A conceptual framework for a Victorian soil-landscape inference system

Nathan Robinson[A], Steve Williams [B], Richard MacEwan [C]

[A]Department of Primary Industries, Epsom, Victoria, Australia, Email Nathan.Robinson@dpi.vic.gov.au
[B]Future Farming Systems Research, Department of Primary Industries, Epsom, VIC, Australia, Email Steve.Williams@dpi.vic.gov.au
[C]Future Farming Systems Research, Department of Primary Industries, Epsom, VIC, Australia, Email Richard.MacEwan@dpi.vic.gov.au

Abstract
The demand for soil information over varying spatial and temporal extents to supply a range of modelling, planning and reporting needs is increasing. A soil-landscape inference system (derived from soil inference systems – SINFERS) provides a system for managing the evolution of digital soil mapping products including spatially continuous or classified soil properties in a logical and ordered manner. The Victorian Soil-Landscape Inference System (VSLIS) concept is a three component model with input, inference and output (or product) phases that resemble a basic and logical workflow. VSLIS will deliver estimates of uncertainty for soil parameters in high demand by modellers, and deliver information products to users via the Victorian Resources Online (VRO) web portal using the Victorian Soil Information System (VSIS) as its data engine. The VSLIS operation will require multiple inference methods, even for the same soil parameter, to be spatially assigned and constrained to satisfy Victoria's physiographic and pedological diversity and input data qualities.

Key Words
Soils, inference systems, modelling, data.

Introduction
Many biophysical and socioeconomic models require soil data to predict agricultural production and potential environmental impacts. Increasingly the demand for soil information has shifted from point measurements (e.g. soil site) to spatially continuous soil parameter surfaces at varying scales. A major limitation for modelling farming system impacts to the soil resource is the paucity of soil site data in the Victorian Soil Information System (VSIS) (MacEwan 2007). The utility of soil information collected in previous government projects may be limited depending upon the scale and purpose of the original survey (e.g. land resource assessment, land capability assessment, irrigation survey, research investigation) and therefore potentially inappropriate for new applications.

Soil inference systems (SINFERS) (Dale *et al.* 1989) were refined by McBratney *et al.* (2002) when exploring the use of Pedo-Transfer Functions (PTF's) as the knowledge rules for inference systems. Soil–landscape inference systems are derived from soil spatial inference (Lagacherie and McBratney 2007). They consist of procedures that use soil and landscape ancillary datasets to derive spatially continuous or classified soil properties. These provide the basis for inferring soil variables including soil functions,
The VSLIS is largely supported through the provision of ancillary datasets or predictive environment covariates (e.g. radiometrics, digital elevation models, satellite imagery, soil maps etc.) in addition to geo-referenced soil observation data of various forms (e.g. laboratory analysis, morphology descriptions, spectroscopy, hydraulic properties and time-sequence measurements). The provision of this data must be serviced by quality systems to properly support the VSLIS. Inadequate ability to access, integrate and manipulate soil data has associated impacts on derivative information products including modelling parameters. These limitations are being addressed through refinement of the VSIS (Williams *et al.* 2009a) and information integration of the broader spectrum of soil information resources from associated systems. A soil–landscape inference system can be self-updating and provide users with the latest system predictions from most recent soil observations stored in the VSIS and ancillary datasets. This will ensure modellers will always receive the most up-to-date inference system predictions.

Victoria's landscapes and current soil information status
Victoria has a diversity of landscapes, reflecting many different processes acting on the land within the earth's crust over a considerable span of time (Jenkin 1982). Landscapes are shaped by geomorphic and pedogenic processes operating at microscopic to megascopic scales'. Approaches to predict the occurrence and uncertainty of soil-landscape parameters will need to be sensitive to both Victoria's physiographic and pedological diversity, and the availability of input data and its inherent qualities (McBratney *et al.* 2002).

Current status of soil and land data (sites and maps)

In Victoria there are over 350 documented soil and land surveys and studies that have been undertaken during the last 80 years. These surveys range in scale from 1:10 000 (large scale soil survey) to 1:250 000 (small scale soil/landform and land systems).

Soil site data from past surveys is being continually entered into the VSIS. Advances in the design of the Australian Soil Resource Information System (ASRIS) have been integrated into the VSIS. The resultant data model shares much in common with the thinking and concepts underpinning the OpenGIS implementation standard for Observations and Measurements (O&M) (OGC 2007). Currently there are 2800 higher quality soil sites in VSIS with a process to further enter soil site data across the state.

The established need for soils data in Victoria

A workshop 'Soil landscape parameters for modelling' held in Bendigo in March 2009 identified the critical (sensitive) soil parameters needed for models used by the Victorian government (Table 1). The key soil parameters for models relative to their sensitivity or dependence in these models are linked to modelling domains (hydrological, growth, carbon and other). Many of the hydrological and physical parameters are common to most model domains. The VSLIS will be used to estimate uncertainty for soil parameters in high demand by modellers, and deliver these data to users.

Table 1. Landscape model domain sensitivity to soil parameters (source Victorian DPI)

Hydrological							Growth							Carbon							Other	
Air-dry moisture content	Critical Lower Limit / Permanent Wilting Point	Drained Upper Limit / Field Capacity	K_{sat}	Infiltration rate	Rooting depth	Moisture characteristic	Clay %	Sand %	Silt %	Stones %	Bulk density	Soil structure	Soil texture	C/N ratio	CEC	NH_4	NO_3	Organic carbon	P	Total N	Carbon fractions	Soil depth
▓	▓	▓	▓	▓	▓	▓	▓	▓	▓	▓	▓			▓	▓	▓	▓	▓	▓	▓		▓
▓	▓	▓	▓	▓	▓	▓	▓	▓	▓	▓	▓			▓	▓	▓	▓	▓	▓	▓		
▓	▓			▓	▓	▓	▓	▓	▓	▓	▓			▓	▓	▓	▓	▓	▓	▓		▓
▓	▓	▓	▓	▓	▓	▓	▓	▓	▓	▓	▓			▓	▓	▓	▓	▓		▓		▓

Note: the highly sensitive parameters are shaded

Overview of the Victorian Soil-Landscape Inference System (VSLIS) concept

Although the primary purpose of the VSLIS is to deliver soil–landscape parameters for modellers, it also delivers a conceptual framework to assist the management of the associated activities and processes. It will become the system that records and applies (over time where possible) the inference approaches that are deemed appropriate for the different Victorian landscapes (based on expert knowledge of the parameters being inferred, landscape and data limitations). This will provide the basis for managing the evolution of these inference systems and their associated products in a logical and ordered manner. The VSLIS will integrate with the existing DPI Modelling Information and Knowledge Environment (MIKE) (Williams *et al.* 2009b) and extend it from a passive register of modelling activity and model metadata to an active engine guiding modelling and inference activities.

The VSLIS in its simplest conceptualisation is a three component model with input, inference and output phases (Figure 1). In this form it also resembles a basic workflow. During the input phase there will be workflows associated with data packaging, delivery, and quality assurance and other "filtering" (ie spatial and temporal) and processing of data for use. The inference phase contains a complex of workflows involving analysis and modelling. These are equivalent in concept to the 'kepler' workflows described by Barsheghian *et al.* 2008. In the output phase workflows are involved in taking the output of the inference processes and assembling these into usable products. Where possible workflows will be automated enabling some products to 'live' (ie. evolve as new data or improved inference models are added). Initially the system will operate to provide a conceptual framework guiding projects and manual inference activities. This need will always be present as some forms of inference will always have a cognitive or tacit knowledge component and resist automation.
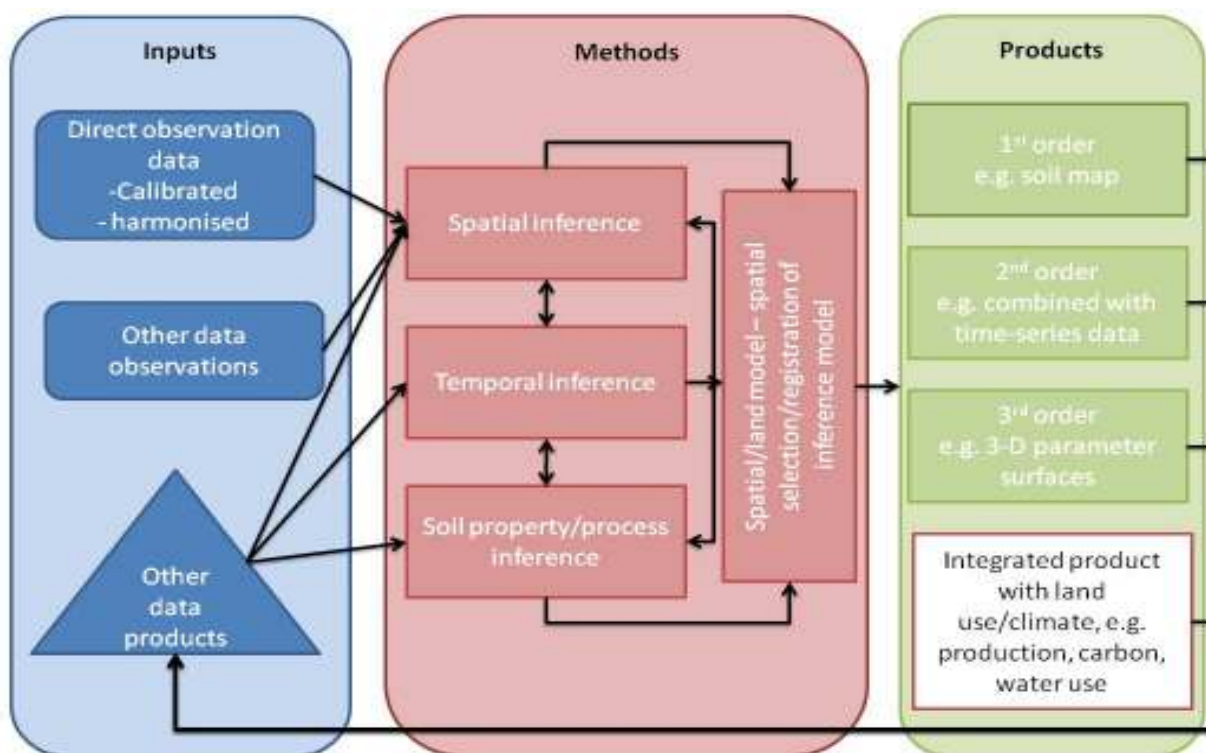
Figure 1. Proposed conceptual framework for the Victorian Soil-Landscape Inference System

Key framework elements
The key elements of the VSLIS are described below according to the component phases in the system.
Input Phase: The phase is characterised by a spectrum of services associated with input data and its metadata. These services support workflows associated with feeding data into inference processes and need to allow data filtering functions including spatial and temporal selection. Mature standards for some functionality will require further evolution and development to meet the VSLIS needs. In some cases outputs from inference processes will be fed back into the data store for use by other inference processes. It will take significant effort to create the requisite metadata to support workflows within this phase.
Inference Phase: VSLIS Inference is soil parameter centric with three domains recognised; (1) soil property, (2) temporal change and (3) soil spatial inference. Inference models and rules associated with these domains may vary in their level of workflow integration and in some cases may be quite discrete elements. Each model or rule will have individual spatial assignments to the Victorian landscape. The establishment of a systems environment to register and manage linkages between of both this spatial alignment and the available and evolving data inputs for inference elements is a significant requirement and challenge.
Output Phase: The VSLIS has three distinct orders of products based on product dimensionality (see figure 1). First order products have two or less dimensions (ie x,y or lat,long) and are represented by traditional graphs and maps. The addition of temporal or depth (z) to these products elevates it to the second order and including both progresses it to third order. Integrated products result when VSLIS products (of any order) are combined with other data and data products, often via analytical or modeling processes. Typical integrated products include land suitability and land impact mapping. Primary workflows in this phase consolidate the outputs from different inference approaches across the Victorian landscape for specific soil parameters.

Inference system selection design concept
VSLIS operation must allow multiple inference methods, even for the same soil parameter, to apply in different areas within the Victorian landscape. This requires that individual inference models be spatially assigned and constrained. This concept is used in modelling systems based on regular spatial grids such as the Catchment Assessment Toolkit (CAT) (Hocking *et al.* 2009) and Platform for Environmental Modelling Support (PEMS) (Chan *et al.* 2008) and those based on irregular spatial units such as the MIKE (Williams *et al.* 2009b). A workflow engine will be required where the inference process is fully automated. This will not only need to manage the workflows and processes associated with the inference step but also workflows to harmonise and integrate the various results. Open source tools, particularly for the latter purpose have been developed in Germany by the Humboldt project (HUMBOLDT 2009) and are freely available.

Current system status and implementation

Although many of the parts of the VSLIS have current stand-a-lone equivalents in the research community the system should be regarded as still in specification and design. Development is intended to be incremental and will take advantage of existing investments and initiatives where possible. That said a substantial part of the input phase is well under-way with the VSIS and MIKE systems intended as major component systems supporting the VSLIS. There are many efforts underway globally that are improving soil inference models and approaches. The interoperability and system linkages within the VSLIS have yet to be fully designed. On the output side most current products are first order, a few are second order with no known third order products.

Conclusions

Co-ordination of future developments in Victoria should be facilitated with support and engagement of the National Soil and Terrain Committee to also inform other states of Victorian progress, but also potential opportunities to collaborate with these states on areas of interest in development of a Digital Soil Map of Victoria. The approach embodied by the VSLIS should be regarded as essential in practically assisting the realisation of synergies between current soil related research and development activities within DPI. Creating the VSLIS will streamline and improve both the responsiveness and quality of the knowledge chain for users of soil and land information. This will benefit other research and associated policy formulation and natural resource management. To achieve this, the most significant future area of work for the VSLIS will be the design, development and implementation of the supporting standards, infrastructures and metadata for system workflows and interoperability. Additionally in support of users it is believed that although first order products can be produced and disseminated outside the system (ie paper or file based), in contrast the higher product orders will increasingly require interactive approaches and support within the system from visualization technologies of rising sophistication.

References

Barsheghian D, Altintas I, Jones MB (2008) Accessing and Using Sensor Data within the Kepler Scientific Workflow System. Environmental Information Management Conference, September 10-11, 2008, pp 26-32.(Albuquerque, New Mexico).

Chan TO, Beverley C, Ebert S, Garnett N, Lewis A, Pettit C, Thankappan M, Williams SR (2008) Platform for Environmental Modelling Support: A Grid Cell Data Infrastructure for Modellers. In 'Landscape Analysis and Visualisation: Earth and Environmental Science'. (Eds C Pettit, W Cartright, I Bishop, K Lowell, D Pullar, D Duncan) pp. 97-117. (Springer Link)

Dale MB, McBratney AB, Russell JS (1989) On the role of expert systems and numerical taxonomy in soil classification. *Journal of Soil Science* **40**, 223-234.

Hocking MJ, Beverley C (2009) Prioritising landscape investment using a catchment-based process model. 18th World IMACS / MODSIM Congress, Cairns, Australia, 13-17th July 2009, pp 4015-4021.

HUMBOLDT (2009) HUMBOLDT Towards the harmonisation of spatial information in Europe. Web: http://www.esdi-humboldt.eu/home.html

MacEwan RJ (2007) New Dimensions for Agricultural Landscapes, Final Report for ORL 1.1 (MIS 05188). Department of Primary Industries.

McBratney AB, Minasny B, Cattle SR, Vervoort RW (2002) From pedotransfer functions to soil inference systems. *Geoderma* **109**, 41-73.

OGC (2007) Open Geospatial Consortium Implementation Standard - Observations & Measurements, OGC Document 07-022 Version 1.0 (Eds S Cox)

Williams SR, Hunter DC, Robinson N (2009a) Victorian Soil Information System, Conference proceedings of Surveying and Spatial Sciences Institute Biennial International Conference, Adelaide, Aust. Oct 2009, pp

Williams SR, Pettit C, Hunter DC, Cherry D (2009b) A Prototype Metadata Tool for Land Use Change and Impact Models – A Case Study in Regional Victoria, Australia. 11th GSDI International Conference: SDI Convergence – Research, Emerging Trends and Critical Assessment, (Eds: Loenen, Besemer & Zevenbergen) June 2009, Delft, pp 137-150.