Digital soil-landscape mapping by image clustering

Daniel Brough[A,B] and Robin Thwaites[B]

[A]Department of Environment and Resource Management, Indooroopilly, QLD, Australia, Email daniel.brough@derm.qld.gov.au
[B] School of Natural Resources, Queensland University of Technology, Brisbane, Australia, Email r.thwaites@qut.edu.au

Abstract
This paper examines a novel approach to examining soil-landscapes for digital soil mapping by using image clustering approaches. The move to digital soil mapping is creating a paradigm shift for many soil surveyors and is relying on new tools and techniques. In Queensland, and more generally world-wide, the requirements to support many environmental objectives are “stretching” the existing land resource information base to the limits of its intended purpose when collected over the last half century. In an attempt to utilise the knowledge contained in legacy mapping of Land Systems in Queensland, image clustering techniques have been tested to identify their suitability to spatially disaggregate existing mapping. Since a Land System, is a recurring pattern of soils, terrain, geology and vegetation an approach that identifies spatial patterns and digital objects is thought to provide a method of extracting this knowledge for inclusion in digital soil mapping approaches in the Inland Burnett Catchment.

Key Words
Digital Soil Mapping, Clustering, Land Systems.

Introduction
Recent improvements in computation, information technology and the types of tools available to soil scientists have led to paradigm shift in land resource assessment. Soil survey and land resource assessment is now moving towards digital soil mapping. With digital soil mapping is moving out of the research phase to become a semi-routine technique that is used in land resource assessment (McBratney *et al*. 2003). The increased access and use of geographical information systems (GIS), digital elevation models, geophysical tools, remotely sensed data and a myriad of other datasets has created a trend towards more quantitative resource assessment. The use of new tools and techniques has raised the awareness of the need to capture and communicate the knowledge gained by surveyors (Bui 2004).

With new natural resource management requirements it has become apparent that existing qualitative and non-digital land resource information needs to be re-interpreted. The non-sustainable management of natural resources has a significant impact on our quality of life (Hillel 2000). The small-scale Land Systems mapping of Queensland and Northern Australia is a prime example of this type of data needing re-interpretation. This re-interpretation fulfils the need to provide more quantitative soil-landscape attribute information at improved scales and resolution. A digital soil-landscape process is used to disaggregate Land Systems mapping in the Burnett Catchment of South East Queensland (Figure 1).

Methods
Existing Land Resource Mapping
To overcome the lack of information available following the Second Wold War when there was increasing development pressures in Northern Australia, Christian and Stewart (1953, 1968) introduced the concept of land systems mapping. They reason that land systems mapping would allow for a reduction in effort to gather knowledge about an area but would provide a broadscale framework on which further intensive studies can be undertaken in areas where special features are important. They rationalise that a hierarchical approach to sub-dividing and describing the landscape during a reconnaissance survey is best suited to covering large tracts of land and providing a framework for the description of the landscapes. In the study area there are two existing land system mapping projects that are complemented by two 1:50,000 soil surveys.

Environmental Correlation Datasets
Several datasets for environmental correlation exist for the catchment, including a Digital Elevation Model (DEM), gamma-ray spectrometry, geology and climate data. DEMs and airborne geophysics have been shown as useful tools for predicting soil-landscape attributes in digital soil mapping (DSM) studies.

The DEM has a pixel size of 25 metres and was built in ANUDEM (Hutchinson 1989) from 1:100,000 scale topographic data that is hydrologically correct. A range of derivatives have been generated to use as environmental covariates, these include slope, topographic wetness index, relative elevation and curvatures.

While the list of DEM derivatives listed above is by no means exhaustive they have previously been found to be useful for DSM, for example Ziadat (2005) used DEM derivatives to predict soil attributes in Jordan.

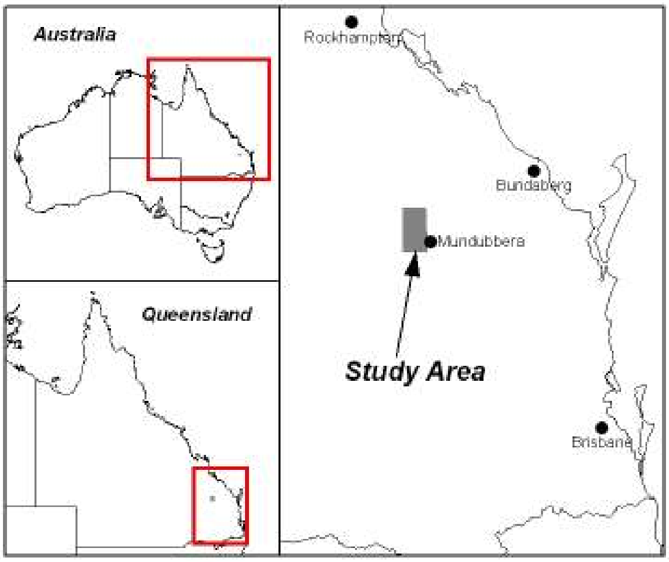

Figure 1. Study area location.

Airborne gamma-ray spectrometry (radiometrics) measures the abundance of Potassium (K), Thorium (Th) and Uranium (U) in soils by detecting the gamma-rays emitted due to the natural radioelement decay of these elements (Wilford 2002). relates to the parent material and geochemistry of the soil (and other weathered materials). Weathering modifies the distribution and concentration of radioelements compared to the original bedrock. Understanding the bedrock and soil responses has proven invaluable not only for mapping soils based on parent materials but also for understanding geomorphic processes (Wilford, Bierwirth and Craig 1997). K, Th and U behave quite differently under weathering situations. Cook *et al.* (1996) utilised radiometric data for digital soil mapping in Western Australia.

The environmental correlation datasets where analysed with both statistical methods and expert opinion as to there usefulness for clustering. The main statistical technique used was an analysis of variance and some decision tree methods to further evaluate the utility of the large number of environmental correlation layers and propose those to be used for clustering. Expert opinion was included as an overarching method to validate the statistical techniques and ensure the correlation dataset actually had a meaning to the soil surveyor.

Each of the environmental datasets required manipulation to set them to a standardised extent and cell size, but also to a standard data range (0 - 255). The standard data range ensures no one variable can significantly outweigh the others when calculating variance in “attribute space”. A number of the datasets where also manipulated to adjust their distribution to limit the effects of skewed data. A prime example of this is Slope (as derived from the DEM) that has a significant skew in the towards flatter areas with a small number of high slope areas (Figure 2).

Image Clustering Method
The spatial disaggregation methodologies applied to the Inland Burnett Catchment area based on a Multiscale Object Modelling approach developed over a number of years for the analysis of remote sensing data, including medical imaging. Hay *et al.* (2005) developed a method (Multiscale Object Specific Segmentation - MOSS) which involves three specific stages in the development of a series of multiscale polygons for forest inventory. The three stages are object analysis, object upscaling and finally the merging

of regions.

The merging step is named Size Constrained Region Merging (SCRM), since the regions created are determined by a user defined size parameter and was proposed by Castilla (2003) to handle multiple bands in a Landsat Image. SCRM is an image smoothing and merging mechanism that produces features that represent individual image objects and maybe converted to a vector layer with associated attributes compiled against it.

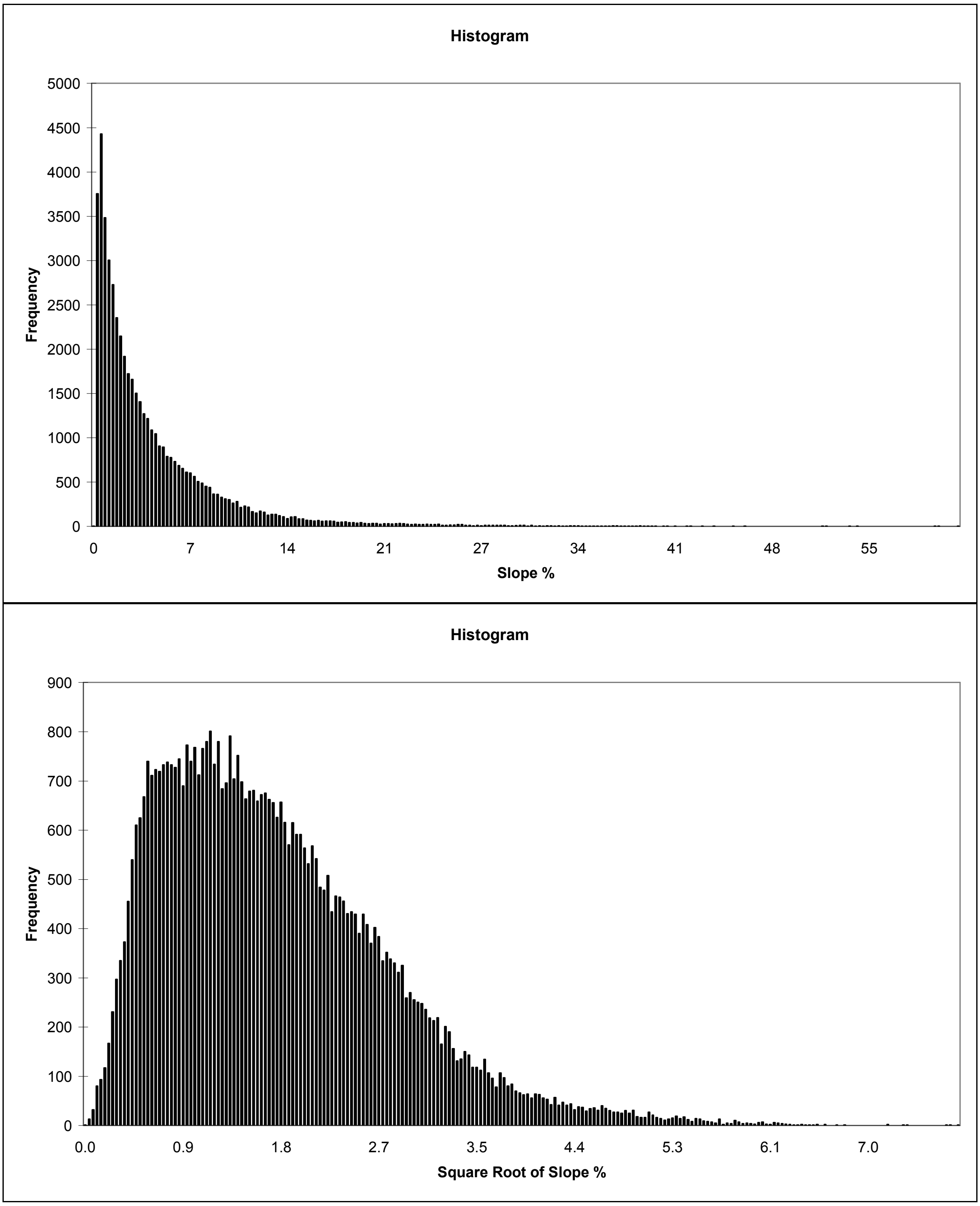

Figure 2. Histograms of DEM derived Slope percentage and the transformed slope.

Digital Soil Modelling

The most appropriate method for digital soil modelling based on image clustering techniques is still being evaluated. Several techniques exist to model and map the improved soil information in the pilot area of the Burnett Catchment. The preference for soil attribute modelling would be to use existing knowledge from land systems survey to identify areas that are known. Several methods exist for the DSM component of this project; these include the ASRIS methodology (McKenzie *et al.* 2005) to define attributes levels for soils, land units and land systems that may be complemented by extrapolation with fuzzy or Bayesian models. The use of tree based approaches used by Zhu *et al*. (2004) may also be valid, especially with the parent-child relationship that exists from the multi-scale polygons identified by the MOSS method.

Conclusion

The Multiscale Object Specific Segmentation method has the ability to provide another tool in the set of methods used for DSM. The ability to identify the environmental correlation datasets used impact greatly on the usefulness of the method. With the right layers selected, the method has the ability to rapidly dissect the landscape into units through a similar process to Air Photo Interpretation. The rapid building of landscape units has the potential to improve the speed at which DSM assessments can be undertaken in areas with existing small-scale land resource information.

References

Cook SE, Corner RJ, Groves PR, Grealish GJ (1996) Use of airborne gamma radiometric data for soil mapping. *Australian Journal of Soil Research* **34**, 183-194.

Bui EN (2004) Soil survey as a knowledge system. *Geoderma* **120**, 17-26.

Castilla G (2003) Object-oriented analysis of Remote Sensing images for land cover mapping: conceptual foundations and a segmentation method to derive a baseline partition for classification, PhD edn, Polythecnic University of Madrid, Madrid, Spain.

Christian CS, Stewart GA (1968) Methodology of Integrated Surveys. In 'Aerial Surveys and Integrated Surveys, Proceedings Toulouse Conference 1964'. Toulouse. pp. 233-280. (UNESCO).

Christian CS, Stewart GA (1953) 'General Report on Survey of Katherine - Darwin Region, 1946'. (CSIRO: Australia).

Cook SE, Corner RJ, Groves PR, Grealish GJ (1996) Use of airborne gamma radiometric data for soil mapping. *Australian Journal of Soil Research* **34**, 183-194.

Hay GJ, Castilla G, Ruiz JR, Wulder MA (2005) "An automated object-based approach for the multiscale image segmentation of forest scenes", *International Journal of Applied Earth Observation and Geoinformation* **7**(4), 339-359.

Hillel D (2000) 'Salinity Management for Sustainable Irrigation'. (World Bank Publications).

Hutchinson MF (1989) A new procedure for gridding elevation and stream line data with automatic removal of spurious pits. *Journal of Hydrology* **106**, 211-232.

McBratney AB, Minasny B, Mendonça Santos ML (2003) On digital soil mapping. *Geoderma* **117**, 3-52.

McKenzie NJ, Jacquier DW, Maschmedt DJ, Griffin EA, Brough DM (2005) *Australian Soil Resource Information System Technical Specifications, Version 1.5,* Australian Collaborative Land Evaluation Program, Canberra.

Wilford J (2002) Airborne Gamma-ray Spectrometry. In 'Geophysical and Remote Sensing Methods for Regloith Exploration'. (Ed. E Papp) pp. 46-52. (CRC LEME: Canberra, Australia).

Wilford JR, Bierwirth PN, Craig MA (1997) Application of airborne gamma-ray spectrometry in soil/regolith mapping and applied geomorphology. *AGSO Journal of Australian Geology and Geophysics* **17**, 201-216.

Zhu J, Morgan CLS, Norman JM, Yue W, Lowery B (2004) "Combined mapping of soil properties using a multi-scale tree-structured spatial model" *Geoderma* **118**(3-4), 321-334.

Ziadat FM (2005) Analyzing digital terrain attributes to predict soil attributes for a relatively large area. *Soil Science Society of America Journal* **69**, 1590-1599.