Integrating disparate data sources in an agricultural GIS

Michael G. Street and Dr. Dion J. Wiseman
Brandon University

Introduction

The modern agricultural industry is in the midst of a technological revolution. GPS (Global Positioning System) receivers and integrated yield, moisture, and protein monitors will soon be standard options on new machinery. In addition, variable rate seeding and spraying systems are expected to be commonplace within a few years. In anticipation of the demand for geomatics technology and spatial information, enterprising rural entrepreneurs have initiated the growth of a rural-based agro-geomatics industry. These relatively small operations offer a variety of imagery (e.g. CIR aerial photography, B/W digital orthophotography, Landsat TM, and SPOT) distributed by section or quarter section, GPS receivers, and mapping, image analysis, and soil testing services.

The agricultural industry's adoption of technology has traditionally been quite typical in that it is driven by profitability. The relatively rapid incorporation of fertilizers and pesticides, for example, occurred as a result of the significant yield increases and profits that were realized. However, when such an economic advantage can not be demonstrated, due either to real or perceived technological limitations, the adoption rate may be comparatively low.

Technology has to be affordable, profitable, and easy to use. These critical barriers to the adoption of geomatics technology in the agricultural industry have yet to be conquered. GPS equipped yield moisture and protein monitors produce vast quantities of spatial information that are in themselves of little use until properly interpreted and incorporated into a farm management strategy. The integration and effective utilization of these new technologies presents a challenge that the ultimate benefactor, the farmer, is currently ill equipped to accommodate.

Objectives

This paper is the result of ongoing research conducted on behalf of the Manitoba Zero Tillage Research Association (MZTRA) addressing issues related to the development and implementation of an agricultural geographic information system in a rural Manitoba setting. The objective of this research is to develop and evaluate methodologies for integrating and effectively utilizing the variety of data currently available with consideration of those factors believed to impede the adoption of these technologies.

Vegetation indices derived from NOAA AVHRR, Landsat MSS, Landsat TM, and SPOT HRV red and near infrared bands have been used extensively for monitoring the presence and vigor of natural vegetation and cropland (Tucker, 1979; Tucker et al., 1984; Goward et al., 1985; Tucker et al., 1985; Philipson and Teng, 1988; Lloyd, 1990; Teng, 1990; Eidenshink, 1992; Richardson and Everitt, 1992; and Zhu and Evans, 1994). More recently, indices derived from high-resolution video systems have been similarly used (Mausel, 1993).

In most cases, vegetation indices are based on relatively course resolution imagery. For example, Doraiswamy and Cook (1995) attempted to assess spring wheat yield utilizing 1.1-Km AVHRR imagery. Of primary concern were issues associated with mixed pixels and interference with non-wheat areas; both a function of image resolution. However, with the exception of Tucker (1979), few CIR derived vegetation indices have been used for crop assessment or yield prediction. These comparatively high-resolution data may provide information regarding crop health and ultimate yield at a sufficient level of detail and early enough in the growing season to be incorporated into the current years farm management strategy. This paper evaluates and compares the relationship between wheat yield and the Normalized Difference

Vegetation Index (NDVI) derived from Landsat TM and CIR photography.

Study Area and Data Sources

The MZTRA Research Farm is a one-section facility located 18 km north of Brandon, Manitoba. Private corporations, Agriculture Canada, Manitoba Agriculture, Ducks Unlimited, and resident staff conduct research on behalf of association farmers. The farm is equipped with the latest technology, such as differential GPS receivers and yield and moisture monitors. Further, remote sensing data is routinely acquired to support ground experimentation.

Yield data was collected during the 1997 harvest at a rate of one sample per second. A John Deere 6620 combine was equipped with an AgLeader-2000 yield monitor, which had been calibrated to +/- 5%. An Omnistar 4000 Differential GPS receiver, accurate to 1m, provided georeferencing. Geo-referenced points and corresponding yield values were stored on a PCMCIA memory card and later imported as an ASCII text file.

Prairie Agri-Photo Ltd. is contracted by the MZTRA to acquire CIR aerial photography of the research farm each year. In 1997 photography was acquired on July 27th at a height of 7600 ft, using a 70mm aerial camera with an IR filter. This photo was scanned at 600 dpi resulting in a ground resolution of approximately 1-meter.

Landsat TM imagery acquired August 8th 1997 was provided by Agriculture Canada. The image was geo-referenced and NDVI values calculated prior to receipt of the data. The image was then subset to an area slightly larger than that of the research farm.

Methodology

Image Rectification and Preprocessing

In order to integrate these disparate data each had to be geo-referenced to a common projection. In addition, the scanned CIR photo had to be rectified to correct for distortion inherent to aerial photography. Ground control points were collected by mounting a

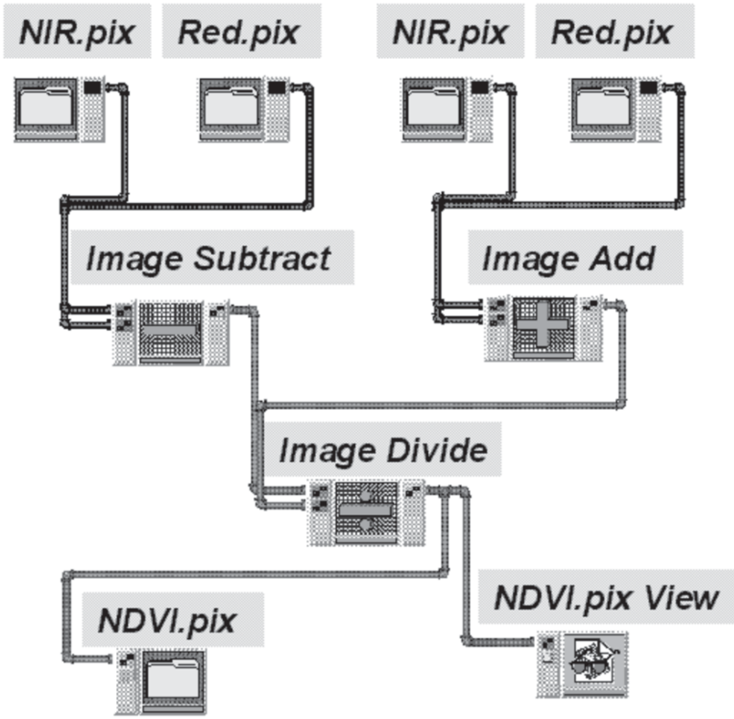

Figure 1: *Derivation of NDVI in PCI Visual Modeler.*

differentially corrected GPS receiver on an ATV and visiting a number of sites across the study area that were visible on the scanned photograph. These control points were then used to rectify the scanned photo using the PCI Image Works GCP-works function.

The Landsat TM and yield data were also imported into PCI ImageWorks. The spatial accuracy of the Landsat image was checked against the control points and as a result the entire image was shifted 7 meters to the east. The yield data was inspected for anomalous points (i.e. points not differentially corrected due to signal noise) which were edited from the file.

Calculation of the NDVI from CIR Photography

The NDVI is a simple ratio of reflected near infrared and red light that compensates for variations in illumination, surface slope,

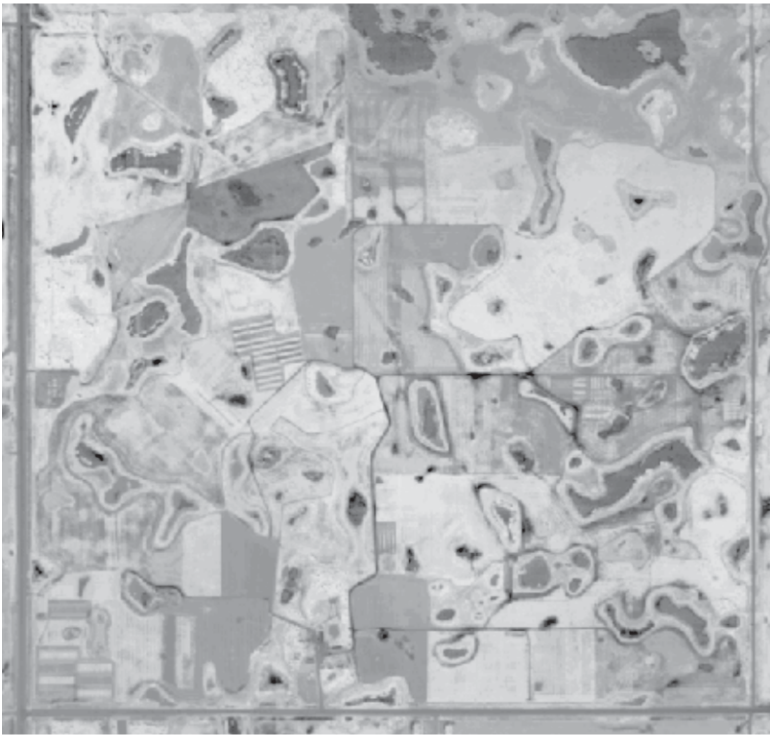

Figure 2: *NDVI image derived from scanned CIR photo.*

and viewing aspect. Values range from –1 to 1, with vegetation ranging from 0.1 to 0.6. Higher values are associated with lush, healthy vegetation; non-vegetated surfaces such as soil, rock, or water typically have negative NDVI values (Avery and Berlin, 1992).

$$\textbf{NDVI = NIR-Red/NIR+Red}$$

The scanned CIR photo was split into 3 bands; near infrared, red, and green. The infrared and red bands were saved as separate georeferenced files and the green band was discarded. The PCI Image Works Visual Modeler was used to create a short program (Figure 1) that executed the NDVI equation, creating a new, georeferenced NDVI raster image (Figure 2).

Comparison of Landsat TM and CIR NDVI Values

The Landsat and CIR derived NDVI images were then displayed in linked viewers so that visual comparisons of NDVI values could be made. In addition, two cross sections were created to evaluate difference and similarities between the Landsat TM and CIR NDVI values over different cover types and over the same cover type. The first cross section was approximately 1500 meters in length and covered a variety of features such as water, roads and several crops. The second was a diagonal profile approximately 250 meters in length across a single wheat field.

Visual analysis of these profiles suggested that the Landsat TM NDVI values were significantly influenced by feature boundaries resulting in mixed pixels along the edge of features that were not truly representative of either feature type. Since it was believed a comparison of Landsat TM and CIR derived NDVI values, which included these mixed pixels, would introduce significant error into the resulting correlation, a random sample of 50 “pure” Landsat TM and corresponding CIR NDVI values was selected. The Pearson product moment correlation was performed on these data to evaluate the strength of the correlation between Landsat TM and CIR NDVI values.

Comparison of CIR NDVI and Wheat Yield

An attempt was made to transform the yield data to raster format by interpolating a grid surface based on the original point coverage. However, the resulting raster image possessed a distinct linear orientation due to the irregular distance between successive sampling points as compared to adjacent combine swaths. In addition, in a few instances excessively high yield values occurred where the swath had rolled in front of the combine. Several spatial filters were applied in an attempt to smooth the data but ultimately this approach was abandoned.

Alternatively, a random sample of 100 yield points was selected from the wheat field and the corresponding CIR NDVI values were recorded. A simple linear regression analysis was performed to determine the correlation between these variables and how well

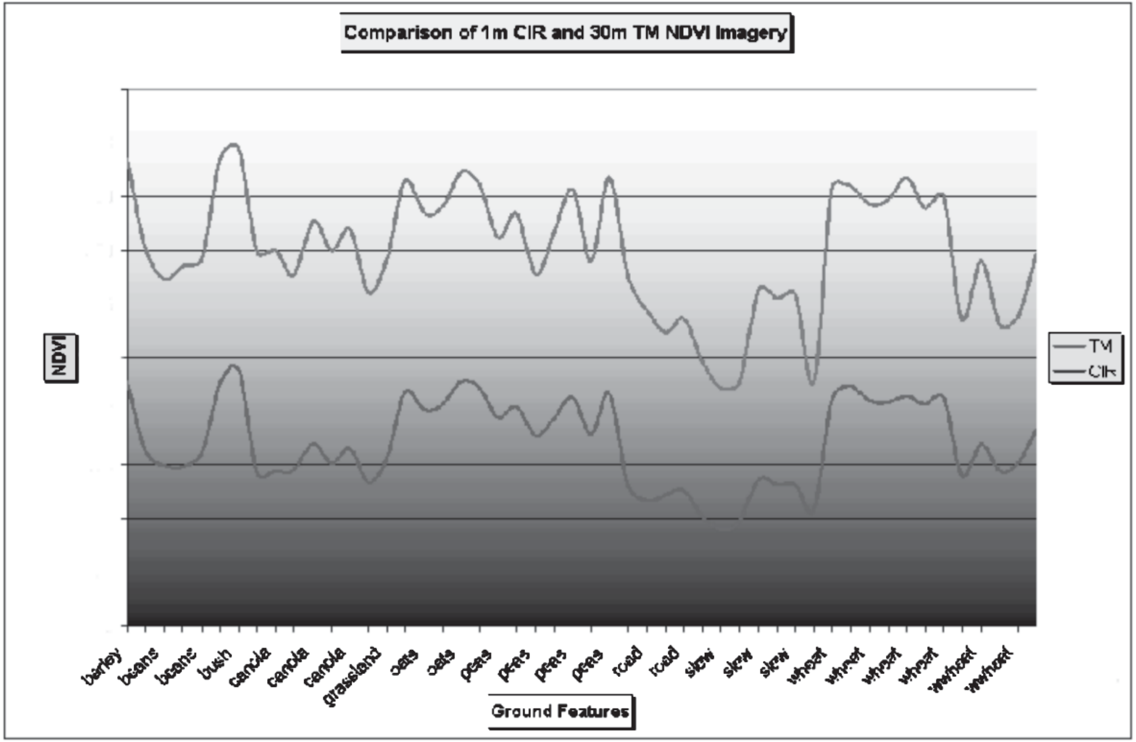

Figure 3: *Comparison of Landsat TM and CIR NDVI values by ground feature.*

mid-season CIR NDVI values are able to predict ultimate wheat yield.

Results

Figure 3 illustrates the strong relationship between Landsat TM and CIR NDVI values over a variety of landcover types. The Pearson product moment correlation coefficient over all landcover types was 0.68; however, when canola was removed from the sample the correlation increased to 0.95. It is believed that this variation was due a change in the spectral characteristics of canola between the acquisition date of the CIR photograph (July 27) and Landsat TM imagery (August 8). The strong relationship between these variables when mixed TM pixels are eliminated from the sample suggests that Landsat TM and CIR derived NDVI data are comparable even though the resolution of these data are significantly different

Figure 4 depicts the results of regression analysis. The correlation between mid-season NDVI and wheat yield was 0.79 and the coefficient of determination 063. This suggests that in our study area 63% of the variability in wheat yield can be predicted by mid season CIR derived NDVI values.

Conclusions

Given the significant correlation between Landsat TM and CIR derived NDVI values it seems reasonable to suggest that comparatively high resolution CIR photography may provide an affordable alternative to Landsat TM imagery for use in an agricultural GIS. The resolution of CIR photography reduces errors associated with mixed pixels and provides the level of detail required for analysis of this scale.

Regression analysis suggests that in this case there seems to be a significant relationship between mid-season CIR NDVI values and wheat yield, and while 63% of the variation in yield can be predicted by the NDVI, a large proportion of the variability is unaccounted for. Nonetheless, it would appear that it may be

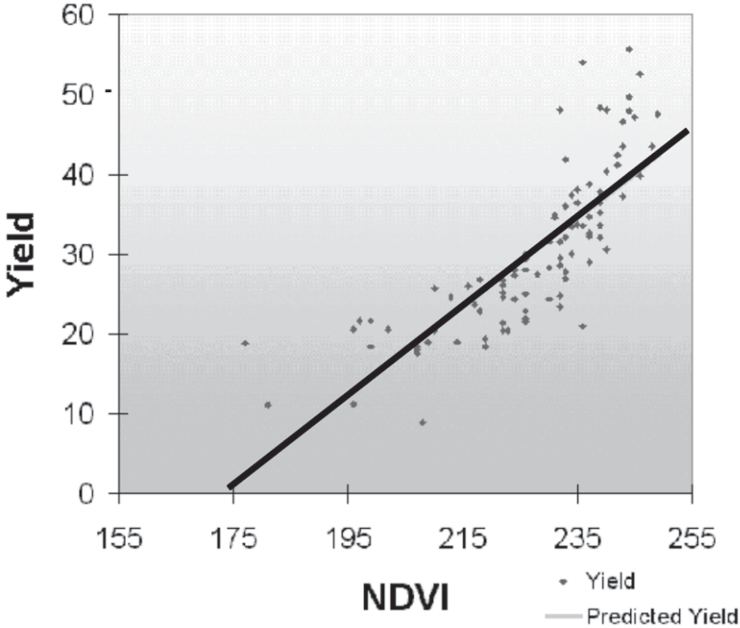

Figure 4: *Results of regression analysis.*

possible to predict wheat yield utilizing readily available and relative inexpensive data early enough in the season to allow this information to be integrated into that years farm management strategy.

References

AVERY, T.E. and BERLIN, G.L. 1992 *Fundamentals of Remote Sensing and Air Photo Interpretation*. New York: Macmillan.

DORAISWAMY, P. and COOK, P.W. 1995 'The spring wheat yield assessment using NOAA AVHRR data' *Canadian Journal of Remote Sensing*, 21: 43-51.

EIDENSHINK, J.C. 1992 'The 1990 conterminous U.S. AVHRR data set' *Photogrammetric Engineering and Remote Sensing*, 58: 809-813.

GOWARD, S.N., TUCKER, C.J., and DYE, D.G. 1985 'North American vegetation patterns observed with NOAA AVHRR' *Vegetation*, 64: 3-14.

LLOYD, D. 1990 'A phenological classification of terrestrial vegetation using shortwave vegetation index imagery' *International Journal of Remote Sensing*, 11: 2269-2279.

MAUSEL, P. W., EVERITT, J. H., ESCOBAR, D.E., and KING, D. J. 1993 'Airborne videography: Current status and future perspectives' *Photogrammetric Engineering and Remote Sensing*, 58: 1189-1195.

PHILIPSON, W.R. and TENG, W.L. 1988 'Operational interpretation of AVHRR vegetation indices for world crop information' *Photogrammetric Engineering and Remote Sensing*, 54: 55-59.

RICHARDSON, A.J. and EVERITT, J.H. 1992 'Using spectral vegetation indices to estimate rangeland productivity' *Geocarto International*, 1: 63-69.

TENG, W.L. 1990 'AVHRR monitoring of U.S. crops during the 1988 drought' *Photogrammetric Engineering and Remote Sensing*, 56: 1143-1146.

TUCKER, C.J. 1979 'Red and photographic infrared linear combinations for monitoring vegetation' *Remote Sensing of Environment*, 8:127-150.

TUCKER, C.J., GATLIN, J.A., SCHNEIDER, S.R. 1984 "Monitoring vegetation in the Nile Delta with NOAA-6 and NOAA-7 AVHRR data' *Photogrammetric Engineering and Remote Sensing*, 50: 53-61.

TUCKER, C.J., TOWNSHEND, J.R.G., and GOFF, T.E. 1985 'African land-cover classification using satellite data' *Science*, 227: 369-375.

ZHU, Z. and EVANS, D.L. 1994 'U.S. forest types and predicted percent forest cover from AVHRR data' *Photogrammetric Engineering and Remote Sensing*, 60: 525-531.