Systems integration: enhancing soil information delivery to Victoria

Pat Johnstone[A], Nathan Robinson[C], Mark Imhof[B], David Rees[D], David Hunter[C], Doug Crawford[B], Stuart Wells[A], Ron Walsh[A], Angela Fadersen[C], Steve Williams[C] and Matt Kitching[A]

[A]Future Farming Systems Research Division, Department of Primary Industries, Werribee, 621 Sneydes Road, Werribee, Victoria, 3030, Australia, Email pat.johnstone@dpi.vic.gov.au
[B]Future Farming Systems Research Division, Department of Primary Industries, Ellinbank, 1301 Hazeldean Road, Ellinbank, Victoria, 3821, Australia, Email mark.imhof@dpi.vic.gov.au, doug.crawford@dpi.vic.gov.au
[C]Future Farming Systems Research Division, Department of Primary Industries, Epsom, Cnr Midland Hwy & Taylor Street, Bendigo, Victoria, 3554, Australia, Email david.hunter@dpi.vic.gov.au
[D]Future Farming Systems Research Division, Department of Primary Industries, Parkville, 32 Lincoln Square North, Parkville, Victoria, 3053, Australia, Email david.rees@dpi.vic.gov.au

Abstract
Data, process management and information sharing are among the biggest challenges facing the scientific community today. The integration of heterogeneous primary and derived data sets arising from soil and land survey activities is imperative to provide easily accessible information required for informed decision making regarding future farming systems and their impact on the environment. Currently the Department of Primary Industries, Victoria (DPI) utilises a suite of software applications in tandem to maintain and manage its digital and hardcopy land resource data sets. An overall soil survey information strategy is evolving to enable the multiple systems to work together as an integrated system. The integration of these systems will provide a sound foundation for building a comprehensive knowledge management strategy based on both current and historical data sets. This will result in improved access, increased usability, efficient information exchange and improved standardisation and consistency of data and its derivatives. Rapid access to samples and data is being used for calibration and prediction of soil chemical and physical properties using MIR. Full integration of these systems will provide an efficient foundation for building a comprehensive knowledge management strategy. This paper presents the concept for the integration of the diverse data sets that comprise DPI's soil site data collection.

Key Words
Data and information management, legacy data, soil archive.

Introduction
DPI and its predecessors have collected soil site data over many years and these sites remain relevant today but not readily accessible for widespread utilisation. DPI's collection comprises physical samples, as well as hardcopy and digital records stored in a number of current and defunct systems. A multidisciplinary team, involving soil scientists, soil surveyors, spatial information scientists, quality systems specialists and laboratory information management systems (LIMS) specialists is developing a systematic framework to enable the integration of soil site data. The Victorian soil site archive comprises almost 30,000 physical soil samples (derived from soil survey and research studies carried out over the past 80 years). Associated with these samples are records and analytical data sets held either as hardcopy or in digital form in a variety of systems. DPI is creating a durable soil sample archive and associated digital inventory to optimise access and utilisation of current and historical samples into the future. A systematic process for the digital capture and storage of ancillary data (consisting of many hundreds of air photos (annotated with soil mapping boundaries and sites identified), field books, working maps and analytical data) is being developed. Preserving this material through digital capture will enable future entry of data associated with many thousands of soil sites across Victoria into a soil information system (SIS) while preserving part of DPI soil and land survey history.

Methods
Victorian Soil Archive
The soil sample archive comprises almost 30,000 soil samples that are currently stored within more than 50 wooden crates (1.2 m x 1.2 m x 0.6 m) containing cardboard boxes and containers of soil labelled only with sample number. A variety of plastic and cardboard containers, which were never intended for long term storage, have been used over the years and many of these are deteriorating with age. At present there is only a manual list of the sample numbers stored in each crate. The samples are being progressively transferred to archive quality containers (high-density polyethylene white jars with white polypropylene screw caps). An

1 – 6 August 2010, Brisbane, Australia. Published on DVD.

inventory is being developed and a process of linking the sample number to the ancillary survey record is being undertaken to enable georeferencing and linkage to the Victorian Soils Information System (VSIS). Once linkages to the original survey and geographic references have been established the samples will be labelled with a barcode and key sample information as per protocols established for the Australian Soil Information System (ASRIS) using an archive quality label.

Historical records

The data collected from previous soil/landform and agricultural and landscape research sampling programmes consists of landform, morphological, physical, chemical and biological descriptor records stored as hardcopy reports as well as maps, aerial photos, field books, site and locality summary sheets, and associated data tables (examples are provided in Figure 1). Although summaries of this information have been collated into hardcopy reports or linked to landscape classification systems such as Land Systems of Victoria Zones, much of the raw data needed for ongoing research, monitoring and systems modelling remains in poorly maintained paper-based archives. Records from 1930 to the present day are currently stored in a range of boxes and filing cabinets. The survey methods have been evolving over time and as such earlier survey data may be incompatible or require interpretation and translation to align to current soil survey data standards. Consequently not all historical survey data is of equal utility or value.

Figure 1. Example of page details in hardcopy field books; Aerial photo with site details and locations identified; Historic soil survey maps; Agriculture lab submission sheets. These records provide opportunities to track sample numbers associated with soil survey sites and link them to their laboratory data and geographic location.

Victorian Resources Online (VRO)

DPI's Victorian Resources Online (VRO) website is the key portal for accessing information about Victoria's natural resources (e.g. soil, landform, water, biodiversity) and their management. It contains maps and associated information, downloadable reports as well as many related links to other information sources. A wide range of information products generated from Victorian soil and land survey work over the past 80 years is made available online. The website is a repository for many contemporary and historical reports. Many of these have been digitally captured (scanned and/or re-typed) and made available as downloadable pdf format reports that have been customised to optimise loading and readability. A 'Soil and Land Survey Directory' facilitates searching of over 100 soil and land survey reports via key words or by a specific Local Government Area or Catchment Management region (refer to Figure 2).

Figure 2. Victorian Catchment Management Authority regions.

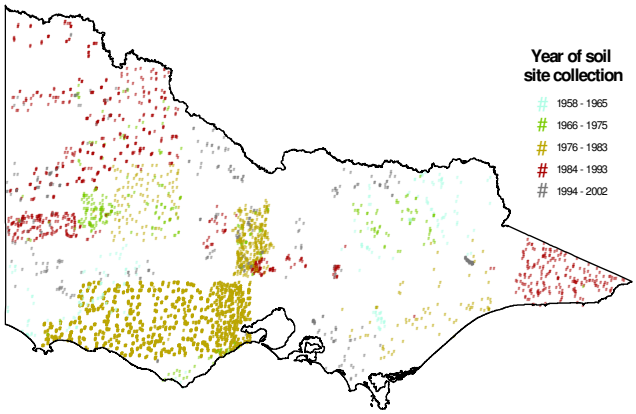

Figure 3. Soil site locations and periods (years of collection, for site data currently in VSIS).

Victorian Soil Information System (VSIS)
The VSIS is being used to capture soil site data (both field and laboratory generated) in a systematic way that will allow the data to be stored, utilised and mapped by a diverse range of users at a local, regional, state and national scales. It is a primary centralised database that provides access to soil site data that adheres to national soil and land survey standards. VSIS consists of an entry module (Victorian Soil Entry System - VSES) used to ensure data adheres to current national standards for soil and land survey guidelines (McKenzie *et al*. 2008) and for field survey (National Committee on Soil and Terrain 2009), and an output (query) module to interrogate the data. Key historical soil sites are progressively georeferenced and then manually entered into the Victorian Soil Entry System (VSES). VSIS contains "high quality" data for 2,800 sites (refer to Figure 3). A further 1000 soil sites have been entered in the VSES and are under going quality checking protocols prior to upload into VSIS.

Laboratory Information System (LIMS)
The DPI-Werribee laboratory utilises a LIMS (relational database hosted on a SQL server) to manage all soil samples that have been submitted for analytical testing. This system is used to register, label, track, report and invoice all samples submitted for testing. Data is stored and tracked in the system using the sample number as the key identifier. Although the system can store geographic reference information it was seldom provided by those submitting samples for analysis. Additional information provided by the submitter for quality checking purposes, including references to the original soil site location and observations was often stored in a free form text field. Unfortunately, due to the myriad of information and formats used in the free text field it is of little use for linking data to the soil sites. Mostly it provides additional information for checking the links established by other mechanisms (e.g. field books, original soil observation and description cards, or survey indexes). Analytical extraction procedures and methods have evolved with advances in technology. During data extraction analyst knowledge and method references are utilised to describe the analytical process in terms of established National Standards (Rayment and Higginson 1992) for soil testing. The LIMS has operated through implementation of two software platforms including LabWare (current system containing records for soil testing from 2006 to the present day) and SMS (defunct system containing records for soil testing from 1988 to 2006).

Data Track
Data Track is a DPI business initiative to further develop policies, procedures and protocols for data storage and management to improve the security and custodianship of existing soil data assets and improve their accessibility, utility and quality. An important part of this process is the development business capability, capacity and intelligence to ensure best value is derived from the DPI soil data assets. Data standards and management protocols are being used to improve the soil information data quality, consistency and availability and thereby preserve them for future users. The initial design and development of the VSIS was fostered within this initiative to provide a flagship model for management of a major data area for the DPI Future Farming Systems Research (FFSR).

Results
The Soil and Land Directory will be a basis for inventory and categorisation of the hardcopy records not yet digitised. The information is progressively inventoried and categorised with respect to Victoria's Catchment Management Regions and then scanned and stored. Three key surveys from the Corangamite, Mallee and East Gippsland regions have been used to refine and develop the system for integrating this information. Samples and records are linked to the original records/maps and then georeferenced to facilitate entry into VSIS. Electronic links are being established between VSIS, VRO and LIMS-SMS / LIMS-LabWare outlined in Figure 4. These links will provide considerable gains in efficiency and accuracy for the further inclusion of historical information from 1988 in VSIS by significantly reducing manual data entry. A mechanism for the provision of a VSIS link during sample registration for all future soil site samples will ensure efficient information exchange between the systems.

Mid infrared (MIR) spectroscopy is being used by DPI as a powerful tool for rapid multi-parameter screening of soil properties. Soil chemical and physical analyses are time consuming and becoming increasingly expensive. Recent research has demonstrated the utility of using MIR spectroscopy to complement and perhaps one day replace aspects of traditional soil analysis. MIR calibrations are being developed from the chemical and physical data that has already been collected for the samples in the DPI Victorian Soil Archive for the major Victorian agricultural soils. In conjunction with this, MIR will be used

to generate results for other analytes that were not part of the original analytical data set. The georeferencing of these samples will enable mapping of soil properties such as pH, EC, exchangeable cations, organic carbon, total nitrogen, and physical properties such as plant available water capacity, particle size analysis and even clay mineralogy once satisfactory calibrations have been established.

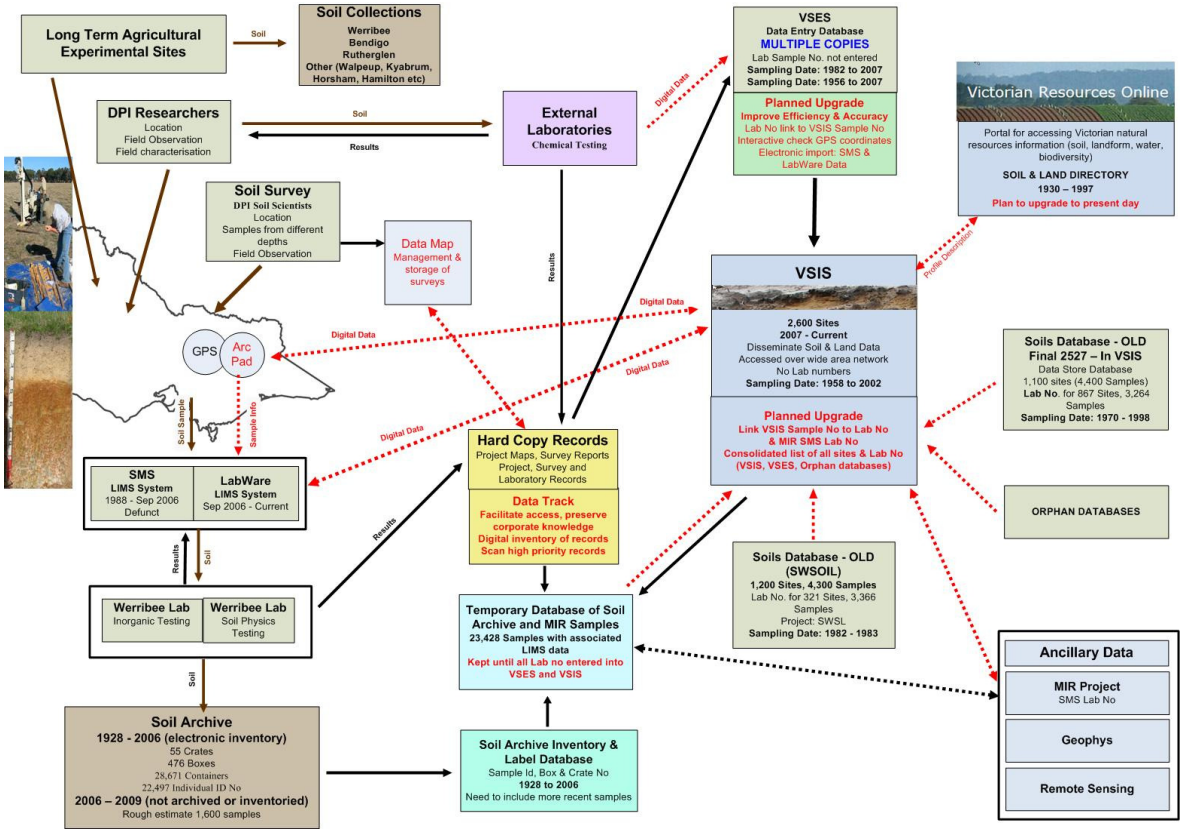

Figure 4. Relationships between soil survey data and archive management systems for Department of Primary Industries, Victoria.

Conclusion

Consideration of future farming systems and the impact on the environment is increasingly using modelling, underpinned by data to understand land use impacts and increase productivity in response to environmental factors including climate change. The cost of developing new data sets has demanded re-analysis and comparative analysis of existing soil samples and information. Currently there exists only a manual record of these samples. The biggest challenge to systems integration currently is the volume of historical samples to process and archive and the diversity of record keeping and data management systems that were used. Linking soil samples to original survey and sample number, and then establishing geographic references is proving most challenging. Most survey reports did not include the sample number in the final tables and so ancillary records need to be traced. An electronic inventory will greatly facilitate this process. The integration of historical and laboratory systems (LIMS) with VSIS and VRO will support the growing need for raw data to support research, monitoring, modelling and soil health initiatives.

References

McKenzie NJ, Grundy MJ, Webster R, Ringrose-Voase AJ (2008) 'Guidelines for Surveying Soil and Land Resources. Australian Soil and Land Survey Handbooks Series Volume 2'. (CSIRO Publishing: Collingwood).

National Soil Committee on Terrain (2009) 'Australian Soil and Land Survey Field Handbook, Third Edition, Australian Soil and Land Survey Handbooks Series Volume 1' (CSIRO Publishing)

Rayment G E, Higginson FR (1992) 'Australian Laboratory Handbook of Soil and Water Chemical Methods' (Inkata Press: Port Melbourne).