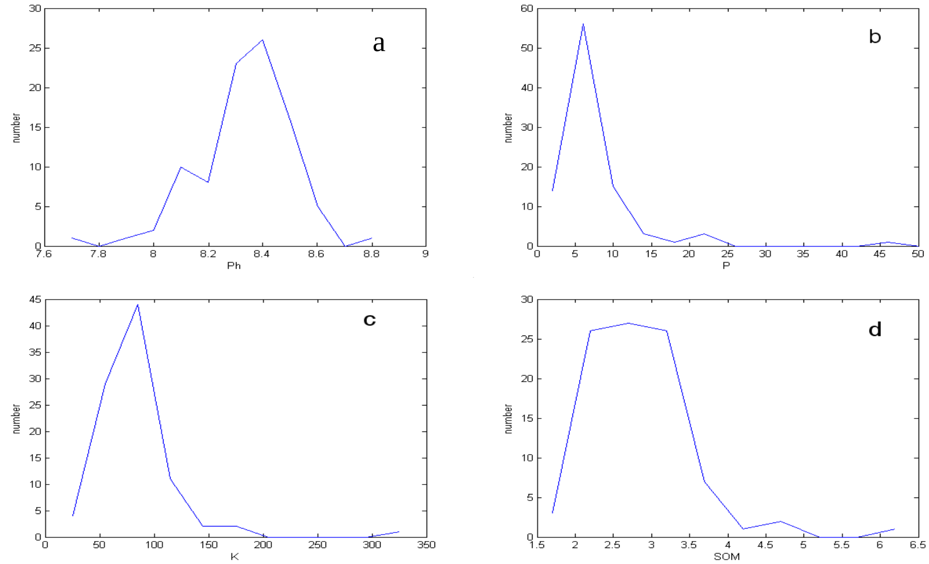

Fig. 2. The histograms of measurements of pH (a), P (b), K (c) and SOM (d) for the first testing field.

Field 1

The first field of 238.7 acres is located in Polk County, Minnesota, with soybeans planted in 2004 and wheat in 2005. The soil sampling conducted in the fall of 2005 after the harvest analyzed the contents of phosphorus (P), potassium (K), soil organic matter (SOM) and pH. Fig. 2 shows the histograms of the measurements. It is interesting to note that the distributions of P, K and SOM all exhibit a heavy tail towards higher values, while pH shows a moderate tail towards smaller values.

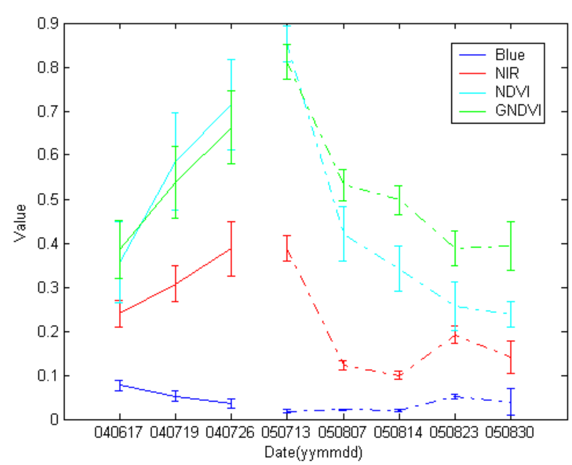

Fig. 3. Temporal variation of the mean (and the standard deviation) of surface reflectance by Landsat at the blue and the NIR bands and vegetation indexes of NDVI and GNDVI for the growing season of 2004 (solid line) and 2005 (dashed line).

Eight Landsat images between June 2004 and August 2005 covered the first field. The temporal variations of surface reflectance measured by Landsat at the wavelengths of the blue and the NIR along with the corresponding NDVI and GNDVI are shown in Fig. 3. The cloud cover limited the satellite coverage to the first half of the 2004 growing season and to the second half of the 2005 growing season.

Due to strong absorption by chlorophyll pigments at the blue wavelengths, the reflectance of band 1 of Landsat typically decreases as chlorophyll concentration increases with maturation and then increases as chlorophyll concentration decreases towards senescence. The trend is opposite to the reflectance at the NIR (and the green and red, not shown), which is positively linked with leaf cellular structure. The maximum reflectance in the NIR and the maximum vegetation indexes (NDVI and GNDVI) occurred on July 26 for 2004 (soybeans) and July 13 for 2005 (wheat). These maximums likely occurred when crops reached full canopy.

The optimal number of zones determined by ZoneMAP was 3 when two NIR images were used and the resulting zone map is shown in Fig. 4-a. Each zone is clearly defined into distinctive domains defined by NIR reflectance of 20040726 vs. 20050713 (Fig. 4-b). The histogram distribution of SOM within each zone (Fig. 4-c) showed that these subfield zones have separated SOM into three classes, with the mean for each class being 3.15, 2.95, and 2.42. The means of pH values for each zone are 8.42, 8.19, and 8.36, respectively. The ANOVA test showed that classification of SOM and pH based on the management zones created using two NIR images at two growing seasons were significant, with Pr < 0.0001.

The distributions of K and P are less distinctive than SOM and pH values;

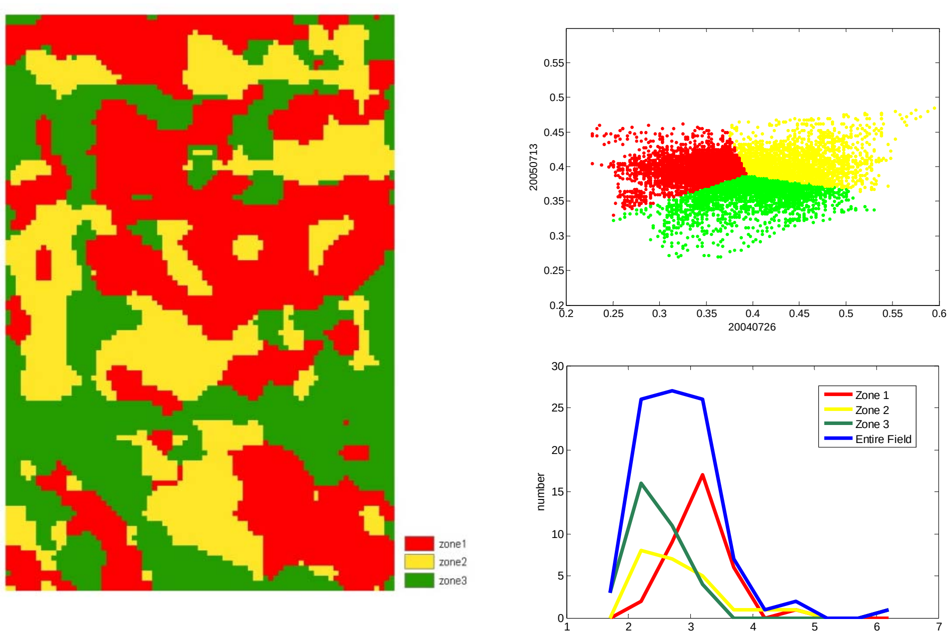

Fig. 4. Management zones (a, left) and the corresponding scattering plot (b, UR) created using a pair of reflectance measurement at NIR by Landsat on July 26 2004 and July 13 2005. The histograms of SOM within each zone and for the entire field is plotted in c (LR).

however they do show a clear trend of variation among different zones. Higher values of K and lower values of P were found in zones 1 and 2, which have high and medium concentration of SOM, while lower K and higher P values mostly occur at zone 3, whose SOM concentration is the lowest. Similar variations of nutrients with SOM were reported by Fleming et al (2004; 2000), who suggested that lower productivity areas would remove less P resulting in a buildup of the soil test P-level.

The management zones classified using NDVI or GNDVI did not separate SOM and pH into classes as distinctive as the zones based on the NIRs. Within zones created with NDVI or GNDVI, the spread of SOM in zone 2 covers that of both zones 1 and 3. Similarly less effective classifications were found for pH, K, and P as well. This result is not surprising. We found relatively high correlation (~ 0.4) between SOM and the reflectance at the NIR as compared to ~ 0.3 between SOM and the two vegetation indexes. Since reflectance at the NIR is primarily affected by the leafy structure of crops, our results suggested that cellular structure of canopy correlates well with the soil organic matter concentration.

The results mentioned above were based on the two Landsat images acquired at the full canopy stage of the growing seasons. We also tested classifications using images at other growing stages. The subfield zones were not as effective in terms of partitioning SOM. However, they may be more effective for other applications, say fertilizer application.

Field 2

The second field is located in Potter County of South Dakota with an area of 112.0 acres. The rotation of crops from 2003 to 2005 was corn, sunflowers and spring wheat. Using a yield map collected in 2003 (Fig. 5-a) and an NDVI map by Landsat on August 25, 2004 (not shown), the farmer created four subfield zones (Fig. 5-b), to determine the application rates of urea for the year of 2005. As a result of this variable-rate application, the spring wheat planted in 2005 delivered a much more uniform yield (Fig. 5-c). While the mean yields of each crop are about the same, 116.78 bu/ac for corn and 114.29 bu/ac for spring wheat, the standard deviation (SD) was reduced from 30.76 bu/ac for corn of 2003 to 19.64 bu/ac for spring wheat of 2005.

Some farmers do not have yield monitoring capabilities, so we tested whether replacing the yield map with satellite imagery, preferably close to the time of harvest, can generate an equally good zone map. The NDVI derived from Landsat on September 1, 2003 (Fig. 5-d) showed some correlation with the corn yield of 2003 (Fig. 5-a). And the zone map created using the NDVIs from September 1 2003 and August 25 2004 is shown in Fig. 5-e, whose zones 1, 3 and 6 roughly correspond to the zones with low, moderate and extra high rate in Fig. 5-b, respectively.

Even though the management zones shown in Figs. 5-b and 5-e are similar, there are marked differences, especially for the zone designated as “extra high rate”. Actually, we have tested with combinations of different bands, indices, or dates, and none can reproduce the zone map created using yield data. This corroborates one argument that has been frequently stressed by our end users who

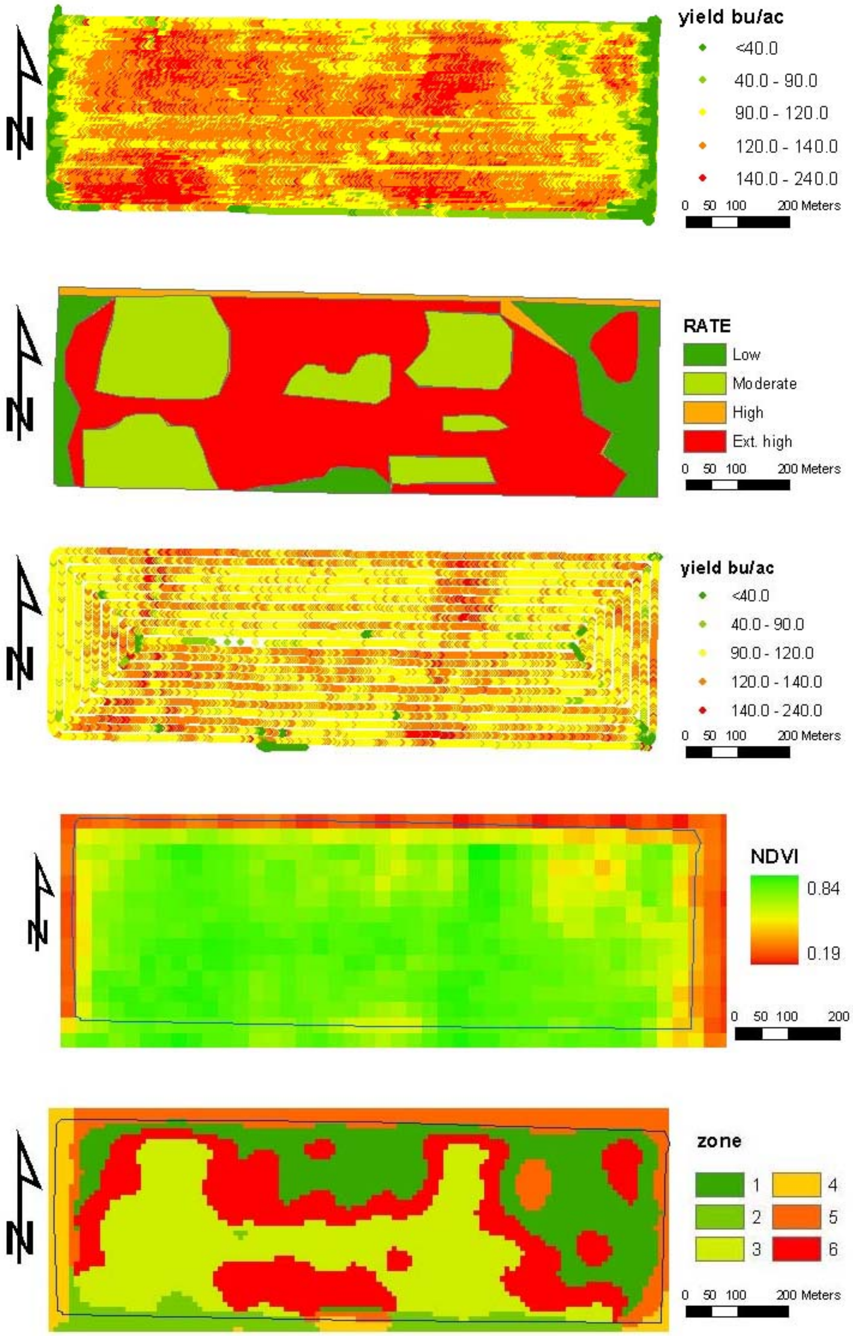

Fig. 5. Using 2003 yield map of corn (a, top) and 2004 NDVI map by Landsat of August 25, 2004, the farmer created the management zones (b, 2nd row) as a basis for determination of variable rate fertilizer application resulting in a more uniform yield for 2005 spring wheat (c, 3rd row). The zone map (e, bottom) when the yield data of 2003 is replaced with NDVI data of September 1, 2003 (d, 4th row)

are early adopters of precision agriculture, “it is critical to monitor the yield and use it in zone management”.

CONCLUSIONS

With the rising costs of raw materials and chemicals, a rapidly degrading natural environments, and increasing global population, we believe precision agriculture is a critical step towards sustainability. ZoneMAP is a web-based decision support tool designed to promote precision agriculture. With its ease of use, extensibility, and access to a rich archive of existing remote sensing data, we hope this tool entice more users to adopt variable rate application through subfield zone management. By testing with real field data from our end users, our results confirm that remote sensing data can be effectively incorporated into delineation of management zones despite many limitations. We also recognize that field surveys of soil attributes and nutrient conditions are important and sometimes cannot be replaced by today's remote observations. However, a preliminary mapping of subfield zones using remote sensing data may help to design a cost-effective plan.

ACKNOWLEDGMENTS

The funding for the project was provided by NASA/USRA through grants NCC8-259 and 03450-08. We thank Mr. Dan Forgey, Mr. Gary Wagner and many other end users for sharing their data as well as knowledge. Without their help this research would not be possible. Funding for Lijian Shi was provided through NASA grant NNX06AE16G.

REFERENCES

Bhatti, A.U., D.J. Mulla, and B.E. Frazier. 1991. Estimation of soil properties and wheat yields on complex eroded hills using geostatistics and thematic mapper images. Remote Sensing of Environment 37:181-191.

Brock, A., S.M. Brouder, G. Blumhoff, and B.S. Hofmann. 2005. Defining Yield-Based Management Zones for Corn-Soybean Rotations. Agron J 97:1115-1128.

Daniel, K.W., N.K. Tripathi, K. Honda, and E. Apisit. 2004. Analysis of VNIR (400–1100 nm) spectral signatures for estimation of soil organic matter in tropical soils of Thailand. International Journal of Remote Sensing 25:643 - 652.

Fleming, K.L., D.F. Heermann, and D.G. Westfall. 2004. Evaluating Soil Color with Farmer Input and Apparent Soil Electrical Conductivity for Management Zone Delineation. Agron J 96:1581-1587.

Fleming, K.L., D.G. Westfall, D.W. Wiens, and M.C. Brodahl. 2000. Evaluating Farmer Defined Management Zone Maps for Variable Rate Fertilizer Application. Precision Agriculture 2:201-215.

Franzen, D.W., D.H. Hopkins, M.D. Sweeney, M.K. Ulmer, and A.D. Halvorson. 2002. Evaluation of Soil Survey Scale for Zone Development of Site-Specific Nitrogen Management. Agron J 94:381-389.

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.

Johnson, C.K., D.A. Mortensen, B.J. Wienhold, J.F. Shanahan, and J.W. Doran. 2003. Site-Specific Management Zones Based on Soil Electrical Conductivity in a Semiarid Cropping System. Agron J 95:303-315.

Metternicht, G. 2003. Vegetation indices derived from high-resolution airborne videography for precision crop management. International Journal of Remote Sensing 24:2855-2877.

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.

Read, J.J., L. Tarpley, J.M. McKinion, and K.R. Reddy. 2002. Narrow-Waveband Reflectance Ratios for Remote Estimation of Nitrogen Status in Cotton. J Environ Qual 31:1442-1452.

Salisbury, J.W., and D.M. D'Aria. 1992. Infrared (8-14 [mu]m) remote sensing of soil particle size. Remote Sensing of Environment 42:157-165.

Seelan, S.K., S. Laguette, G.M. Casady, and G.A. Seielstad. 2003. Remote sensing applications for precision agriculture: A learning community approach. Remote Sensing of Environment 88:157-169.

Sullivan, D.G., J.N. Shaw, and D. Rickman. 2005. IKONOS Imagery to Estimate Surface Soil Property Variability in Two Alabama Physiographies. Soil Sci Soc Am J 69:1789-1798.

Wetterlind, J., B. Stenberg, and M. Söderström. 2008. The use of near infrared (NIR) spectroscopy to improve soil mapping at the farm scale. Precision Agriculture 9:57-69.