The Current and Future State of Australian Agricultural Data

Ross Darnell (pictured), Michael Robertson, Jaclyn Brown, Andrew Moore, Simon Barry, Rob Bramley, Michael Grundy and Andrew George

Commonwealth Scientific and Industrial Research Organisation

Digital technologies are currently underpinning revolutions in business and society, driving efficiency gains and opening up entirely new business models and opportunities. There is significant concern in the agricultural sectors that Australia could miss out on the benefits of digital technologies because of gaps and deficiencies in our data infrastructure or a lack of digital innovative thinking. In reviewing cross-sectoral data it has become apparent how haphazard the development of data and knowledge assets has been in some cases. While the value of information and knowledge about Australia has been recognised, there has not been a fully coordinated strategy around its prioritisation, collection and realisation of value. At the highest level, there needs to be an assessment of whether tools and platforms developed for other markets will be fit for purpose for Australian enterprises. The Accelerating Precision Agriculture to Decision Agriculture (P2D) project reviewed the current and future direction of the major data sets and analytical methodologies across different agricultural sectors and developed 13 recommendations that address the needs, decisions and priorities facing these industries. The P2D project was led by the Cotton Research and Development Corporation (RDC) and was jointly funded by the Department of Agriculture and Water Resources, the Rural R&D for Profit program and all 15 RDCs. Here, we present a summary of the P2D project report on whether digital systems developed for other agricultural markets will fit the purpose for Australian enterprises and some of the key recommendations.

Agricultural industries have long been innovators in the use of technology, including information and communications technology. In recent years, though, a confluence of factors has raised the prospect of digital agriculture: the use of digitally-captured information in the operations of farms as a matter of course. The cost of collecting data about farm resources (such as soils, plants, animals and equipment) is falling rapidly. Generic computing platforms and technologies such as the cloud and machine learning are becoming ubiquitous, even in bandwidth-constrained rural settings. Existing and new businesses are adapting to provide services that take advantage of these technologies. Nevertheless, there is still considerable debate about the potential business value of digital products and services in agriculture. It could be argued that many current offerings are speculative. In addition, the platforms and technologies are only now maturing so the industry has not yet consolidated.

Many of the systems in digital agriculture have been developed in the United States (US). This reflects the larger US market, the location of the major agribusinesses in that market and the much bigger pool of venture capital. Some of the major 'platforms' (a term we will return to) are in the process of establishing Australian operations and/or franchises. In a comprehensive report (Perrett et al. 2017), the Australian Farm Institute (AFI) identified a number of potential challenges to the existing US technologies being applied in Australia. In particular, they noted that the lack of publicly available soil and weather data meant that approaches developed for the US corn and soybean industries would not directly transfer to the Australian market. The US has publicly available detailed soil and weather information that Australia lacks.

The significant recent Accelerating Precision Agriculture to Decision Agriculture (P2D) project explored these issues, delving more deeply into how the provision of data in Australian agriculture can help to realise the potential benefits from the adoption of digital technologies. Here we summarise our main conclusions and recommendations; more detail can be found in the P2D report (Barry et al. 2017).

Major Cross-Sectoral Data Sets

Much of the data required to do digital agriculture will be collected on-the-fly on individual farms. There are however some information bases that are highly useful, which can be applied across industry sectors and for which data quality can be improved by taking a whole-of-landscape approach. These cross-sectoral data are where the policy settings can be most effective. We highlight some of the reasons why in the following sections.

Soils

Unlike the US and some European countries (where farm-scale soil maps have been produced), Australia has not had a long-term, detailed soil survey program. Some broad-scale and consistent mapping is available in some jurisdictions, and soil information (plus streams of soil data) is now available from farms and agribusinesses based on proximal and remote sensing technologies.

In 1992 the Australian Collaborative Land Evaluation Program (ACLEP) was established to develop a coordinated approach to land resource assessment across Australia. The program included all Commonwealth, state and territory agencies involved with land resource assessment. Despite many achievements, the collaborative model forged by ACLEP is no longer viable because of inadequate funding and the lack of a formal institutional mandate.

A recurring issue has been provision of the enabling infrastructure for collecting, curating and analysing soil information. The need for

Example 1: A decision support tool to predict yield from rainfall

Yield Prophet Lite (Lane 2017) is a decision support tool that estimates potential yield values for your crop given different rainfall amounts and fertiliser application rates. It also gives the rainfall likelihood for the remainder of the growing season.

The tool requires weather, soil water availability, soil nitrogen and carbon levels to estimate the likelihood of achieving yield level categories.

The tool is built on the French and Schultz rainfall yield model. It has been successful because of its ability to be understood and used quickly, present complex probabilities simply, and directly answer the question a farmer has about rain and nitrogen without needing to access other additional sources of information.

Weather forecast information is sourced from the Bureau of Meteorology's POAMA (Predictive Ocean Atmosphere Model for Australia) seasonal climate forecasting models, while soil information requires farmer values and soil carbon estimates drawn from the Soil and Landscape Grid of Australia.

new arrangements to achieve more open access to information has been recognised in a range of reviews and reports (eg Campbell 2008; NCST 2013; ITPS 2015; Keogh & Henry 2016; McKenzie et al. 2017). In addition, it is now clear that the digital revolution has created exciting new possibilities that overcome many past obstacles.

The National Committee on Soil and Terrain (NCST) and the Australian Soil Network (ASN) proposed a new soil-information system for Australia that supports the best features of the current system, takes advantage of new technologies and avoids the restrictions in the ACLEP model (NCST 2013). Central to these proposals are the establishment of the Australian National Soil Information Facility (ASIF).

Collection and use of soil information by private actors in agriculture has been an important part of business practice associated with the provision of services and advice, supply of inputs and in various forms of planning, reporting and assessment. These data have tended to be held as key intellectual property or to protect privacy and commercial interests. With appropriate policy and technical settings, it is envisaged that these rich data sources can be part of ASIF and therefore broaden the scope and timeliness of the public integrated system.

Changes in the structure of private actors in the agricultural advisory system, increased soil information capacity in agribusiness and increased capacity on-farm to collect and monitor soil status is now providing new private-sector opportunities in soil information supply and demand. In addition, the new web services provided by ASIF as envisaged here can also form the basis for new businesses in the knowledge economy or provide the context for new soil information services. For example, potential exists for locally-based soil data marketplaces; especially if the data streams available from farm machinery, soil sensors and appropriately interpreted proximal and remote sensing are included. Here we envisage local and intense soil information available on-farm or across farm communities being used in context with ASIF.

Example 2: Measuring and monitoring sugarcane yield

Jason owns a 100 hectare sugar farm in North Queensland's Herbert River District. Like all growers, he has no choice which mill his cane is delivered to, but by sharing data about his production system with the mill, he wants to ensure that he harvests his cane in such a way that sugar production at the mill is optimised, so as to maximise his return.

His cane is yield monitored during harvest as the harvester also has on-the-go Commercial Cane Sugar (CCS) sensing. Each bin has an RFID (radio frequency identification) tag which connects via a reader and Bluetooth to the harvester so that, once delivered to the mill, the provenance of each harvested bin and its production details are known. Both the harvester and the haulouts have machine guidance to minimise the risk of soil compaction.

The data collected are uploaded to a data cloud and on non-harvest days, Jason and his agronomist can download the maps of yield, CCS and tonnes of sugar produced and see how these compare with previous seasons, with the predicted yield and CCS and also with his soil and Digital Elevation Map (DEM).

These data are valuable inputs to his decision about further ratooning, moving to a fallow and thence replanting, and in conjunction with his high-resolution soil data, decisions on which varieties he should plant next and whether drainage or other land preparation works are warranted.

Weather and climate

Australia's climate is one of the most variable in the world and can be traced as a major influence on our interannual food productivity (Hochman et al. 2012). The best way to adapt to such a system is to be able to forecast it, from the next 10 minutes to the next decade, and to translate this to optimal decision making in the face of inherent uncertainty.

The challenge is that weather models tend to have skill at the daily time scale only up to 10 days; beyond this time, chaotic factors begin to dominate. Sub-seasonal forecasts must aggregate to longer probabilistic time scales (eg the likelihood of receiving more than 25 millimetres rainfall in January) or harness synoptic features that can persist beyond this threshold, such as the Madden Julian Oscillation in the tropics.

Climate forecasts at any time scale are provided on a global grid, where the value in that grid describes the average conditions throughout the grid box. The typical agricultural user is interested in only their location, not the averaged conditions. In the absence of farm or paddock-scale gridded products, some method of downscaling and calibration is therefore required.

Improvements to the skill of weather and climate forecasts can continue to enhance Australia's resilience in agriculture.

Remote sensing imagery

The primary uses of remote sensing (RS) imagery in agriculture have been in the detection and mapping of classes of land cover of interest, or the change in land cover responses over time, or a combination of the two. To date, this information has been at a coarse spatial grain.

New and increasing numbers of RS sensors and platforms are becoming available. Most important are those from the national space agencies (government) and the private sector, particularly the miniaturised satellites and sensors mounted on aeroplanes, drones or unmanned aerial vehicles (UAVs). This new generation of satellite sensors is beginning to provide both high spatial resolution and high repeat frequencies, making it feasible to detect changes in time at paddock and sub-paddock scales.

However, the trend towards increasing numbers of sensors and platforms with higher spatial, temporal and spectral resolutions will result in increasing data volumes. Someone, somewhere, will need to accommodate the flow, storage and processing of massive volumes of spatial data – and if it is to be useful for farm operations, then satellite information needs to be available within days rather than weeks.

The emergence of new data streams, particularly from proximal-sensing technologies (such as drones, UAVs and ground-based sensor networks) is expected to accelerate. It is easy to imagine a future where a property (or a collective of properties) has its own fleet of autonomous drones that launch several times a day to provide high-accuracy, high-resolution sensing of the full estate. These will combine with ground-based and animal-based sensor networks; RS imagery will then be just one part of a whole, integrated information ecosystem that provides seamless, real-time and management-relevant information products.

Big data analytics and data assimilation schemes will be the core of such an information ecosystem. These will also involve automated data analytics, artificial intelligence and fully integrated user-interface networks. A large challenge here is the difficulty faced in integrating data sets of different scales, type and quality.

Land use

Information on land use in Australia has been of interest since the early days of European settlement. Land use information allowed people to understand opportunities and plan businesses. It helped governments to develop regions and learn and respond to the novel Australian environment.

Land use information is potentially useful to industries and governments for a range of purposes. Knowledge of the location and extent of industries can support government development of infrastructure, better regional planning and policies, and allow commercial providers to target products and services. Land use information can also be used to support market access requests, plan and execute effective responses to incursions of pests and diseases, and plan effective surveillance strategies. Australia's National Committee on Land Use and Management Information provides coordination and standards to underpin the production of products.

There is significant demand for improvements to land use products to provide better resolution, accuracy, timeliness, and land use change. Improvements in accuracy will come from better quality information sources (eg from new RS products). Improved timeliness extends to the ability to make within-season assessments of land use. These would be valuable to predict yields and use this information to inform logistics and marketing.

The impediments to reaching this future state are both technical and institutional. The technical challenge is to find better ways of accessing and integrating information. Digital agriculture platforms, citizen science (via public and/or industry members), administrative information such as research and development levies, and improved RS products are all potential new data streams that can significantly improve the timeliness and quality of land use information in the future. The challenge is to integrate these sources efficiently.

The institutional issue is that while there are a wide range of potential beneficiaries that can get value from this data, the costs associated with developing and operationalising products are too large to be met by a single user. Development of new products requires cooperation between all potential beneficiaries. A more distributed business solution is needed.

What Needs to be Done?

Data that is of importance to the agriculture sector is often of interest more broadly; obvious examples are weather and RS information. The Rural Research and Development Corporations (RDCs) are the natural participants in the broader alliances that need to be formed in these cases. Where information meets cross-sectoral needs, but the benefit is mainly to agriculture, RDCs should commit to joint acquisition, provision and dissemination of the data sets. An example of a class of data on which RDCs should focus their investment effort is the acquisition of functionally relevant soils data, especially beyond the broadacre cropping zone.

The current ways of handling cross-sectoral data and assets reflect needs, decisions and priorities that have changed over time. The current arrangements may be appropriate for a future where the opportunities for predictive analytics in the agricultural sector are both greater and very different. We recommend development of a strategic plan around cross-sectoral data assets. The plan should be a living document that identifies the needs and the pathways to achieving these needs, such as public, public/private, and private investment.

To illustrate current and potential situations where data and analytics can be used for the benefit of the producer, we have included examples (on pages 42 and 43) of existing and potential innovations in the digital agriculture ecosystem that draw on assimilation of data to improve outcomes.

Analysis-ready data: standards and portal

While the existence of data and knowledge is necessary to facilitate digital agriculture, it is not sufficient. Substantial work needs to be done to develop analysis-ready data. Data sets that have to be wrangled into useful forms create a significant barrier that condemns the data to being accessible only to a small set of experts. An effort is needed to make the cross-sectoral, locally-collected data FAIR[1]: findable, accessible, interoperable and reusable.

In the P2D project, we located a variety of portals containing links to large numbers of data sets of varying quality. These collections help to make data findable, but they are a limited resource for new participants in the digital agriculture sector. There is a clear need to go beyond simple data portals that collate raw information, and to invest

1 https://www.force11.org/group/fairgroup/fairprinciples

in organising the data sets into agreed formats that are well-documented and 'analysis-ready'.

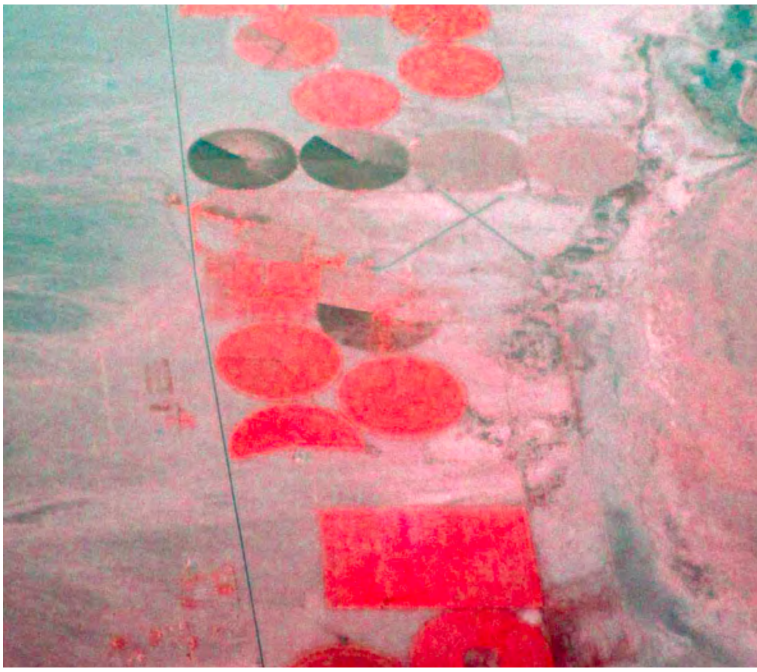
RDCs should collectively advocate for FAIR storage and dissemination of data sets that are valuable across the rural sector and that are also widely used in other industries. Classes of data where RDCs should actively advocate for secure and FAIR access – and should be joined by the rest of the industry – include: satellite imagery (especially via Geoscience Australia and particularly to ensure reliable access to the next generation of public-sector satellites such as Sentinel-1 and Sentinel-2); historical and forecast climate information; and improved monitoring of land use.

Everyone has a platform: data support tools for digital agriculture

The perception of commercial opportunities in digital agriculture has seen an explosion of platforms looking to gain a foothold in the market. Rather than starting from computer models and interfaces designed by agricultural scientists and targeted at particular decisions, these new platforms are based on ideas and models that have been successful in other digital industries.

Despite all being marketed as platforms, these new software tools are actually highly diverse, and they reflect different views of where the opportunities in the rural sector. At least four broad types of platform can be seen emerging in the North American (and to a lesser extent in the Australian) rural industries sector:

- **Aggregated views of information:** these tools are similar in purpose to traditional monitoring/diagnosis tools, but present a decision maker with multiple data streams (for example presenting current weather and forecasts, soil moisture and commodity prices side-by-side). These applications are analogous to the use of 'dashboards' to provide synthesised management information in government and industry. The weakness of these products is their inability to integrate information and the scales at which input data are available.
- **Mobile apps:** these are based on simple, easy-to-use interfaces and are targeted at very particular problems. They are often linked with other technology such as drone-mounted or in-field sensors. Examples include the NSW Drought Feed Calculator[2] and The Yield's app for irrigation in horticulture.[3] These tools exploit the ubiquity of smartphones and the well-developed ecosystem to market and deploy apps. The major impediment to using them in Australia is broadband coverage.



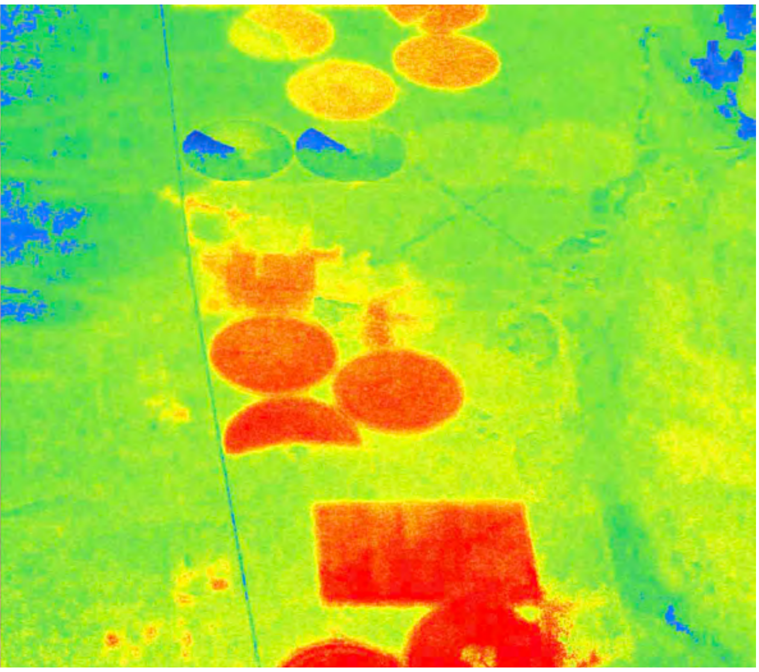
- **Federated analysis platforms:** these are based on gaining access to data from multiple enterprises and using it to learn to predict, or to benchmark, commercially important quantities such as prices of inputs, commodities, or yields. The resulting analytics can, in principle, be used for any of the purposes described above. These applications mimic the classic big data model where the flow of data permits continuous improvement of the analytics. In Europe and the US, their success is critically dependent on the availability of publicly curated soils and weather information at appropriate scales. Variants of such platforms can also provide

2 https://www.dpi.nsw.gov.au/animals-and-livestock/nutrition/costs-and-nutritive-value/feed-cost-calculator

3 https://www.theyield.com/products/sensing-plus-for-agriculture

privileged access to suppliers and markets; in these cases, the platform can mimic the Costco business model in which membership provides access to improved buying power.

- **'Pure' platforms:** these are platforms in the narrow sense; their purpose is to provide software infrastructure through which multiple third parties can transact business, exchange data and access digital and professional services. They typically include cloud-based storage, standard data formats and access control; access is on a subscription basis. Pure platforms are powerful tools and if successful can become dominant players. There are preliminary indications that major software companies are developing pure platforms for agriculture.



Dashboard and app platforms can largely be left to the market. The capacity to federate and analyse data, and the pure platforms that enable data exchange and transmission, on the other hand, should be receiving policy attention. The sharing of data between organisations and individuals is one of the key opportunities opened up by digital technologies, but it is also one of the most challenging. Shared data can be used in a variety of ways. The ability to gain additional value from private investments occurring in data collection should not be underestimated, but neither should the current lack of infrastructure and proven business models to support this.

A platform or platforms are needed for owners and users of agricultural data to exchange, market and value-add data for a variety of end purposes. Such platforms are a positional good: the benefits of common access will be large. An ambitious but worthy challenge would be for industries and policy managers to explore the feasibility of an industry-good pure platform, with a business model and appropriate protocols around use and rights of owners and users, that could catalyse data exchange and encourage competition and innovation rather than dominance by a single market actor.

Analysis-ready people: training for digital agriculture

The new opportunities being facilitated by digital technology rely on a range of skills. To participate fully, people need knowledge of digital technologies to understand opportunities. They need knowledge of agriculture to understand the true value proposition of information and services, and they need to understand the opportunities that digital technologies will provide for business-process innovation across the sector. While there is a pool of people with some or all of these areas of expertise, that pool is not yet large enough to harvest the coming opportunities.

As a matter of urgency, changes to university and vocational educational and training must be devised to ensure a supply of agricultural data scientists and data-ready farm workers. Evidence indicates that the Australian university system in particular is not producing sufficient agronomists with the required skills and that current incentives to change this situation are insufficient.

Conclusion

Australian agriculture is at an important juncture. Digital technologies are disrupting existing business models and transforming all aspects of the Australian and world economy. There is an

existing and successful agricultural technology and advisory sector in Australia but the pool of resources to develop new technology is limited. There are major investments in new digital agriculture platforms occurring internationally. The challenge is to ensure that the potential efficiency gains from digital agriculture are achieved.

That Australia could miss out on the benefits of digital technologies because of gaps and deficiencies in our data infrastructure is a significant concern. Whether Australia can mitigate the risks and grasp the opportunities depends on a number of issues. At the highest level there needs to be an assessment of whether tools and platforms developed for other markets will be fit for purpose for Australian enterprises. Australian farming systems have similarities and differences to those overseas. A key policy question is whether these differences will lead to significant barriers or delays to entry and whether this risk can be mitigated in some way.

Where existing platforms are fit for purpose, there is opportunity to leverage off international investments to access cutting-edge technology developed in other markets. Understanding what potential barriers exist is key to enabling this leverage. Are there gaps in our data and knowledge holdings and information infrastructure that make the Australian market less attractive for investment and development? Where products do not exist, there need to be opportunities – and no barriers – for Australian companies to develop appropriate technologies.

All industry sectors and RDCs will need to make these assessments taking into account the different business models, production systems and technology providers. However, there will also be opportunities for cross-sectoral investment and strategy around data. The P2D report (Barry et al.) considers these cross-sectoral issues and is a component of a broader project examining issues around privacy, ownership, architecture and availability of data in Australian agriculture.

References

Barry, S, Darnell, R, Grundy, M, Moore, A, Robertson, M, Brown, J, Gaire, R & George, A (2017), *Precision to decision – current and future state of agricultural data for digital agriculture in Australia,* CSIRO, Australia.

Campbell, A (2008), *Managing Australia's soils: a policy discussion paper,* prepared for the National Committee on Soil and Terrain (NCST) through the Natural Resource Management Ministerial Council (NRMMC), available at: http://www.clw.csiro.au/aclep/documents/Soil-Discussion-Paper.pdf

Hochman, Z, Gobbett, D, Holzworth, D, McClelland, T, van Rees, H, Marinoni, O, Navarro Garcia, J & Horan, H (2012), Quantifying yield gaps in rainfed cropping systems: a case study of wheat in Australia, *Field Crops Research,* vol. 136, pp. 85–96.

ITPS (2015), *Status of the world's soil resources (SWSR) – technical summary,* Food and Agriculture Organization of the United Nations and Intergovernmental Technical Panel on Soils, Rome, Italy, available at: http://www.fao.org/3/a-i5126e.pdf

Keogh, M & Henry, M (2016), *The implications of digital agriculture and big data for Australian agriculture,* Research Report, Australian Farm Institute, Sydney, Australia.

Lane, V (2017), Simplifying farm system modelling outputs via user experience methods, in Proceedings of the 18th Australian Society of Agronomy Conference, Ballarat, VIC, available at: http://www.agronomyaustraliaproceedings.org/

McKenzie, NJ, Hairsine, PB, Gregory, LJ, Austin, J, Baldock, J, Webb, M, Mewett, J, Cresswell, HP, Welti, N & Thomas, M (2017), *Priorities for improving soil condition across Australia's agricultural landscapes,* report prepared for the Australian Government Department of Agriculture and Water Resources, CSIRO Agriculture and Food, Canberra.

NCST (2013), *Establishing the Australian Soil Assessment Program,* a report prepared by the National Committee on Soil and Terrain for the Soil Research Development and Extension Reference Group, 22 March 2013, available at: https://soilstrategy.net.au/wp-content/uploads/Australian-Soil-Assessment-Program-FINAL-August-2013.pdf

Perrett, E, Heath, R, Laurie, A & Darragh, L (2017), *Precision to decision – analysis of the economic benefit and strategies for delivery of decision agriculture – draft final report,* Australian Farm Institute, Sydney, August.

About the Authors

Dr Ross Darnell is a Research Scientist and Team Leader with the CSIRO's Data61 Business Unit focusing on survey and experimental design and statistical modelling of environmental and agricultural systems.

Ross joined CSIRO in 2008, having extensive experience as a statistical consultant in research and academic institutions in Australia and the United Kingdom and was previous to that a biometrician for New South Wales Agriculture.

Dr Michael Robertson is the Deputy Director of CSIRO Agriculture and Food. He is based in Perth, Western Australia. His research background is in crop agronomy and simulation modelling and recently he has been overseeing CSIRO's efforts in digital agriculture.

Dr Jaclyn Brown is a climate scientist in CSIRO's Agriculture and Food Business Unit. She is passionate about finding innovative ways to link the last in climate research and forecasting with the needs and decisions in agriculture. She studies everything from the physics of ocean currents that drive climate modes to how farmers use apps to understand weather forecasts.

Dr Andrew Moore leads CSIRO's Digiscape Future Science Platform that is harnessing the digital revolution for Australian farmers and land managers. After training as an ecologist, Andrew joined CSIRO in 1989. Prior to Digiscape, he worked on building mathematical models of agro-ecosystems, especially in livestock production and mixed farming, and on applying the models via computer decision support tools.

Dr Simon Barry is a member of the CSIRO Data61 Executive Team and Research Director of Analytics Research Program. Simon joined CSIRO in 2007 and is currently leading the analytics research program that turns data and models into knowledge for informed decision making and action, in partnership with other scientists, government agencies, businesses, and industry.

Dr Rob Bramley is a Senior Principal Research Scientist based at CSIRO Waite Campus in Adelaide where he is also Site Leader. He has been active in precision agriculture R,D&E in the wine, sugar and grains sectors for over 20 years and is best known for pioneering the development of precision viticulture.

Mike Grundy is the Research Director – Soil and Landscapes in the Agriculture and Food Business unit of CSIRO. He has led major multi-disciplinary natural resource assessment activities for over 30 years. In recent years, he led the Agriculture and Land Use elements of the Australian National Outlook and the process that led to the Australian Soil and Landscape Grid.

Mike is a Fellow and Honorary Life Member of the Australian Institute of Agricultural Science and Technology and a Board Member of the International Soil Conservation Organisation. He serves on a range of international and national science/policy advisory groups.

Andrew George has over 15 years' experience as a statistician. His career has been focused on the development of new statistical methods for uncovering the genetic basis of heritable traits in animals, plants, and humans. Andrew has worked both abroad and in Australia. He is now a Principal Research Scientist at the CSIRO.