Geophysical Tools and Digital Elevation Models: Tools for Understanding Crop Yield and Soil Variability

Paul Rampant[1] and M Abuzar[2]

[1]Primary Industries Research Victoria, Box 3100, Bendigo Delivery Centre, Bendigo, Victoria 3554, Australia. Email: paul.rampant@dpi.vic.gov.au
[2]Primary Industries Research Victoria, Private Bag 1, Ferguson Road, Tatura, Victoria 3616, Australia.

Abstract

Since the adoption of precision agriculture land managers require tools to map soils at a finer scale to understand variability in yield maps. The increasing input costs to horticultural crops are also instigating the need for finer detailed soil information to enable land managers to profit from variable rate technology. This paper discusses electromagnetic induction (EMI), gamma ray spectrometry (GRS) and digital elevation model (DEM) technologies with regard to mapping of soil variability to understand both within season yield variability and temporal yield variability. While the use of some of these technologies for delineating soil types is not uncommon, the strength of the combination of these techniques to understand yield variability is the main focus of this study.

The decision tree classifier, See5, was used to predict soil types and yield data from the various combinations of geophysical and terrain attributes. Individually the geophysical data were relatively poor predictors of the soil and yield information, misclassifying 18–53% of the trial area. Combined, the geophysical data misclassified only 7–16% of the area.

The terrain attributes were found to predict soil types very well, misclassifying only 2.3% of the area. However the terrain attributes alone misclassified area 6% when predicting yield zones and over 11% when predicting the yield classes for an individual year.

Using all the geophysical and terrain data, the soil types could be predicted very well with less than 2% of the area misclassified and could also predict yield zones quite well misclassifying 5% of the area. The predictions of yield for an individual year were always worse than for soil types and yield zones.

Introduction

Yield maps have been used for some time and the equipment to monitor crop yields is becoming more common. While many land managers find this data an accurate representation of yield variability across the paddock, they find it of limited value as a planning or management tool. Seasonal variability and the lack of understanding of the cause of the spatial variability are the two main limitations of this data. Crop yield is a function of pedologic (soil characteristics), climatic and anthropic (human) factors. Although factors such as rainfall, frost and hail can be monitored we have little control over them. We have some control over some of the pests, and we have control over the anthropic factors of crop management. The factors that vary relatively little over the course of time are pedological.

With the development of precision agriculture and increasing input costs in horticulture, there is a stronger need to understand soil variation at a finer scale than previously demanded. Some of the technologies that have been used to assist with the identification of soil properties and delineation of soil variability are: electromagnetic induction (EMI) which measures the soil apparent electrical conductivity, digital elevation models (DEM) which model the topographic features, and gamma ray spectrometry (GRS) which measures the natural gamma radiation emissions emanating from the near surface soil and rock.

Ground-based EMI surveys have been used for many years to infer areas of saline soils (Cameron et al. 1981), to delineate soil spatial variability (Johnson et al. 2003) and as a surrogate for some soil characteristics such as soil moisture and clay percentage at paddock scale (Sudduth et al. 2001) and depth to claypans (Doolittle et al. 1994). The most common instruments used for soil apparent electrical conductivity measurement in Australia are the EM38 and the EM31 (from Geonics Ltd). Significant soil profile sampling needs to occur before an EMI survey can be interpreted for most soil characteristics.

There is a complete coverage of airborne gamma radiometrics for the state of Victoria, (Australia) that has varying degrees of density with run spacings at 250 m and 400 m. GRS data is normally collected over 256 frequencies. The parent material and degree of weathering determine the signature emitted from the soil and rock. These data have been collected primarily for geological exploration, but have been extremely useful for assisting in delineating soil units at scales of 1:100 000 or broader. The relationship between GRS and soil characteristics has been documented (Cook et al. 1996). At the paddock scale the relationships between soil characteristics and airborne GRS data become too weak to be beneficial, this is primarily due to the density of typical airborne surveys.

Due to many new grain harvesters being delivered pre-fitted with yield monitors, crop yield maps are becoming quite common. The seasonal variability of these yield maps seems to offer more confusion to the land managers than solutions. As well as the relatively consistent factors that contribute to crop yield, such as pedologic and topographic attributes, there are many that vary from season to season, such as climatic, biological and anthropic factors. The complex interaction of these factors tends to limit the usefulness of crop yield maps alone (Corwin et al. 2003). Kitchen et al. (2003) noted that EMI and topographic measurements were both important in accounting for yield variation, but soil apparent electrical conductivity accounted for yield variability better than did topographic properties.

The aim of this paper is to discern which geophysical and digital elevation model derived layers best delineate the soil types how they relate to the crop yield variation across the target area. In this way we may better understand the extent that soil and topography have in the yield variability.

Materials and methods

Paddock description

The 160 ha paddock chosen for this study is in north-western Victoria in an area known as 'the Mallee'. This area has an average annual rainfall of approximately 370 mm with a growing season rainfall of approximately 250 mm. The chosen paddock has a variation in elevation of about 10 m over a distance of more than 2 km (Figure 1). The steepest slopes are 3% and average is 1%. There is considerably more relief in the northern section of the paddock than the south. For this reason the paddock was divided into two along an irrigation channel for the purposes of analysis.

A soil map was produced in the field (Figure 2) using a differential global positioning system (GPS) and on-screen mapping. Soil was initially mapped using visual characteristics such as surface colour and texture, and was then improved by assessing 40 soil cores sampled across the trial site. Assessing the variation in germination and vigour of a stand of volunteer wheat, which remained in the ungrazed paddock, assisted delineation of soil types. This paddock has a large range of soils with surface textures ranging from deep sands to heavy clays.

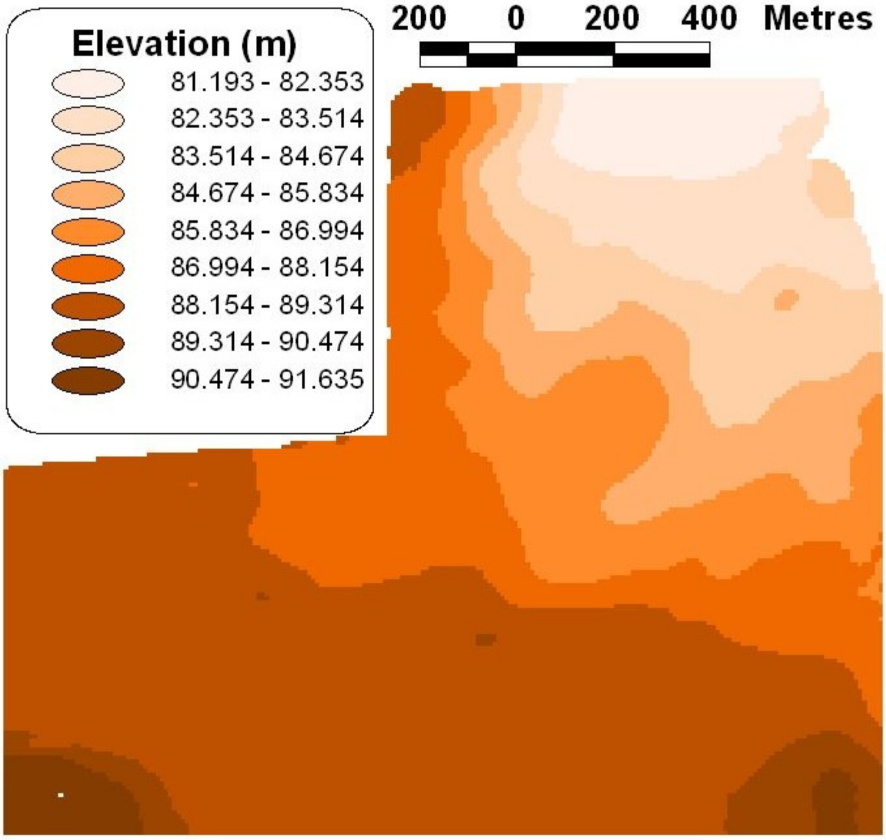

Figure 1. Elevation (ASL) of trial site

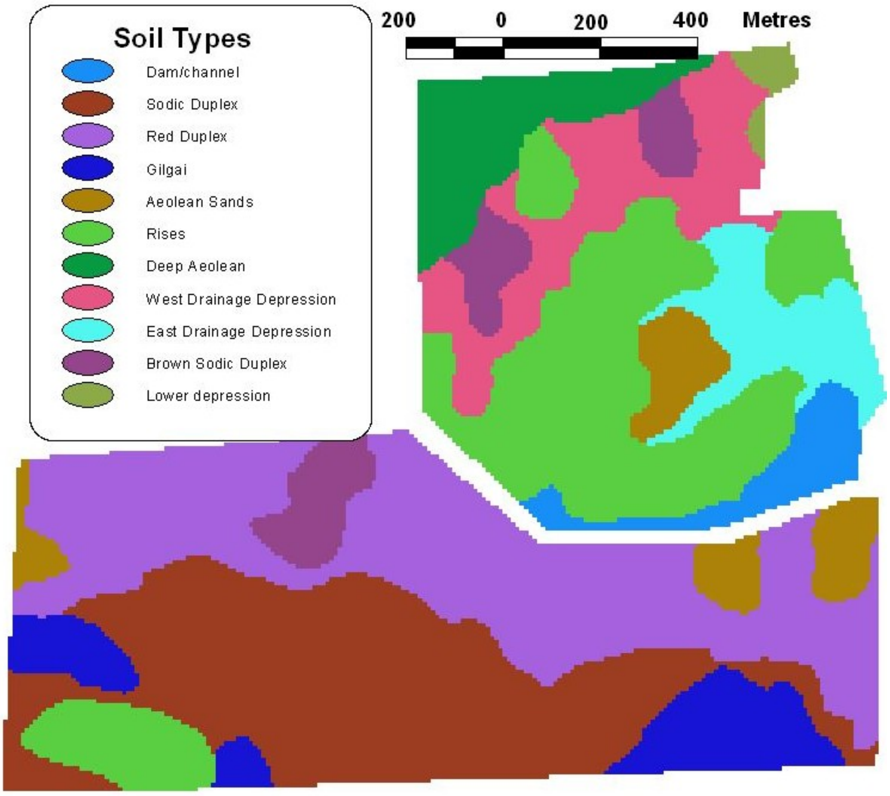

Figure 2. Simplified soil map of site

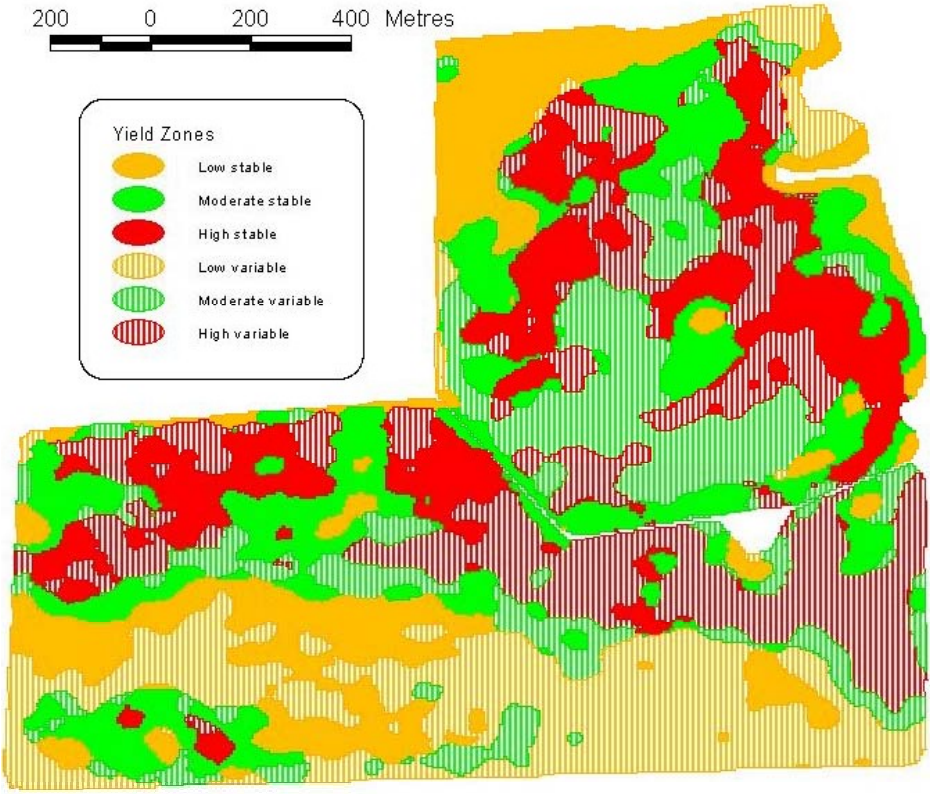

Figure 3. Yield zones as determined by grain yield and satellite data. This data was delineated by level and temporal variation.

The yield data was logged for three seasons over a seven-year period by a harvester mounted grain yield monitor. This data was combined with crop biomass information, which was derived from Landsat satellite data, for a number of years that were cropped but the yield data not collected over the same

seven-year period. This data was analysed and delineated by the level of grain yield and the variability over these years (Figure 3).
The decision tree classifier See5 (Rulequest Research 2000) was applied using a number of combinations of available datasets in order to identify the combinations that best predicted the soil units. Some major advantages of this process were: there was no assumption of linearity of the data, it could be used to classify categorical data, and the process was easy to interpret It was utilised in this case due to the non-linearity of the data and for the ability to handle continuous and continuous data to predict classified soil and yield information. This decision tree classifier uses a fairly simple binary tree classifier that simply divides the data into two branches at each node. This process was then run with different input data in an attempt to predict the classified soil and classified yield data. Due to the difference in relief between the northern and southern sections of this paddock, the data was processed separately.

Surveying/ Proximal Sensing methodology
An EMI survey was conducted with both a dual dipole Geonics EM38DD and a Geonics EM31 instrument. The EM38 was towed behind an all-terrain vehicle with both the horizontal and vertical pole data constantly logged. Traverses were spaced at 20 m intervals. The EM31 was mounted in the vertical polar mode beside the all-terrain vehicle (Figure 4).

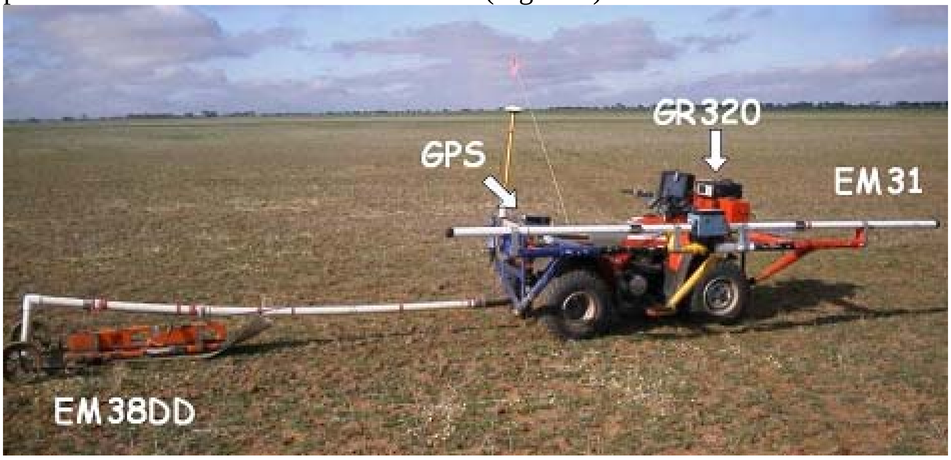

Figure 4. All-terrain vehicle fitted with the EM38, the EM31, gamma ray spectrometer (GR320) and differential GPS.

For the determination of the gamma ray spectrometry an Exploranium GR320 instrument with a large detector crystal (GPX256 which is 256 cubic inches or about 4.2L) was utilised. It was mounted on an all-terrain vehicle along with the EM38 and EM31 and the data was logged continuously.
The computer logging the geophysical data was also logging the data from a differential GPS. A NavCom Starfire 2050 GPS unit was used, as it has almost no in-field set-up time, and is accurate to about 10 cm (horizontal) when moving continuously. This GPS not only determined an accurate position of the collected geophysical data, but also provided accurate elevation data (relative accuracy of ±10cm with 95% confidence) which was later the basis of a DEM. From this DEM the terrain attributes of aspect, slope and relative elevation were determined.
In this study the relative elevation was found to be one of the most useful factors for delineating soil types. Relative elevation is a calculation from the DEM comparing the focus cell to the surrounding cells to determine if it is relatively low or elevated. In this paddock three relative elevations were generated assessing radii of 15 m, 30 m and 50 m. This indicates that this methodology is capable of determining position in gilgai if the DEM is of fine scale and high accuracy.
Conductivity data from the EM38 (horizontal and vertical poles),EM31(in vertical pole) and gamma ray spectrometry (regions of interest for the Potassium, Thorium and Uranium spectra) data were extrapolated to continuous layers using an inverse-distance weighting procedure. The potassium, thorium and uranium ranges of frequencies were extracted from the GRS data and extrapolated individually, along with the total count. The EM38 horizontal and vertical poles were treated as separate layers (Figure 5).

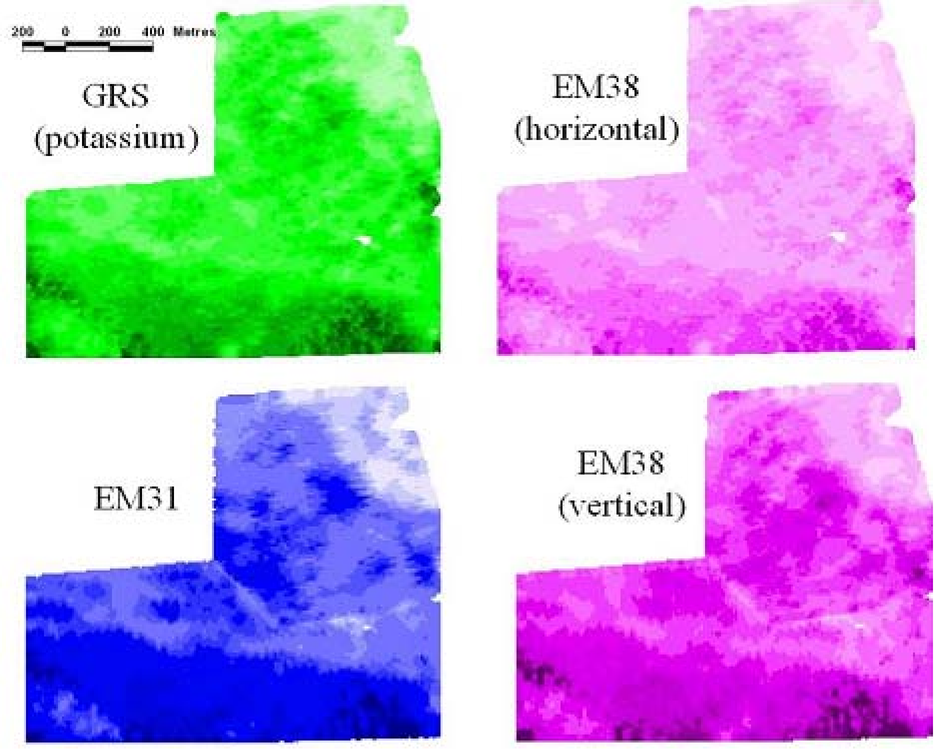

Figure 5. Extrapolated layers derived from some of the geophysical data logged on the all-terrain vehicle. Shown is the potassium range of gamma radiometrics (as GRS (potassium)), and electromagnetic induction (as EM31, EM38 (horizontal) and EM38 (vertical))

Results and discussion

The accuracy for each attempt of prediction of the target layers using different input layers is reported of as the percentage of misclassified data (Figure 6). For the purposes of this keeping this paper brief, not all the layers analysed were included in the diagrams and discussion. Examples of other layers included other regions of interest from the gamma radiometric spectrum, and other DEM derived data such as topographical wetness index and slope. These layers also showed similar trends to the layers discussed in this paper and were only omitted for brevity. This showed that EM38 (horizontal and vertical dipoles) alone incorrectly classified 21% of the area in the southern section and 35% of the North when predicting the soil types, whereas GRS alone misclassified 20% in the southern section and 18% in the north. Currently the data from EM38 surveys is being utilised to attempt to understand the variability in yield. This study demonstrates that EM38 data only partially represents the soil types. EM31 determines the soil apparent electrical conductivity to 6 m and is slightly less affected by the surface soils than is the EM38. When the EM31 data was processed with the EM38 data the errors dropped to 15% in the south and 22% in the north, showing that the apparent conductivity of the soil beneath the root zone still assists with prediction of the soil within the upper 1.5 m. When EM38 and GRS were used together to predict the soil types the misclassification dropped to 10% across the whole paddock. Combining the data from both of the EMI tools and the GRS reduced the errors to 7.4%. Elevation alone produced errors of greater than 38%; however, when used in conjunction with the other DEM derived terrain attributes the errors were less than 3%. When all of the terrain attributes, including elevation, were used in conjunction with any or all of the geophysical datasets the errors were 1.4% in the south and 1.8% in the northern section. This tends to indicate that the use of individual geophysical layers will assist in the delineation of soil types and hence the understanding of crop yield variability. This relationship, however, is limited and could be significantly enhanced with the inclusion of the terrain layers that may be derived from data simultaneously collected, and is often overlooked.

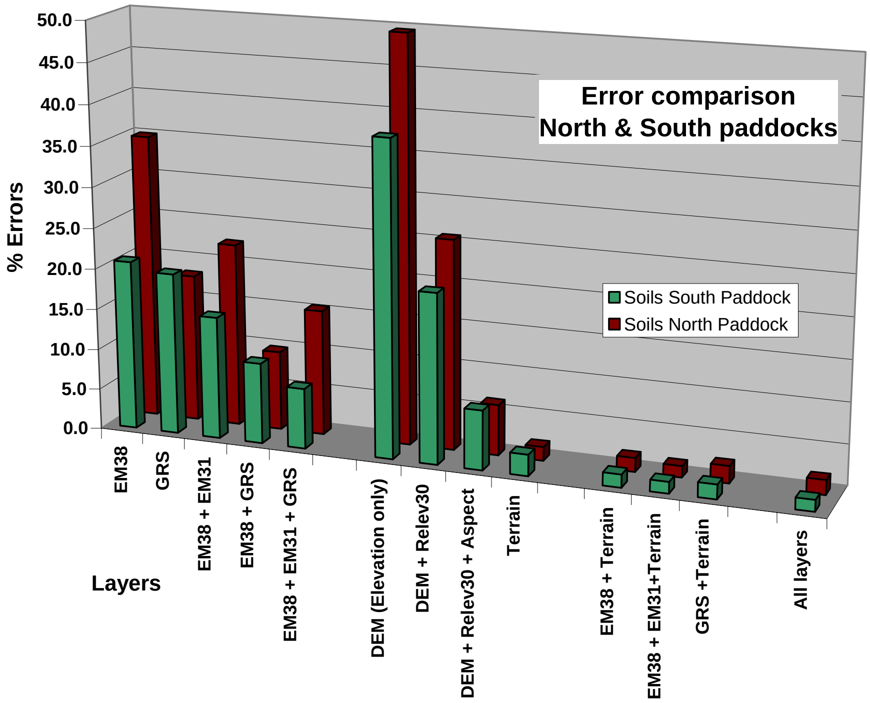

Figure 6. Chart indicating percentage of errors when predicting soil types for individual sections of paddock using See5 with different combinations of data. (EM38 includes both horizontal and vertical dipole, GRS = gamma ray spectrometry, DEM = elevation only derived from the digital elevation model, Relev30 = relative elevation at 30 m, Terrain = elevation and all relative elevations and aspect.)

Although there was low relief at this site there was still a strong relationship between the soil units and the landform. This was evident from the low errors when See5 processed all the terrain attributes with no geophysical data. The northern section of the paddock had greater relief than the southern section but the errors on predicting the soil types from terrain alone were only marginally higher in the southern paddock. When the terrain attributes were added to the geophysical data the errors were very low. In areas of less relief, these terrain variables may not be as useful; however, this study indicates that, even in areas of low relief the terrain attributes still add considerable to the delineation of soil units. Contractors often collect elevation data when conducting an EM survey, this may be overlooked when land managers are attempting to interpret the EM or yield data.

Further validation would be required in order to determine the true error of prediction of these soil types. At present the error only indicates how well See5 can predict the soil types as mapped in the paddock using the collected data. The decision tree may better define the soil types than is represented by the original soil mapping, as the original soil mapping is a single surveyor's interpretation of the distribution of the soils across the paddock.

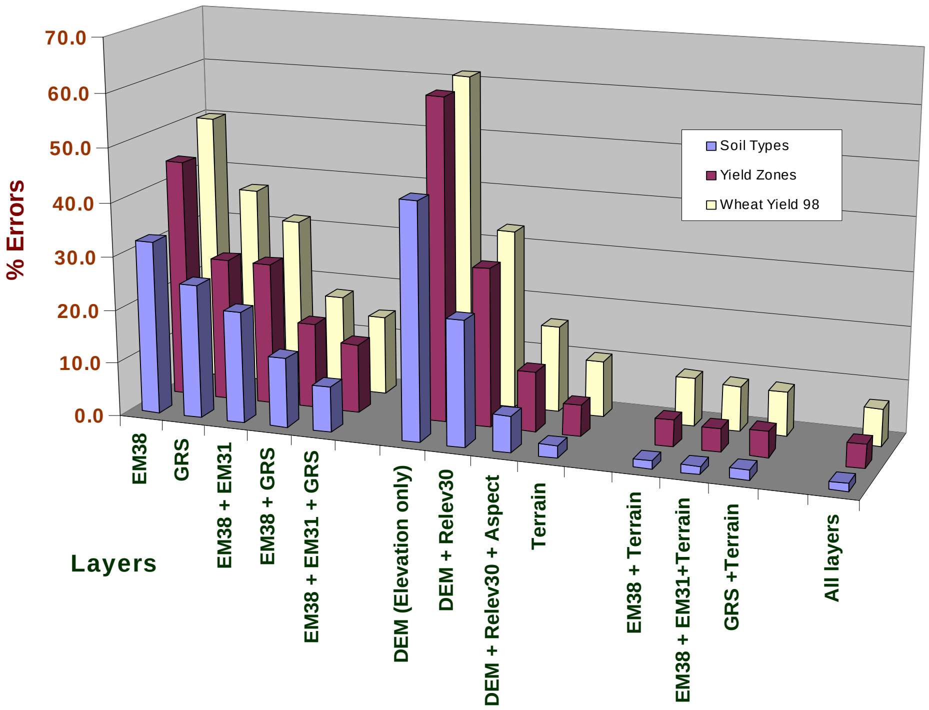

Figure 7. Errors on prediction of soil types, yield zones and wheat yield for 1998 with combinations of collected data for the entire paddock. (EM38 includes both horizontal and vertical dipole, GRS = gamma ray spectrometry, DEM = elevation only derived from the digital elevation model, Relev30 = relative elevation at 30 m, Terrain = elevation and all relative elevations and aspect.)

The prediction of yield for any individual year is would be less accurate than the prediction of the temporal yield classification (yield zones). The temporal classification effectively attenuates the effect on yield of unusual climate in any one year. When using the geophysical and terrain data to predict the yield zones and the yield values for one year (1998), the general trend of the prediction of the soil types was the same for the various combinations of data layers (Figure 7). However the area misclassified increases from 1.5% when predicting the soils with all the data to 4.4% when predicting the yield zones and increases again to 6.9% when predicting the yield for a single year. This result is to be expected but it also confirms that the primary long term yield variation is due to the underlying soil and landscape characteristics.

Conclusion

In the southern section of the paddock the GRS alone provided a much better predictor of the soil types than EM38 alone whereas in the northern section they were quite similar. Overall GRS was a better predictor of soil type and yield data than EM38. Although individually EM38 and GRS data are useful for delineating soil types at paddock scale, when combined they form a much stronger relationship (generally improving the accuracy of prediction by about 10%). When combined with the terrain attributes derived from the DEM, the accuracy of predicting the soil types decreased the misclassification to less than 2%. In this case, although subtle, the landform features corresponded well to the soil types with the terrain attributes alone causing up to 2.6% errors in prediction. Considering the minimal added cost and time to produce a DEM and create suitable derivatives, this work clearly shows the potential of incorporating topographic data with geophysical data when delineating soil types. In fact, in this study, the terrain attributes alone predicted the soil types better than all the geophysical data combined. This may not be the case on land with less relief. Terrain characteristics are add considerably to the accuracy of both soil types and yield and should be used to assist in the interpretation of geophysical data for the purposes of understanding yield variation in precision agriculture.

References

Cameron DR, de Jong E, Read DWL, Oosterveld M (1981) Mapping salinity using resistivity and electromagnetic techniques. *Canadian Journal of Soil Science* **61**, 67–78.

Cook SE, Corner R, Groves, Grealish GJ (1996) Use of airborne gamma radiometric data for soil sampling. *Australian Journal of Soil Research* **34**, 183–194.

Corwin DL, Lesch SM, Shouse PJ, Soppe R, Ayars JE (2003) Identifying soil properties that influence cotton yield using soil sampling directed by apparent soil electrical conductivity. *Agronomy Journal* **95**, 352–364.

Doolittle JA, Sudduth KA, Kitchen NR, Indorante SJ (1994) Estimating depths to claypans using electromagnetic induction methods. *Journal of Soil and Water Conservation* **49** 572–575.

Johnson CK, Eskridge KM, Wienhold BJ, Doran JW, Peterson GA, Buchleiter GW (2003) Using electrical conductivity classification and within-field variability to design field-scale research. *Agronomy Journal* **95,** 602–613.

Kitchen NR, Drummond ST, Lund ED, Buchleiter GW (2003) Soil electrical conductivity and topography related to yield for three contrasting soilcrop systems. *Agronomy Journal* **95**, 483–495.

Sudduth KA, Drummond ST, Kitchen NR (2001). Accuracy issues electromagnetic sensing of soil electrical conductivity for precision agriculture. *Computers and Electronics in Agriculture* **31**, 239–264.

Rulequest Research, (2000) See5, Rulequest Research, Sydney, Australia.