Transforming Scientific Research and Development in Precision Agriculture: The Case of Hyperspectral Sensing and Imaging

A thesis presented in partial fulfilment of the requirements for the

degree of

Doctor in Philosophy

in

Agriculture

At Massey University, Manawatū

New Zealand

Megan Zeta Cushnahan

2021

Abstract

There has been increasing social and academic debate in recent times surrounding the arrival of agricultural big data. Capturing and responding to real world variability is a defining objective of the rapidly evolving field of precision agriculture (PA). While data have been central to knowledge-making in the field since its inception in the 1980s, research has largely operated in a data-scarce environment, constrained by time-consuming and expensive data collection methods.

While there is a rich tradition of studying scientific practice within laboratories in other fields, PA researchers have rarely been the explicit focal point of detailed empirical studies, especially in the laboratory setting. The purpose of this thesis is to contribute to new knowledge of the influence of big data technologies through an ethnographic exploration of a working PA laboratory. The researcher spent over 30 months embedded as a participant observer of a small PA laboratory, where researchers work with nascent data rich remote sensing technologies.

To address the research question: "How do the characteristics of technological assemblages affect PA research and development?" the ethnographic case study systematically identifies and responds to the challenges and opportunities faced by the science team as they adapt their scientific processes and resources to refine value from a new data ecosystem. The study describes the ontological characteristics of airborne hyperspectral sensing and imaging data employed by PA researchers. Observations of the researchers at work lead to a previously undescribed shift in the science process, where effort moves from the planning and performance of the data collection stage to the data processing and analysis stage.

The thesis develops an argument that changing data characteristics are central to this shift in the scientific method researchers are employing to refine knowledge and value from research projects. Importantly, the study reveals that while researchers are working in a rapidly changing environment, there is little reflection on the implications of these changes on the practice of science-making. The study also identifies a disjunction to how science is done in the field, and what is reported. We discover that the practices that provide disciplinary ways of doing science are not established in this field and moments to learn are siloed because of commercial constraints the commercial structures imposed in this case study of contemporary PA research.

Acknowledgements

The first acknowledgment must go to my supervisory group, comprising Associate Professor Dr Brennon Wood and Associate Professor Dr Mathew Henry for their consistent support, guidance, and patience during the running of this project. They continually pushed me to deepen my reasoning and clarify my arguments, and I'm very satisfied with the outcome of their efforts. They taught me much about academic research in the humanities and fostered many enjoyable conversations not only about the particular project we worked on together but about academe in general. Furthermore, they were valued colleagues throughout my studies, inviting me to read and discuss matters in our field. These experiences not only were personally enjoyable and intellectually stimulating but they also significantly benefited the research reported in this thesis.

I give my thanks to faculty staff and my fellow students in the Massey University School of Agriculture, especially Dr Miles Grafton, Dr Doug Drysdale, the Social Sciences PhD Group, and my companions in the 'PhD Morgue' for their unwavering support, guidance, and kindness. I gratefully acknowledge the *Pioneering to Precision* research team, who allowed me inside of their world for 30 months and to C-DAX for partly funding my stipend.

Many friends outside Massey University provided support and camaraderie. First, I would like to say a special thank you to my dear friends Matt and Marie Henry, whose care, wine, child-minding, and encouragement were vital to the completion of this project. Thanks also to my colleagues at Transpower New Zealand for their support, laughter, patience, silliness, encouragement, hugs, cards, and of course chocolate.

Finally, as always, many thanks are due to my family. Undertaking a PhD as a working mother of two young children was never going to be easy, and without my family's baby-sitting, cooked meals, hugs and reassurance, this thesis would never have made it to fruition. In particular, my parents, Alice and Ross Hood, encouraged me at every stage as I pursued my goal of earning a doctoral degree and their guidance during my formative years was what gave me the ability to do so in the first place. I am grateful every day for your unconditional love, patience, support, and astute ears when the going got tough.

My extended family are to be similarly thanked, especially my in-laws Pat and Marie Cushnahan, my brother Neil Hood, and my late grandfather, Neil Candy, who was a crucial influence through his encouragement of his granddaughters to follow an academic path.

Thank you to my beautiful children Joseph and Margaret, who are the pride and joy of my life. I love you more than anything and I appreciate all your patience and support during Mummy's PhD studies.

Finally, this work is dedicated to my husband Tommy, who devotedly accompanied me on my PhD journey as he finished his own. The past six years have been hard in many ways, but they have also been the best years of my life, for which he is decisively responsible.

Table of Contents

List of figures

List of tables

List of abbreviations and acronyms

HSI	Hyperspectral sensing and imaging
IP	Intellectual Property
P2P	Pioneering to Precision (project)
PA	Precision Agriculture
PGP	Primary Growth Partnership
WSNs	Wireless Sensor Networks

Glossary

Māori words and phrases used in this thesis are italicised. Oxford Dictionary of New Zealand English has been used to indicate those Māori words now regarded as part of New Zealand English and are therefore not italicised.

Some of the following glossary definitions are used in full, or adapted from Natural Resources Canada (2015):

Actant: An actant is a human or non-human involved in an activity under study (MacLeod, Cameron, Ajjawi, Kits, & Tummons, 2019).

Agency: Agency is the ability to act and/or exert power which is distributed across networks of people and things (MacLeod et al., 2019).

Assemblage: An assemblage is a complex tangle of natural, technological, human and non-human elements that come together to accomplish both intended and unintended outcomes (MacLeod et al., 2019). In this thesis, assemblages relate to both assemblages of technologies (communities of machines) and assemblages of data from multiple sources.

Classification: When image pixels are the same colour, or nearly the same colour, an image "classification" computer program can recognize this and group such pixels together. Such a grouping is called a "class" and the process of doing the grouping is called "classification".

Data mutability: In data science, a mutable object is one whose state can be modified after it is created. In terms of the P2P this refers to the nature of aerially acquired reflectance data, which undergoes significant processing before analysis. Because the processing operations are black-boxed (especially the early processes performed within the device itself), it may be impossible to return to the original state of the data unless you have access to the original dataset. This has implications for data commensurability, especially between calibrations of the device. An immutable object is an object whose state cannot be modified after it is created and in the context of this thesis refers to all ground-reference data.

Digital number: In remote sensing systems, a digital number (DN) is a variable assigned to a pixel, usually in the form of a binary integer in the range of 0–255, i.e. a byte.

Electromagnetic spectrum: The range of energy which contains parts or "bands" such as the visible, infrared, ultraviolet, microwave (radar), gamma ray, x-ray, radio, and which travels at the speed of light. Different parts of the electromagnetic spectrum have different wavelengths and frequencies.

Emergence: The term emergence suggests that reality is less stable and predictable than we typically acknowledge. In this view, teaching and learning consist of both intended and unintended, predictable, and unpredictable, elements (MacLeod et al., 2019). It suggests that knowledge is always unfolding, surfacing moment-to-moment through a series of complex negotiations between an ever-evolving assemblage of actors.

Exhaust data: In reference to data sources, Kitchin and Lauriault (2018) separate captured data (deliberately sources) and exhaust data (inherently produced as a by-product of another process).

Ground reference data: Accurate measurements or observations of some property on the ground which can be used as a label or description of what a potential overhead image represents.

Ground truthing: Remote sensing analysts must be sure that their image analysis is accurate. This is done by field where they go out to the actual places shown in the images and confirm that what they think they see on the image is true. This term was often incorrectly used by the P2P researchers to refer to ground reference data, which were actually being collected to inform the algorithms for the project.

Image: The picture that is a result of the sensing process. A remote sensing image can be displayed on a computer monitor or it can be made into a printed copy.

Mosaic: A large image made by combining smaller images. For example, to get an image of a whole farm using the AisaFENIX, we combine many images. This is tricky because the images are taken at slightly different times and they could look different in colour or brightness between each pass of the aircraft.

Pedalogy: The branch of soil science that integrates and quantifies the formation, morphology, and classification of soils as natural landscape bodies (Sposito & Reginato, 1992).

Pixel: The smallest unit in a digital image. A satellite image is made up of a matrix of many pixels, each having its own digital value.

Platform: This is what carries a sensor - usually a satellite or an airplane. A remote sensing platform could also be a hot-air balloon, a tall tower, etc. (related words: satellite, aircraft, sensor)

Proximal: Samples taken from very near their target.

Reflection: Reflection occurs when radiation (light, radar signals, etc.) bounces off a target. It is very important in remote sensing how that reflection happens, how much is reflected and how the radiation is changed in the process of reflection, because it tells us much about the target that caused the reflection.

Remote sensing: Remote sensing is the action of collecting images or other forms of data about the surface of the Earth, from measurements made at some distance above the Earth, processing these data and analysing them.

Resolution: Spatial resolution describes how clearly you can see detail in a picture. Consider the focusing done by a camera. If the picture is blurry and you can't see small objects, the resolution is poor (low resolution). If the picture is sharp and you can see small objects, the resolution is good (high resolution). Resolution is also used in describing colour detail (how similar colours are) and even time detail (how close in time things happen).

Satellite: A satellite is a natural or human made object continuously orbiting above the Earth or another planet or star. A remote sensing satellite carries one or more instruments for recording images of the Earth, which are transmitted to a receiving station using radio waves.

Scanner: While a camera would take a picture of an area all at once, a scanner is a device that examines an area point by point until the entire area has been imaged. These points become the pixels in a digital remote sensing image.

Sensor: A sensor is the device that records a remote sensing image, much like a camera.

Target: Targets are the features being studied in a remote sensing image.

Validation: All remote sensing products must be validated because uncertainty is produced in every data processing procedure. Validation involves not only quantifying uncertainty, but also providing thresholds to determine whether the product is reliable.

"If you want people to have some grasp of science, you must show how it is produced." — **Bruno Latour**

Chapter 1: Introduction

In the early-1980s, a new field of agricultural science emerged called *site-specific farming*. In response to growing interest in this new approach to agricultural research, a small group of American agricultural researchers and statisticians came together at a workshop in 1990 with the goal of leveraging off recent advances in geolocation technologies to help farmers account for the variability in their crop yield and soils (M. A. Oliver, Bishop, & Marchant, 2013). Few researchers at that conference could have predicted the trajectory of the field, now known as precision agriculture (PA).

The arrival of PA brought with it a new wave of researchers from diverse fields such as engineering, mathematics, remote sensing and computer science who aimed to substantially improve the agronomic (Ortiz, Balkcom, Duzy, Van Santen, & Hartzog, 2013), environmental (Brown, Dillon, Schieffer, & Shockley, 2015; Schieffer & Dillon, 2015) and economic (Griffin, Lambert, & Lowenberg-DeBoer, 2005; Schieffer & Dillon, 2015; Schimmelpfennig & Ebel, 2016) performance of agricultural enterprises (Mulla, 2013; N. Zhang, Wang, & Wang, 2002). Research in the field flourished, promising to address some of society's 'wicked problems' such as food security challenges, soil degradation, water pollution and more recently climate change mitigation (Talebpour, Türker, & Yegül, 2015).

The work of many researchers was focussed on improving, or even replacing farmer decision-making and a legion of domain experts was engaged to help inform new systems. This interdisciplinary approach fitted well with contemporary theories hypothesising that different interests and perspectives are required to respond to the complex problems facing modern society (Alrøe & Noe, 2014). While much progress has been made, many of the gains thought possible from widespread adoption of PA remain elusive (Mintert, Widmar, Langemeier, Boehlje, & Erickson, 2016). Recently, many firms and researchers have entered the PA arena with the goal of applying "big data" techniques in an attempt to improve returns derived from PA techniques, technologies and practices (Ryan, 2019; Sonka & Cheng, 2015b; Wolfert, Ge, Verdouw, & Bogaardt, 2017).

There is much scholarly debate about the significance of the arrival of big data for the established PA research community (Carbonell, 2016; Carolan, 2020; Tremblay, 2017; Wolfert et al., 2017). Like many other scientific fields, the arrival of big data and advances in data analytics is driving the PA research community to re-examine its method of scientific research (Bell, Hey, & Szalay, 2009; Hey, Tansley, & Tolle, 2009; Kitchin & McCardle, 2014; Tolle, Tansley, & Hey, 2011).

While there are many excellent studies hypothesising on the future of PA for farming, such as Schrijver (2016) and Mulla (2013), most focus on the implications for farmers, with little or no discussion on the consequences for science-making. Importantly, these studies do not recognise the critical changes in the science-making practices and processes that are required to analyse and refine value from complex data assemblages. This research gap gives rise to key questions that this thesis seeks to address.

What is precision agriculture?

Precision agriculture first emerged as a specialist field of agricultural science in the early 1980s, although the term 'site-specific agriculture' was initially used to describe the concept. In 1990 the term 'precision agriculture' appeared as the title of a workshop held in Great Falls, Montana, USA (M. A. Oliver et al., 2013).

While there are many definitions and concepts pertaining to PA found in literature (Bramley, 2009; Khosla, 2010; McBratney, Whelan, Ancev, & Bouma, 2005; M. A. Oliver et al., 2013; United States House of Representatives, 1997), there is no nomothetic definition that fully reflects the field's technologies, processes and practices. For example, one frequently cited definition from the National Research Council of the USA states "Precision agriculture is a management strategy that uses information technologies to bring data from multiple sources to bear on decisions associated with crop production" (Dixon & McCann, 1997, p. 17).

Another early definition by Robert and Stafford (1999) represented PA using the three "Rs"; Right time, Right amount and Right place. Later, Khosla (2010) added two further criteria; "Right Source" and "Right manner", to narrow the definition. In 2019, the International Society for Precision Agriculture released their official definition for PA, which reflects the importance of data to precision agriculture:

> Precision Agriculture is a management strategy that gathers, processes and analyzes temporal, spatial and individual data and combines it with other information to support management decisions according to estimated variability for improved resource use efficiency, productivity, quality, profitability and sustainability of agricultural production (International Society of Precision Agriculture, 2019, p. 1).

Such a generic definition does little to distinguish the field from many other contemporary fields of enquiry. Indeed, PA is a field of research that is not exceptional in the sense of being easily identifiable and one cannot give an *a priori* definition. However, it does have some characteristics, principles and themes under which academics publish. So, while I am unable to arrive at a Napoleonic definition of PA, what I have found is that a field of practice exists that is characterised by a familiar genre of publication. It is without apology that I resist the temptation to rationalise this vagueness as a problem that needs to be solved before I can explore the processes and practices of the field. Consequently, in Chapters 2 and 3 I attempt to build an understanding of how research in the field operates by examining the science published under the banner of PA, rather than taking the traditional approach of presenting a literature review. While a literature review on an object of enquiry that can be established beforehand is a useful instrument, in this case I first had to identify the object of enquiry itself. Hence, I break with convention of presenting a literature review *on* PA, and instead I provide a review *of* the field of precision agriculture in Chapters 2 and 3.

Remote sensing in precision agriculture

This thesis focuses on a group of PA researchers employed on a project called *Pioneering to Precision* (P2P).[1] The goal of the P2P project was to improve fertiliser practice on hill country farms through remote sensing of the nutrient status of the farms and precision application of fertiliser (Ministry for Primary Industries, 2014). The steep terrain of the hill country landscapes mean it is impractical to accurately gauge pasture nutrient levels using established ground-based data collection methods. To overcome this problem, the P2P researchers worked with aerially acquired data from remote sensing technologies called hyperspectral sensors. These sensors collect data without the need to make contact or take physical samples directly (J. B. Campbell & Wynne, 2011). They collect light reflected from objects as a collection of images taken simultaneously of the same place. These collections of images are called bands (also called channels or layers), which are commonly described by the name (Red or Near-IR for example) and the wavelength of the energy being recorded (Montgomery, 2019).

Remote sensing applications require a mechanism for data capture. For example, human vision detects three wavelengths or spectral bands: red, green, and blue. Our brains combine the data detected by our eyes into a single colour image. All the colours that humans perceive are created by mixing the three additive primary colours: red, green, and blue (Montgomery, 2019). Similarly, standard photographic cameras collect light in red, green, and blue bands, which are combined to create a colour image (T. A. Cushnahan, Yule, Grafton, Pullanagari, & White, 2017). Importantly, like human vision, most sensors used in PA research only reflect light from the upper surfaces of the objects they scan. This means they cannot see through objects. For example, unless the soil surface is free of vegetation, the sensor will only provide reflectance data on the vegetation layer.

---

[1] I use the past tense when referring to the *Pioneering to Precision* project which was originally due to conclude in June 2020. The project was extended for six months due to Covid-19 disruption, so unless a further extension to granted, the project will end around the time this thesis is submitted for examination (White, 2020).

The two classes of sensor are multispectral and hyperspectral. They simultaneously measure data in multiple regions of the electromagnetic spectrum. Plant characteristics have been measured by PA researchers using multispectral sensors on ground and satellite-based platforms for many years (G. Anderson & Yang, 1996). Multispectral sensors usually have 5-10 bands which collect light from the visible and beyond into the Near Infrared (NIR), in discrete portions of the spectrum. In contrast, the hyperspectral sensors used by the P2P team have hundreds of contiguous narrow bands stretching from the visible (VIS) and NIR into the Short-Wave Infrared (SWIR) (Mariotto, Thenkabail, Huete, Slonecker, & Platonov, 2013). Hyperspectral sensors analyse a wide spectrum of light rather than just assigning primary colours (red, green, blue) to each pixel as a human eye would. The light striking each pixel is broken down into many different spectral bands to give more information on what is imaged. These tools lie at the heart of the work of the P2P researchers, who are the focus of this study.

Research aims and research problems

The primary aim of this research is to respond to limitations in our understanding of who contemporary PA researchers are, how assemblages of PA technologies are changing, and how the arrival of data-rich technological assemblages may impact the practices and processes of science-making. Initially, this study was intended to focus on how farmer uptake of technologies such as the hyperspectral sensing and imaging could be improved. However, in the first few months of my research, a theme emerged from the early data analyses indicating that the P2P researchers were struggling to follow the linear science process prescribed in the project set-up. It became apparent that the data generated by the project were unfamiliar and challenging to the group. The flexibility of the grounded theory approach allowed me to adapt my line of inquiry to ask reflective, more relevant questions informed by real-time observations, rather than strictly following a predetermined research path. It was decided that this emerging theme deserved deeper analyses, moving the focus of the study to the researchers themselves and the data characteristics of the nascent technologies they employ.

Hence, to theorise how the arrival of these technologies may transform PA research, I changed the primary research question posed in this thesis to: "How do the characteristics of technological assemblages affect precision agriculture research and development?" To address this research question, I begin by exploring the following questions:

1. What are the characteristics of new precision agriculture technologies?
2. How do information-rich technologies promise something different for precision agriculture?
3. What challenges, risks and opportunities emerge from the assemblage of data rich technologies in precision agriculture?

4. How do big data influence science-making in precision agriculture?

To provide a robust response to these research questions I initially performed an extensive bibliometric analysis of PA scientific publications, including specific analysis of research relating to the use of hyperspectral sensors in PA research. Early analyses revealed that while there is a growing body of scholarly publications in the field of PA, there is little reflection on science-making in the PA publications. This leads to a research gap in understanding the inner workings of contemporary PA laboratories, including the methods and approach of science-making, which make it difficult to build an understanding of how data from nascent technologies, such as aerially deployed hyperspectral sensors may impact the field.

To inform my research, I embarked on an ethnographic study of the P2P researchers, who were working at a contemporary precision agriculture research laboratory. My goal was to articulate a representation of the social reality of a science team and their science-making. For thirty months I walked the thin line between ethnographer and participant following the working researchers into their laboratory, and beyond to the field. The ethnographic study explores the implications of the arrival of big data technological assemblages and new data analytics for PA research and those engaged in it. By studying the doing of science, I hoped to see how science-making practices contribute to the possibilities of knowledge generation in PA that have profound implications for all the actors who have pinned their hopes on the field’s promises.

Recognising that ethics has become a cornerstone for conducting effective and meaningful research (Drew, Hardman, & Hosp, 2007), ethical questions were considered throughout the study. Indeed, while the ethical considerations I outlined in Chapter 4 may initially appear to be roadblocks, in practice they proved integral to the study. Points where the study pivoted to accommodate ethical considerations ended up providing some of the most valuable lines of enquiry in the thesis, such as the decision to focus on non-human actants in the latter stages of the research. This led the thesis to focus on the influence of data in science-making practices and processes and to consider what a radical shift in the data economy could mean for PA research. In this journey what emerged is that the non-linear, iterative scientific method adopted by the researchers is a far cry from the linear classical science traditionally promoted for agricultural science. The thesis explores if the arrival of big data really can help PA researchers realise the promise of the field in addressing some of society’s wicked problems, or if it is the last throes of a utopian attempt to solve problems from a second order perspective.[2]

---

[2] In this thesis I apply the interpretation of first and second order science as applied to agricultural technology development by (Alrøe & Noe, 2014). Alrøe and Noe applied the concept of first and second order perspectives developed by (Marton, 1981) which makes a fundamental distinction between two perspectives, which Marton describes “From the first-order perspective we aim at describing various

Significance and rationale of the study

Investment in the development of PA technologies is rapidly growing and the sector is attracting vast investment in new PA and related technologies. For example, a recent research report on the PA market estimates global sales of US$9.56 billion in 2019 (Market Study Report, 2020). Despite the impressive scale of the field, there has been little contribution from the humanities and social sciences in PA research. PA researchers widely promote the field's promise in helping address some of farming and society's key challenges such as climate change (N. Rao, 2017), water quality and improved food security (Demirbaş, 2018) to promote investment in research propositions (Mogues, 2015). Yet, progress has been slow, and the alarm bells are starting to ring in agricultural policy circles. Indeed, evidence suggests that while PA has succeeded in addressing some of the field's first-order problems (Franzen & Mulla, 2015), to date PA has largely failed to deliver on early promises to improve second order problems such as food security (Schrijver, 2016), climate change mitigation (Alrøe & Noe, 2014) and water quality (Dodd, McDowell, & Quinn, 2016).

Much of the sociological research on technological innovation focusses on why farmers don't adopt technologies or on ways to entice them to utilise technologies. While there are many empirical studies on the changing nature of innovative organisations in other sectors (N. Anderson, De Dreu, & Nijstad, 2004), the emphasis of PA research remains on technical innovations themselves rather than the researchers behind them (Sunding & Zilberman, 2001). The focus on technologies and their assemblages is leading to a lack of visibility of the impacts these technologies have within science-making organisations.

An important recent study performed for the European Parliament by Kritikos (2017) warns that:

> … the increasing use of data in agriculture and the gradual introduction of precision agriculture in European farming in combination with the lack of human resources raise a variety of socio-ethical challenges that resemble those that emerge on multiple occasions when technology is introduced in economic activities where the human element is more than vital (p. 39).

Kritikos (2017) outlines numerous socio-ethical risks associated with big data technologies in PA: monoculture, augmentation of the digital divide, possible data concentration and manipulation, including farmers' dependence on external inputs delivered from high-tech providers and the subsequent lock-in effects, threats against the sustainability of small, local farms, genetic erosion, control and unfair practices (Kritikos, 2017). This work is supported by Eastwood, Klerkx, Ayre, and Rue (2019), who observed that PA innovation activities have focused on technology

---

aspects of the world and from the second-order perspective (for which a case is made in this paper) we aim at describing people's experience of various aspects of the world" (Marton, 1981, p. 177).

development and on-farm use without considering socio-ethical implications and have excluded certain actors such as citizens and consumers. Indeed, historical analysis supports the notion that significant gains in agricultural output have come at a high social and environmental cost. Disappointingly little progress has been made in solving the social and economic problems of smallholder farmers, who have generally benefited the least from this boost in production (Clunies-Ross & Hildyard, 2013; Kritikos, 2017).

Motivation of the researcher

My motivation to improve farmer livelihoods and promote environmentally sustainable agriculture stems from being raised on a third-generation dairy farm during the 1980s. Growing up, I was acutely aware of the huge responsibility my parents felt to leave the land to my generation in a better state than they received it. Like many family farms, financial and agronomic sustainability are important drivers for the operation as the land passes between generations. In a departure from the norm in the 1980s, my parents invested significant time, money, and effort in making our farming enterprise environmentally sustainable. Many weekends were spent planting fence lines and waterways with trees and shrubs my mother had grown from cuttings and seed. My parents formed weirs to create wetland systems in our valleys with a view to enhancing biodiversity and water quality. They taught us that our family relies on the farm not only for our income, but for our wellbeing. After all, we drink the water from our aquifers and eat the produce from our farm. Like my parents, my generation are responsible for handing it to the next generation in a better state than we received it.

My parents based their purchase and production decisions on the best information to hand. Meticulous records of weather (daily) and detailed notes on the production and interventions of each field were maintained. In the early 1980s they experimented with new crops, researched production systems from overseas and sought advice from domain experts. Where budgets allowed, they also invested in early PA technologies. One early experience of PA technology resulted in my father beating an early model of a C-DAX precision sprayer to ‘death’ with a piece of acanthine pipe, such was his utter frustration at its poor performance. Despite his best efforts to render the machine usable, the sprayer was eventually consigned to the bull paddock, which doubled as a cemetery for failed machinery. Dad reverted to using his old low-tech sprayer, disappointed that the promise of this new tech had failed to deliver.

When it came time to choose a career, my school’s careers advisor said I was ‘too smart’ for a career in agriculture. Hence my career took a segue into turfgrass agronomy, completing an Honours degree in Plant Science (Turfgrass) in the 1990s. Soils science, agronomy, plant pathology and quality management became my bread and butter for the next fifteen years. I travelled the world in the service of the urban elite as a golf and turf consultant and had extensive experience in research extension and adult education. Then, following the sale of the research institute that I worked for, I was made redundant. What to do? Serendipitously, an opportunity to study PA arose at Massey University and I was kindly encouraged to apply by the retired head of the school, who was also a past board member for the research institute.

Assumptions, social position, and preconceptions

Even when the formality of data collection is over, ethnographic researchers must look back over their fieldnotes, interview transcripts, memories, and other artefacts to start making sense of the work and seek out themes, discourses and theoretical insights (Gill, 2008). Therefore, it is useful to explore conditions and assumptions that influence the lens through which an ethnographer such as myself sees the world. Unlike many doctoral scholars, academia does not dominate my background. Indeed, the end goal of this research is not even to have an academic future. Rather, this thesis represents a deeply personal geographic return to my home in the Horowhenua, New Zealand and to my agricultural roots.

Having grown up on a family farm, I have an inherent affiliation with and relatability to farmers that is difficult for a 'townie' to emulate. Living on a working farm also provided me with a social and analytical position unique within the research group. The *Pioneering to Precision* (P2P) project team at the centre of this study were based within Massey University, not far from my home. As an alumnus of the school, I was not entirely in exile when entering the laboratory. I understood the workings of the University and had friendships with many of the departmental staff prior to commencing the study. However, as the team was not established until well after my graduation, other than my husband, I only had a passing acquaintance with the other participants prior to undertaking my study. Importantly, I initiated my research as "Tommy's wife", rather than entering the field as an anonymous researcher. In practice this meant that I would never be outside the gender hierarchies in the community and I would be judged accordingly.[3]

When entering the realm of the P2P laboratory I envisaged my technical expertise would be of value to the research project and could perhaps help me interpret more of what was going on within the project than someone with less experience in the biological sciences. On reflection, this work was also influenced by my commercial experiences in that I had some benchmark of what research looks like outside of the University setting. Admittedly, my pre-conceptions of research organisations and their innovation processes probably led me to ask some questions the researchers did not appreciate. Research and development within every successful organisation I had worked for was in some way user led. These organisations had other commonalities that

---

[3] While Tommy is my husband's real name, like all participants, he is represented by a pseudonym in the thesis.

formed my opinions on what successful research looks like too. For instance, organisations having significant commercial success had many ‘research champions’ in the field, including their own research extension staff, who engaged in the entire research process, from generating new ideas, to vetting them, to testing them, to using them or recommending them to other end users. Having worked for industry-funded research groups, most of my colleagues took an ‘industry benefit’ view to both business and apprenticeship. As a graduate, I myself had been taken under the wing of renowned researchers, who shared their time and knowledge with me generously to ensure I would thrive in the turf industry.

In hindsight, it was with naivety that I assumed the contemporary university research environment would be the same. As I grew to know the researchers in the research group, I observed that my motivations for being involved in the project were very different from the early career academics. First, for me the project promised to improve the livelihoods of New Zealand farmers, a passion borne from seeing the tragic fallout of the deregulation of the country’s agriculture industry on farming families and rural communities in the 1980s. The project also provided an opportunity to follow world-leading research in the plant sciences, a lifelong interest of mine. However, these were not the primary drivers for most of the researchers that surrounded me. Most had joined the project from other fields, such as engineering, remote sensing, volcanology, and data analytics. For some, the project provided a much-welcomed regular income after thrifty years as doctoral students. For others, the project was an opportunity to advance their academic careers by working on a high-profile project, although interestingly they actively avoided publishing in the PA space as they could acquire higher academic rankings by publishing in other fields such as volcanology.

Working with others who have such different perspectives and priorities to mine emphasised the reality that my experiences, my perspectives, social position (Latour & Woolgar, 1986) and background assumptions (Hammersley & Atkinson, 2007) all influence the research approach for this study. My personal life shapes the how and why of this research. My circumstances as a working mother, feminist, a New Zealander of heterogeneous European and Scandinavian ancestry and the wife of a doctoral student all affected my perspectives and the assumptions I brought to the work of understanding the science-making in the P2P project. These factors all influenced the choices I made during this research study. Consequently, this thesis, like all ethnographic works, reflects my personal perspectives.

Limitations of the study

The present study includes an ethnographic case study of science-making in a PA research laboratory in New Zealand; as such, the scope of the study is delimited in several ways as discussed below. Only one research team was examined in the study to make the research manageable so that the influence of the arrival of big data on science-making can be studied in considerable depth.

The scope of the data gathered in the study was limited by the physical ability and time available for me to be present at the laboratory, to participate in laboratory activities (including their data collection events), and to collect data.

The laboratory studied in this research was selected in a purposeful manner. The laboratory was employing nascent PA research tools, which may not be employed in typical present-day agricultural research situations. For this reason, the findings of this study cannot be generalised with absolute certainty to any wider population of PA laboratories elsewhere.

This study does not serve the purpose of proving or disproving pre-existing theory. To borrow an analogy by Isabel Wilkerson, author of *Caste* (Winfrey, 2020), I initially approached the case study as a building inspector. In other words, I approached it like I was looking at the PA research system as though it is an old house and my job was to illuminate those spaces and areas and the parts of the structure that are problematic and that will not get better on their own. This study is interpretative and generative in order to produce theoretical ideas about the arrival of big data in PA research laboratories and its impact on science-making and its structures in the contemporary context. One thing that the building inspector does not generally do is make the repairs, hence this thesis does not present itself as having all the answers. However, as the study progressed, I did feel that the thesis had evolved to a point where the findings may help PA researchers understand themselves, their data, and the science-making structures they work within. In time, this may help the PA research community to better adapt to the changing data environment of precision agriculture research.

Thesis overview

This thesis follows an unconventional thesis structure borne out of initial phases of the study, which focused on defining the object of enquiry. When I deliberated the structure of the thesis, I regarded the early chapters as an important prelude to the ethnography. If I had changed the structure of the thesis to a conventional format it would have looked unified and smooth – the current structure captures the wrinkly, explorative work that was performed. Hence, the thesis is organised as follows.

Chapter 1: Introduction

In the introductory chapter I set the scene by providing a background of the reasons for the research and define the nature of the study. I identify the key research questions and explain the significance of the study to the field’s research community. I introduce the not inconsiderable assumptions, social position, and preconceptions that I brought to this research and explore the implications these may have had on the progress and outcomes of this thesis. I conclude with a brief discussion of the limitations of the work and provide an overview of the thesis structure.

Chapter 2: The changing nature of precision agriculture research

Given difficulties of defining PA in abstract terms discovered in Chapter 1, the thesis opens with a bibliometric analysis of science published as precision agriculture (and its close variants) for the period 1996 to 2018 to profile PA development in terms of shifts in emphasis over time. Bibliometric analyses were applied to observe the evolution of the scientific literature in PA and to identify specific characteristics of the related knowledge domain. Specifically, I aimed to understand how the characteristics of technological assemblages may affect precision agriculture research and development. To achieve this, I traced the progress of the field through scientific publication and the results make some useful contributions to the body of knowledge on precision agriculture. I discovered there is rapid growth in the number of scientific outputs published under the PA banner and gained insights into the publication trends in the field, such as first author geolocation and trends in the institutions performing research. I also found that data-rich technologies have arrived in PA research, including a dramatic increase in remote sensing technologies such as hyperspectral sensing and imaging technologies. While the bibliometric analysis provides useful insights into publication trends, it yields little in terms of reflection on how the scientists actually operate.

Chapter 3: A brief history of the evolution of scientific data in precision agriculture

In Chapter 3 I extend my effort to define what precision agriculture is and to determine how the arrival of big data technologies may impact science-making in the field. I build on the bibliometrics analysis by charting the reimagining of agricultural science into modern precision agriculture research through a brief history of the evolution of scientific data in precision agriculture. I reflect on and critically review PA research to date and discover there has been significant research on the technologies that support precision agriculture research, both in terms of physical technologies, such as Global Positioning Systems, and the analytical tools and techniques that supported science-making. In this, and the preceding chapter, I identified that data-rich tools have arrived in precision agriculture research. However, I also found that there has been very little reflexivity amongst the field's research community, which leaves a significant research gap in terms of understanding how these new tools may impact the opportunities and challenges associated with their arrival.

Chapter 4: Research design and methods

In Chapters 2 and 3 I reviewed the literature relevant to this study. While these chapters yielded very specific substantive focus points around under-theorised lines of enquiry, even in the early stages of the study it became apparent that could not be resolved through literature-based study alone. When considering how to fill these gaps I determined that an ANT approach would be useful to describe socially and materially science-making systems. In particular, I selected an ethnographic approach with the aim of contributing to the understanding of the methods and approach of contemporary PA researchers. In this chapter I outline the research approach and methodology, including choices that were made as my research progressed. I discuss how and why I employed sensitivities and tools from the Actor Network Theory approach to follow the human and non-human actors in the P2P project. I discuss the procedures followed for data analysis and then discuss the role of reflective practice. I conclude with a review of ethical considerations made to protect the interests of the study participants and reflect on how they helped shape the study itself.

Chapter 5: Data acquisition in the new data economy

In Chapter 5 I deliberately narrow the focus of the thesis and turn an ethnographic lens to researchers working with nascent remote sensing technologies in a large research project called *Pioneering to Precision* (P2P). I introduce the humans and machines that are central to the P2P project and find that the innovation process is driven by the main actant, a remote sensing tool called the AisaFENIX. I follow the P2P research team into the field to explore how they collect field measurements and discover that the P2P project is framed around the notion that innovation is stable and predictable. In practice, I find that in the new data ecosystem[4] some things we previously thought were stable are not.

---

[4] This thesis employs the concept of ‘data ecosystem’ used by Medyckyj-Scott et al. (2016) who describe it as a system made up of people, practices, values, and technologies designed to support particular communities of practice. In such an ecosystem, ‘data is valued as an enduring and managed asset with known quality’ (Medyckyj-Scott et al., 2016, p. v). This ecosystem approach indicates that there is intent in the data - structure and quality of data are considered, rather than simply being a data lake or exhaust data.

Chapter 6: The all-data revolution: a new data ecosystem for precision agriculture research

In this chapter I extend the ethnographic investigation to identify and explore the ontological data characteristics of airborne hyperspectral imager AisaFENIX. I return to the laboratory to chronicle the efforts of the P2P researchers to prepare the AisaFENIX data and their supporting data for analysis. I find that data from these technological assemblages require significantly more processing to prepare them for analysis than anticipated, and as a result a new scientific division of labour in PA is also emerging.

Chapter 7: Beyond scarcity – the structure of the investigation and discovery process in an all data ecosystem

In this chapter I examine the iterative data investigation process, verification and scientific reporting performed by the P2P research team to refine value from the P2P dataset. I investigate where value resides in contemporary data ecosystems and explore some of the challenges and opportunities that big data brings to PA research. Framed by new data relationships, observations of the P2P researchers reveal that a new scientific method in PA is emerging, and it is a far cry from the linear classical science traditionally promoted for agricultural science. As how science is done changes, so to do other relationships. In this vein, in this chapter I also investigate the scientific division of labour within the P2P research team and examines how they adapted to meet the needs of a new data ecosystem. I discuss the intractability of linearity in scientific reporting and find a disjunction between how science is done in the field, and what is reported in academic literature, which has implications for how the field may progress knowledge in future.

Chapter 8: Embracing the beautiful mess: realising value in contemporary precision agriculture

In the final chapter I explore how researchers and policy makers can support contemporary PA researchers in the big data era. Learnings from the ethnographic study are examined, with data characteristics, and processing methods and practices suggested to assist researchers when repurposing PA data. I conclude the thesis with a reconsideration of how PA research funding and practices could be adapted to meet the needs of contemporary science-making, and, finally, I make some suggestions for further research.

Chapter 2: The changing nature of precision agriculture research

Introduction

In this chapter I aim to situate PA scholarship within a larger scholarly conversation by analysing scientific publishing trends as an indirect measure of the evolution of the field. To support the analyses, the chapter presents evaluative bibliometrics representing research trends based on data from Scopus and other selected sources. A large bibliometric dataset was assembled and analysed to identify the character and paradigmatic extent of shifts in the data and technological ecosystems that PA researchers work within.

Consideration of the characteristics of contemporary data is fundamental to understanding the raw materials upon which the researchers of today build knowledge. Global data generation is growing at an unprecedented rate, with over 2.5 quintillion bytes of data being generated every day (DOMO, 2019). As the production of agricultural data also booms, it is highly relevant to establish if technologies and the data characteristics of the technologies used by researchers are also changing.

This analysis is designed to support investigation of the overarching research question by answering the following questions:

1) What are the key trends in scientific publishing in precision agriculture?
2) Is the nature of data generated by precision agriculture technologies being researched changing?
3) Is the sphere of application of precision agriculture research changing?

Guided by the work of Aken (2004) and Verdouw, Beulens, Reijers, and van der Vorst (2015), the research reported in this chapter uses bibliometric analyses to address these previously unsolved problems. The artefacts developed for the study include; a) a longitudinal dataset of scientific literature produced between 1996 and December 2018, coded to assist aggregated analyses, b) a conceptual framework to support the typology of PA technologies and data characteristics, and c) a suite of bibliometric indicators that facilitate the temporal analysis of publications.

Why data?

Data have been paramount to PA research since its inception, yet there has been little scrutiny of temporal changes in the characteristics of PA data and the implications for the research community. Researchers such as Sonka and IFAMR (2014) explored the idea that the nature of PA data may be changing, suggesting that researchers will need to overcome organisational and technological challenges to refine value from datasets containing massive amounts of highly variable data. Importantly, researchers have identified that technologies with different data characteristics may have different adoption trends (Shaw & Willers, 2006). This suggests that a robust typology may be useful to guide future research on technology development and diffusion strategies (Griffin et al., 2004; Mulla, 2013).

In response to the research questions, the artefacts support analysis by understanding the character and paradigmatic extent of shifts in PA research over time. Existing typologies such as Kamilaris, Kartakoullis, and Prenafeta-Boldú (2017) and Wolfert et al. (2017) influenced the selection and organisation of data-related categories. However, this study is also shaped by the author's own observations of science-making in a PA laboratory using data rich technologies, assemblages, and analytical approaches. Importantly, the researchers followed faced previously undocumented challenges that seem linked to the characteristics of the data generated by the technologies they were developing.

The significance of systematic reviews in theory development

Summaries of past research are widely used to inform new enquiries (Adams, Smart, & Huff, 2017) and literature reviews are essential in academic research to construct a view of existing knowledge and to examine the state of a field (Cropanzano, 2009). Researchers typically collect available evidence on a topic or problem to assess the state of the existing evidence (Linnenluecke, Marrone, & Singh, 2020). However, literature reviews are traditionally narrative reviews that frequently only offer an arbitrary selection of evidence (Tranfield, Denyer, & Smart, 2003). Consequently, systematic reviews such as bibliometric reviews are preferred where the researcher requires a comprehensive background for theory development and testing (Sternberg, 1991). Linnenluecke et al. (2020) suggest that bibliometric analyses can assist researchers by:

> … establishing a context and delimiting a research problem; seeking theoretical support; rationalising a problem and new lines of enquiry; distinguishing what has been done from what needs to be done; identifying the main outcomes of (and methodologies used in prior studies); and avoiding fruitless research (p. 177).

Tranfield et al. (2003) describe the important distinction between narrative reviews and systematic reviews:

> Systematic reviews differ from traditional narrative reviews by adopting a replicable, scientific and transparent process, in other words a detailed technology, that aims to minimize bias through exhaustive literature searches of published and unpublished studies and by providing an audit trail of the reviewers decisions, procedures and conclusions" (p. 209).

A key difference between literature reviews and bibliometric analyses is that the latter requires the collection of a representative or comprehensive dataset of available research (Tranfield et al., 2003). A replicable process is employed to evaluate existing evidence to minimise bias that results from the random inclusion or exclusion of studies in the literature review process (Linnenluecke et al., 2020).

Recent technological advances in the accessibility and quality of bibliographical information sources such as Elsevier's Scopus, Google Scholar and Web of Science (SCImago Research Group (CSIC) OECD, 2016) have stimulated interest in bibliometric analysis in precision agriculture. Liu, Guo, and Guan (2010) were the first to apply bibliometric techniques to scholarship in the field of PA. Their modest study published in 2010 aimed to provide data for scholars on the study of PA and help them identify research hot spots and the development trends of the field. Indeed, the study sets a useful benchmark against which to measure trends in PA publishing today. M. Z. Cushnahan, Wood, and Yule (2017) also capitalised on these innovations to create and code a large dataset to explore changes in PA research, and to identify what shifts have occurred since 1996. They employed bibliometric analysis to identify authorship trends and importantly charted emerging data-rich technologies to inform discussion around data characteristics and assemblages in present-day PA research.

Wolfert et al. (2017) used bibliometric analysis to examine the role of big data in smart farming, which they viewed as an extension of precision agriculture. Their analysis of the role of PA data is of interest as they argued that big data is changing the scope and organisation of farming. The authors used bibliometric data to support their conclusion there has been a major shift in roles and power relations among different players in existing agri-food chains. Similarly, Kamilaris et al. (2017) used bibliometric analysis to explore the influence of big data on smart farming, identifying and analysing papers of interest, and importantly, examining the problem they addressed. Their study stressed the point that big data research and big data analytics are still in the early stages of development in precision farming, a finding inferred from the limited number of scientific publications and commercial initiatives.

Pallottino et al. (2018) pioneered the use of temporal term mapping to identify changes in term clusters of publications from 2000 to 2016. They identify three main research specialisms: one related to agricultural engineering; the second mainly assigned to computer science (this includes sensing and data analysis approaches, and interactions among information and communication technologies); the third area is associated with agronomic studies. These groupings are similar to those used by Liu et al. (2010) who separated publications into agriculture machinery, information obtaining and processing, variable-rate operation and research objects.

On enabling technologies, Pallottino et al. (2018) remark on the rising popularity of some technology terms in research publications, such as Zigbee and LiDAR, and the demise of others, such as the early electromagnetic sensor SPAD. The authors also use term analysis to identify specific linkages between, for example "shape" and "image" to the term "apple", suggesting that term analysis has the potential to identify trends in the focus of PA research for specific crops. They also suggested that other terms such as policy appear with closely surrounding terms such as "food security", "sustainability", "environmental impact", "climate change" and "greenhouse gas". Collectively these studies support the assertion that bibliometric indicators can offer a richer understanding of historical and modern trends in PA scholarship.

Taxonomies of precision agriculture technologies

While there have been previous attempts to classify PA technologies, no robust typologies exist that fully recognise data characteristics, including the volume and diversity, and the technological assemblages that generate them. The classification of PA technologies and their assemblages is fraught with difficulty, mainly because few tools employed in PA were invented for the purpose of PA research. The 'borrowing' of technologies from other fields is commonplace in PA research, making it nigh on impossible to set firm boundaries on what technologies are 'precision agriculture technologies'. Researchers routinely employ tools developed for research in fields as diverse as remote sensing, soil science, meteorology, automotive engineering, and food science. Exacerbating the complexity of analyses is the reliance of PA researchers on technological assemblages to build relationships between measurements and tools.

Early attempts to classify PA technologies by Sunding and Zilberman (2001) and Douthwaite, Keatinge, and Park (2001) divide innovations into 'embodied and disembodied', however, they employed different approaches. Douthwaite et al. (2001) viewed technology or components of technology as a representation of the new knowledge, or existing knowledge used in a novel way. In contrast, Sunding and Zilberman (2001) grouped innovations more broadly into those embodied in capital goods or products, and those not embodied in any physical item. These included non-software items such as a method of improving a system. Table 1 lists the main typographies used in PA research to date.

Table 1 *The main typographies used in PA research*

Classification	**Reference**
Soft' and 'Hard' technologies	Nowak (1997)
According to form (mechanical, biological, chemical, agronomic, biotechnological, informational technologies	Sunding and Zilberman (2001)
According to impact	Sunding and Zilberman (2001)
Embodied and disembodied	Sunding and Zilberman (2001); Douthwaite et al. (2001)
Information-intensive versus embodied knowledge	Griffin (2011)
Agricultural data sources	AgGateway (2014)
Systematic classification of issues and concepts for the analysis of Big Data applications in Smart Farming from a socio-economic perspective	Wolfert et al. (2017)

The separation of information-intensive and embodied knowledge by Griffin et al. (2004) was an important advance as they hypothesised the two categories may have different adoption trends. Later, Griffin revised his framework to explain that technologies that initially fall into the information-intensive category can shift to an embodied knowledge category, thus making the technology easier to use and potentially more likely to be used by farmers (Griffin, 2011). The later work clarifies the category definitions, noting that information-intensive innovations require additional data collection and upskilling of the operator, whereas embodied knowledge needs minimal additional inputs.

In 2014, AgGateway produced a white paper on data privacy and use, which includes a comprehensive section on sources of farm data. While not strictly providing a taxonomy, the white paper's results guide the codes applied in this chapter (AgGateway, 2014). Wolfert et al. (2017) use bibliometric analyses to examine the role of big data in smart farming. They present a useful conceptual framework to facilitate the systematic classification of issues and concepts for the analysis of big data applications in smart farming from a socio-economic perspective. The framework considers the following three classes: network management; the stakeholder network; and business process separately. The authors subdivide business processes into the data chain, the farm management and the farm processes (Wolfert et al., 2017). While the typologies developed by Wolfert et al. (2017) and Griffin (2011) do not fully classify the characteristics of technologies and data in PA, they provide a sound foundation from which a new taxonomy in PA can be formed.

Bibliometric methods

A longitudinal dataset of publications was initially populated using searches of Scopus (Elsevier) databases (https://www.elsevier.com/solutions/scopus). Scopus was selected for this study as the preliminary research examined the coverage of various bibliographic databases in the precision agriculture field and found Scopus provided both wide coverage of the topic and advanced bibliometric features including the ability to download searchable metadata as an .xml file.

The databases were searched using the following search types: TS=("precision agriculture"), TS=("precision farming"), TS=("site-specific agriculture"), TS=("site-specific farming"), TS=("soil-specific agriculture"), TS=("soil-specific farming"), TS=("digital agriculture") and TS=("digital farming"). Bulk search results were downloaded from the Scopus website as an .xml file, which was uploaded to a Microsoft Excel spreadsheet.

Each publication is treated as an individual unit of analysis. Data selected for inclusion in the dataset included the following categories: Title, Authors, Year, Source title, Volume, Issue, Article Number, Page start, Page end, Page count, Cited by, Affiliations, Authors with affiliations, Abstract, Author Keywords, References, Document Type, Source and EI. The preliminary study also revealed that some articles from the canonical publication *Precision Agriculture Journal* are not indexed in the Scopus database, so the dataset was supplemented with the entire *Precision Agriculture Journal* catalogue, sourced from the Springer website (http://www.springer.com/gp/products/journals). Including grey literature such as blogs, magazine articles and government papers was considered as it may signpost technology changes earlier than peer-reviewed scientific articles (Adams et al., 2017; Liu et al., 2010). The criteria that drove the decision to omit these publications were: a) robustness and independence of content cannot easily be determined; b) metadata required for further analyses is not available.

Peer-reviewed conference papers from some PA conferences did fulfil these criteria however and hence papers presented at the following events were included where available:

- International Conference for Precision Agriculture (1998, 2000, 2008, 2010, 2012, 2014, 2016, 2018). Limited data for the 2002, 2004 and 2006 is included where available.
- Proceedings of the First Workshop in Soil Specific Crop Management (1993)
- European Precision Agriculture Conferences (where data available).

Data were unavailable for the International Conference for Precision Agriculture in 2002 and 2004. However, some papers were accessed on ResearchGate (www.Researchgate.com) and professional research websites, e.g. the website for the International Society of Precision Agriculture (www.ispag.org). Data were exported to a single Excel spreadsheet and duplicates removed. Each record was checked for relevance and papers not relating to PA were removed from the working files. Where the abstract text was not complete, internet searches (mainly www.Researchgate.com) were used to populate the field; the publication was removed from the working files where this information was unavailable. This accounted for fewer than five percent of the publications initially included in the dataset. The following bibliometric indicators were used to explore the data to identify key features and discern trends in scientific publication:

- Trends in publication frequency, key publications
- Trends the different types of publication such as journal articles and conference papers.
- Authorship trends
    - Global distribution of research (country of first author)
    - Number of authors per journal article
    - Intercountry authorship
    - Patterns in the social networks of researchers

As the other areas of my research took longer than originally planned, my data was updated to ensure the analyses were still relevant. My final data extraction occurred in late 2019. This included the extraction of updated citation information for each record using the Google Scholar website (https://scholar.google.co.nz/) as a reference. I recorded the total number of citations for each publication against each publication's individual record in Excel. Contemporaneously, I accessed affiliation data for first authors from Scopus, the Springer Journal website and where necessary from other sources. Data sources included Google Scholar author profiles, ResearchGate (https://www.researchgate.net/home) and the International Society for Precision Agriculture's website (https://www.ispag.org/Proceedings). Google Scholar was used for this task as it and its associated platform for academic profiles, provide a precise and accurate picture of the bibliometric community (Martín-Martín, Orduña-Malea, Ayllón, & López-Cózar, 2016). Importantly, Google Scholar offers the widest coverage of the bibliometric social media platforms (Orduña-Malea, Ayllón, Martín-Martín, & López-Cózar, 2015).

Classification of precision agriculture technologies and their assemblages

To present an accurate picture of the trajectory of contemporary PA research, this study attempts to identify key technologies within the research community and explores how they are applied. Existing typologies of agricultural machinery mainly focus on power sources (motor, animal) and where in the production chain the machinery is used, e.g. planting, maintenance, harvest, and post-harvest (Nowak, 1998; Sunding & Zilberman, 2001). While some other sources, such as AgGateway (2014) group data into various sets, this chapter is inspired by the work of (Wolfert et al., 2017). In their bibliometric study of smart-farming publications, Wolfert et al. (2017) take an ANT perspective, by acknowledging data as an actant, deliberately categorising technologies and the data they create. Unlike other taxonomies which often solely focus on technologies themselves, the authors classified publications in smart farming from a socio-economic perspective to demonstrate the scope of the influence big data technologies are having on primary production.

I also coded publications to technology classes to monitor which technologies researchers employ in their research over time. To support analysis of the sphere of application of technologies and practices, I separated technologies into three key domains: recording technology, reacting technology and guidance technology. This approach aligns with that recently taken by Balafoutis et al. (2017). Records were coded to a technology where it was used in experiments or was the focus of the study (not simply mentioned). Publications were frequently coded to multiple categories within a single domain, for example, coded to both the recording technology and reacting technology class.

Conceptual framework

The aim of the framework produced for this thesis is to support aggregate analyses to discover shifts in publication volumes and content over time. I designed a typology to capture who is undertaking research in the field of PA and what the data generated by their research look like. The Open Coding method described by Corbin, Strauss, and Strauss (2015) was employed to inform the typology, which is designed to simplify the systematic classification of scientific outputs. Each publication was manually coded for their main subjects (not just mentions) to domains at the macro-, meso- and micro-level (see Table 2). Codes were then entered into an Excel Spreadsheet against the individual publication record.

Table 2 *Typology of PA research publications employed in this thesis*

<table>
<tr><th colspan="2">Macro Level</th><th>Meso Level</th><th>Micro Level</th></tr>
<tr><td rowspan="4">Broad subject areas and subdivisions</td><td>**Crop type**</td><td>Crop type subgroups</td><td></td></tr>
<tr><td rowspan="2">**Enabling technology**</td><td>Data characteristics</td><td>Data characteristic subgroups
Data intensive</td></tr>
<tr><td>Technology classification</td><td>• Guidance technology
• Recording technology
• Reacting technology</td></tr>
<tr><td>**Measurement**</td><td>• Crop monitoring
• Soil mapping
• Terrain analysis
• Climate
• Water monitoring
• Livestock monitoring</td><td></td></tr>
<tr><td>Specific themes</td><td colspan="3">• Precision agriculture practices
• Management zones
• Technology transfer/ uptake
• Benefits of precision agriculture
• Applications of precision agriculture
• Farmer characteristics
• Research practice
• Geostatistics
• Information management
• Decision support systems</td></tr>
<tr><td colspan="2" rowspan="4">Bibliometric information</td><td>Publication name</td><td rowspan="3"></td></tr>
<tr><td>Publication type</td></tr>
<tr><td>Number of citations</td></tr>
<tr><td>Country of author</td><td>Organisational affiliation of first author</td></tr>
</table>

Coding data characteristics

The typology draws on the work of Wolfert et al. (2017) and Kamilaris et al. (2017) in terms of the classification of PA data characteristics. Publications are grouped into types that allow a technology or assemblage to be classed as ‘data-intensive’. While not all individual technologies may meet all the characteristics of big data, the looser classification of data-intensive was designed to capture data on technologies and assemblages that produce high volumes of data in a short period (near real-time), or data that can be aggregated on a large scale (Sonka & Cheng, 2015b), rather than simply using volume as the key criteria.[5] Technological assemblages coded to the data-intensive category include; but are not limited to a) hyperspectral sensors, b) machine vision, (c) machine learning/ fuzzy clustering, (d) Internet of Things, (e) Artificial Neural Networks, (f) wireless sensor networks and (g) robotics (and associated technologies).

Coding is also used to recognise trends in the popularity of sensors, an important technology in PA research. For this study, the following technologies are classed as sensors; a) Electromagnetic Induction (EMI) technologies, b) hyperspectral sensors, c) Multispectral sensors, d) Light Detection and Ranging (LiDAR) sensors. Publications were coded to individual technologies such as LiDAR. A single publication could be coded to multiple sensor types if employed individually or as an assemblage in the study. This allowed trends in the popularity of the individual technologies to be captured as well as allowing for aggregation to provide an indication of wider trends in sensor use and assemblages. Remote sensing data collection platforms were also categorised so that trends pertaining to platform utilisation can be analysed.

Sensors require a platform to work from and these are pedestrian, vehicle, UAV, aircraft, and satellite (Yule, 2015). For this study, I combine the pedestrian/ handheld and vehicle platforms into the proximal code, which represents those sensors deployed very close to the target (usually ground-based). This is because the data from these platforms do not typically require atmospheric corrections or orthorectification of data or images before the data can be analysed. As the pre-processing requirements for data from unmanned aerial vehicles, satellite and other airborne data collection platform varies, each platform is classified under its own code:

- **Airborne**: includes research publications describing the data collection platform as airborne, aerial, aeroplane, aircraft, helicopter, balloon, and kite.
- **Unmanned aerial vehicles (UAVs)**: includes research publications describing the data collection platform as: Unmanned Aerial System (UAS), Unmanned aircraft, drones, Remotely Piloted Aircraft Systems (RPAS) and Unmanned Aerial Vehicles (UAVs).
- **Satellite**: includes research publications describing the data collection platform as ‘satellite’, e.g. Landsat, GF-1, Sentinel, Hyperion, and Spot. Satellite navigation systems such as GNSS, GPS, GLONASS, Galileo, and Beidou are excluded from this code.
- **Proximal**: includes research publications describing the data collection platform as proximal, tractor-mounted and/or ground-based.

These categories are useful as they have implications for data production, processing, and classification.

---

[5] Due to rapid developments in this area, a unifying definition of what constitutes high volume is difficult to give, but generally it is a term that refers to data volumes so large that traditional data storage and processing applications are inadequate (Wolfert et al., 2017).

Limitations of the bibliometric analysis

Decisions on which publications to include in this study were guided by the benchmark set by Liu et al. (2010). This includes the decision only to include publications in the dataset when a full abstract was available in English.[6]

An important qualification regarding the creation of this typology is that like any classification, I encountered technologies and applications that were difficult to classify or could bridge multiple classes. I have attempted to keep categories simple and some are only illustrated with a few examples in the text. I do however recognise the perennial dangers that lie in the application of any classification scheme (Van Maanen, 2011) and accept that some readers may find the categories too broad, too encompassing and indeed, too categorical. Consequently, the typology assembled for this thesis is not intended to be exhaustive and others may organise and assign different technologies to other classes. My key objective is to create a transparent, replicable classification system that facilitates the organising and directing of thoughts rather than collecting categorical examples and arranging them unequivocally.

The Scopus investigation was not initially restricted by publication date or type. The later decision to limit the temporal analyses of the dataset to 1996–2018 was informed by a preliminary study of relevant literature dating as early as the 1950s, which found few published research outputs in PA before 1996, after which there was a dramatic rise in research activity.[7] While this aligns with the reported expansion of PA research (D. M. Oliver et al., 2012), it should also be noted that when Scopus launched in 2004 its pre-1996 coverage was very limited, so prior trends may not be fully accurate (Elsevier, 2017). While earlier materials are available in indexed repositories, the publication records are often missing digital (searchable) key bibliometric information and hence are excluded from this study for reasons of expediency.

Another limitation relates to the classification of publications. Semantic classification has been criticised as a rigid method that has difficulty dealing with rapid changes in the nomenclature and terminology ambiguities between and within disciplines (Hey et al., 2009). These concerns are relevant to this study as data science and PA are relatively new fields of inquiry and nomenclature is highly variable. The typology created for this study draws on earlier works in the field and is designed to help address problems with variability of nomenclature. Similar terms such as precision farming, smart farming and soil-specific farming are aggregated into precision agriculture; this is particularly important as it allows earlier studies that are not labelled as

---

[6] Some analysis of Chinese-language journals was performed in the preliminary work for this thesis.

[7] Pallottino et al. (2018) limit their study to 2000 onwards because 'prior to that year the published papers were less than 50 per year'. However, in their search terms they omit the terms soil-specific and site-specific, which were commonly used to describe precision agriculture during this period.

‘precision agriculture’ to be included. It is recognised however, that aggregation may lead to some records being classified to the wrong domain. A further limitation of the analysis is that codes are applied to records based on limited information such as key words, abstracts, and titles. Where ambiguity exists, the full paper was referred to where available. That withstanding, such assumptions are likely to have led to an underestimation of some records coded.

The results provide useful insights into publishing trends in PA, particularly the technologies employed and the sphere of application of PA research. However, the study was unable to shed much light on important ontological and epistemological questions concerning contemporary PA research. Author data included some significant data, such as the author’s country, which was useful when considering the social networks of researchers. Importantly however, the affiliation data do not provide reliable insight into the researcher’s background, level, or area of expertise, nor their motivations. Information in abstracts, especially relating to methodology, is limited in detail and information on the scientific process and scientific division of labour is rarely included.

Classifying PA technologies was complex because of the abundance and diversity of technologies used by farmers and researchers. Also, scholarly communication in the field encompasses many activities excluded from traditional peer-reviewed publication channels. These include many conference presentations, informal seminar discussions, face-to-face or telephone conversations, formal journal and book publications, interaction with social media platforms, email exchanges, and grey literature such as website posts, blogs, and digital objects. Indeed, a great deal of PA research is never published in academic publications. Commercial interests concerning the protection of intellectual property mean that in many cases, even when research is performed at educational institutions such as universities, research is embargoed or never published in the public domain.

Trends in precision agriculture publishing

The first PA article in my dataset of publications arrived in 1956 and 148 publications were indexed to the field prior to 1990. This serves as a gentle reminder that PA principles are not entirely new. Rather, it is the arrival of Global Positioning System (GPS) technologies in the early 1990s that facilitates the expansion and application of PA research. The bibliometric indicators support the notion that this field of research is growing rapidly in terms of publication outputs. Documents published in scholarly journals and conferences have increased more than fourfold over the past twenty years from 207 in 1998 to 919 in 2018. Furthermore, the field is increasingly diverse in terms of researcher geolocation, sphere of application and the journals in which research is published.

The bibliometric analyses also provide a glimpse into the extraordinary rise of remote sensing technologies in PA research. The evaluative indicators provide credible evidence that data-rich remote sensing technologies and assemblages such as Hyperspectral Sensing and Imaging (HSI), and the data they generate, are highly influential in modern PA research.

Publication trends

Precision agriculture's broad appeal is evident with 11,597 published papers in the field in 2,432 different journals and professional conference proceedings between 1956 and 2018. PA is commonly assumed to have emerged in the 1980s (Mulla, 2013), however this study reveals a small number of publications are listed in the Scopus database as early as the 1950s. For example, Gunther and March (1956) and Stolzy, Cahoon, and Szuszklewicz (1957), on topics still considered contemporary, such as site-specific irrigation based on soil moisture monitoring. Compared to related fields, PA publications increased rapidly in the decade to 2006 and while the field continues to grow, the rate of growth has slowed compared with other emerging specialist fields such as agricultural robotics. This may be due to the evolution of the field and negative influence of the global financial crisis on university research budgets in the USA and Europe (Wolinsky, 2009).

Figure 1 shows the distribution of PA journal articles and conference papers published between 1996 and 2018. Conference papers and journal articles are the leading forms of scientific communication, together accounting for over 96% of publications. While the publication trends for articles are quite smooth, the number of conference publications regularly spike in line with the hosting of the biennial international precision agriculture conference. While only twelve PA books were listed in the database between 1996 and 2018, they are well supported with citations, with 164 citations per publication.

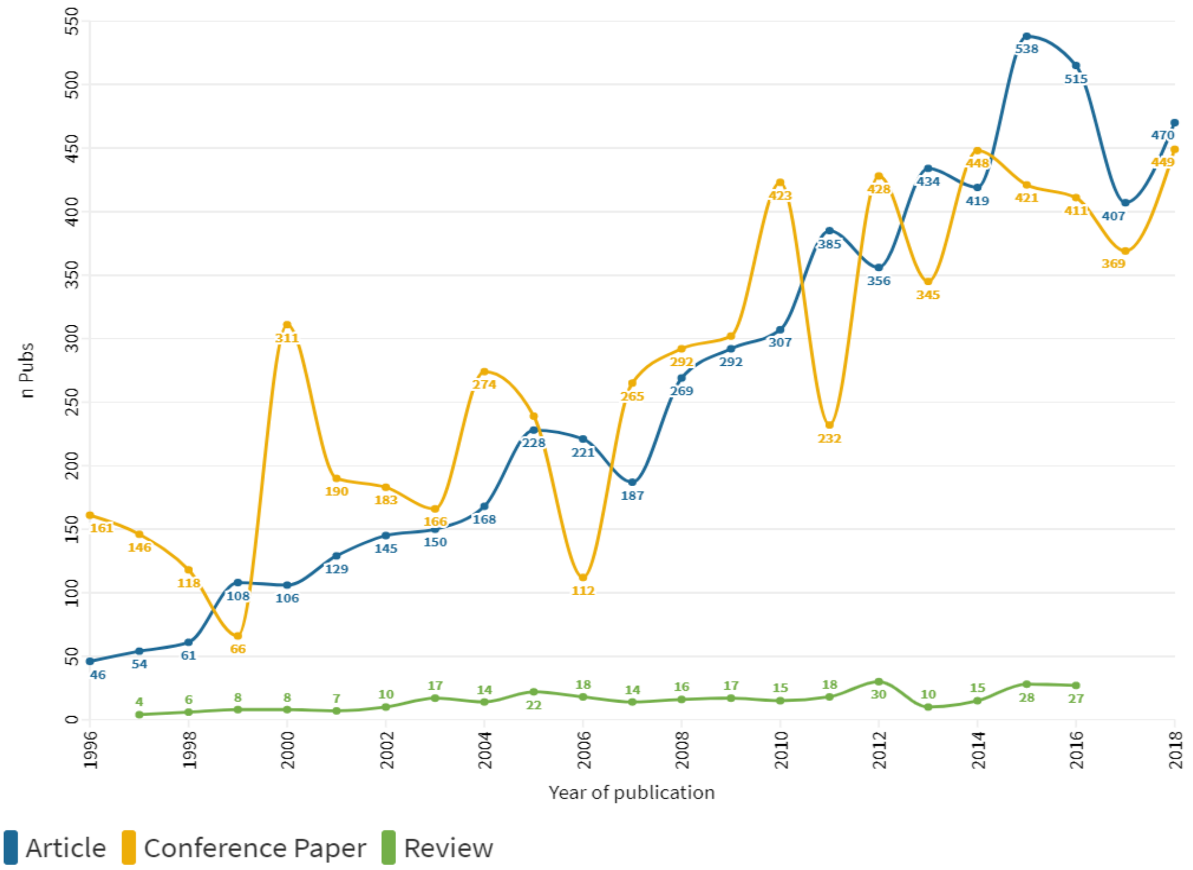


Figure 1 *Number of PA publications by document type, per year, 1996 - 2018*

A shift in precision agriculture research output

The first bibliometric analysis of PA research by Liu et al. (2010) provides a useful benchmark to measure progress. Then, thirteen periodicals had published more than fourteen papers in total since 1990. The top five periodicals accounted for an impressive 79% of the total citation frequencies of PA papers, with agricultural engineering and soil science publications dominating the top ten publications (Liu et al., 2010). Just nine years later, the number of periodicals that have published more than fourteen documents has exploded to 128, illustrating the broad appeal of PA in academic publishing today. Despite the widening sphere of application, there are still dominant publications and the top five publications (*n*pubs) still account for a quarter of all citations. Importantly, the periodicals in which researchers choose to communicate their findings in the past nine years are changing.

Today, the two multidisciplinary publications *Precision Agriculture* and *Computers and Electronics in Agriculture* (both based in the Netherlands) stand as the field's leading publications both in terms of publication numbers and citation frequency. While the influence of agricultural engineering publications also remains, the results in Figure 2 represent a considerable shift away from domain expertise in soil science and agronomy to remote sensing-related publications.

Domain-specific publications that were influential in 2010, such as *Agronomy Journal*, *Journal of Soil and Water Conservation*, *Soil Science Society of America Journal*, *Communications in Soil Science and Plant Analysis* (Liu et al., 2010) have been supplanted in terms of influence; in 2017-2018, collectively these journals only produced eight publications in precision agriculture. In contrast, the specialist remote sensing periodicals *Remote Sensing* (Switzerland), *Sensors* (Switzerland) and *Remote Sensing of Environment* (Netherlands) produced sixty-seven articles in the same period, signalling a shift from domain-expertise to a technology-based research focus.

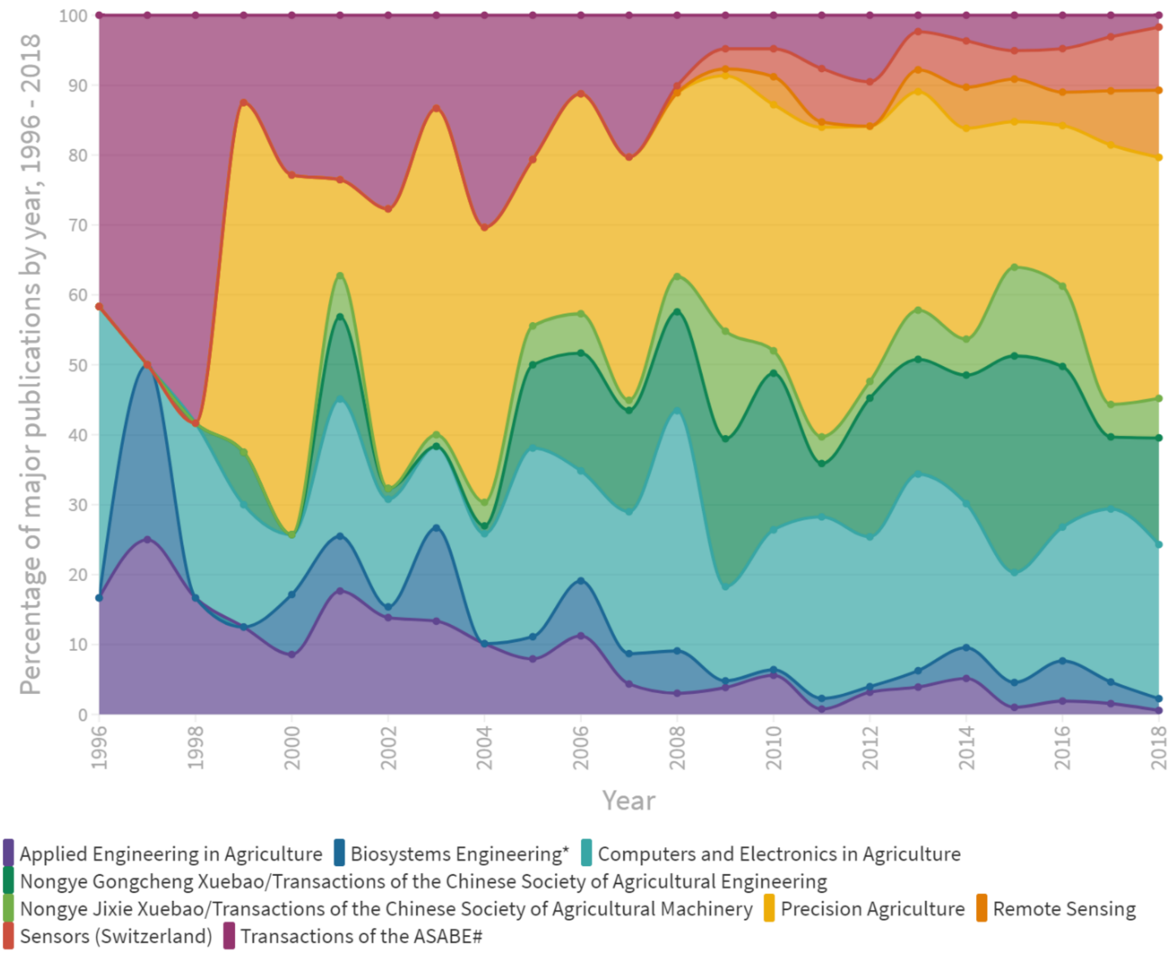


Figure 2 *Statistics of major PA publications, 1996-2018*

*From 2002, “Journal of Agricultural Engineering Research” changed its name to ``Biosystems Engineering”.

#From 2005, “Transactions of the ASAE” changed its name to “Transactions of the ASABE”.

Another feature of the data is the rise of the Chinese-language agricultural engineering periodicals which publish with a full abstract in the English language. *Nongye Gongcheng Xuebao* and *Nongye Jixie Xuebao* are becoming far more influential in terms of publication numbers. Figure

2 also shows a shift in the popularity of American publishing houses; while seven of the top thirteen publishing houses in Liu’s 2010 work were from the USA (Chen et al., 2009), that number has dropped to four. European publishers have superseded the Americans, with four periodicals from the Netherlands, two from Switzerland and one from Belgium featuring in the top thirteen journals up to 2018. Importantly, Dutch contributions include the two highest ranking journals, *Precision Agriculture* and *Computers and Electronics in Agriculture*, while the American *Transactions of the ASABE* has moved from its comfortable top ranking in Liu’s study in 2010 (Liu et al., 2010), down to fourth position in 2018.

Geolocation of research

The geolocation of first authors provides an insight into who is performing research in the field. The dataset was analysed to study the global distribution of scientific production since 1996. Based on the results, PA research has become far more diverse in the past twenty years. In 1996, first authors from only 23 countries published in the field, a number that had more than doubled to 69 in 2018.

In terms of total publications, authors based in the USA still easily lead the world, both in terms of publication number and citations (see Table 3). In Liu’s 2010 study, thirteen countries had a total of over 20 first authorships since 1990 (Liu et al., 2010). In 2019, an impressive 49 countries had 20 first authorships or more, highlighting the increased diversity in the field and cementing PA as a truly global field of study. The countries that have the most first authorships are the USA, China, Germany, Brazil, and India, with a 25%, 16%, 7%, 6% and 4% share of first authorships respectively. Canada, which ranked fourth in publication outputs in Liu’s study in 2010, has slipped to eighth position in 2018, with 3% percent of first authorships compared to 6% in 2010.

Table 3 *First author statistics of all precision agriculture publications, top twenty countries, 1996-2018*

Country	**Year of first publication**	***n*pubs**	**Share of publications (%)**	***n*citations**	**Citations per publication**
USA	1956	3,125	25	98,259	31
China	1987	2,011	16	15,352	8
Germany	1984	819	7	17,801	22
Brazil	1991	698	6	6,472	9
India	1985	496	4	6,779	14
Spain	1993	439	4	12,558	29
Australia	1976	431	3	15,523	36
Canada	1980	377	3	11,718	31
Italy	1981	353	3	7,681	22
United Kingdom	1981	305	2	9,833	29
France	1989	288	2	6,712	23
Japan	1984	254	2	3,305	13
Netherlands	1962	171	1	7,661	45
Greece	1995	153	1	2,585	17
Belgium	1990	152	1	3,006	20
Denmark	1988	151	1	3,402	23
Iran	2002	136	1	1,462	11
Malaysia	2002	130	1	1,214	9
Israel	1988	123	1	1,815	15
Republic of Korea	2000	111	1	936	8
Other		2,310	15	29,281	13

The temporal analysis reveals a shift not only in the influence of Canadian researchers, but more importantly of their American peers. As shown in Figure 3, authors from the USA dominated the field until the mid-2000s, producing over half of all global PA publications in the late 1990s. Subsequently the USA's share of first authors has gradually reduced to a quarter of publications in 2018. Research based in the USA has simply not kept pace with the massive growth in global scientific publishing in PA led by new players based in China, India, Spain, and Brazil. Together these countries have seen a staggering rise in their share of global publications from a little over one percent in 1996 to 33% in 2018 (first authors).

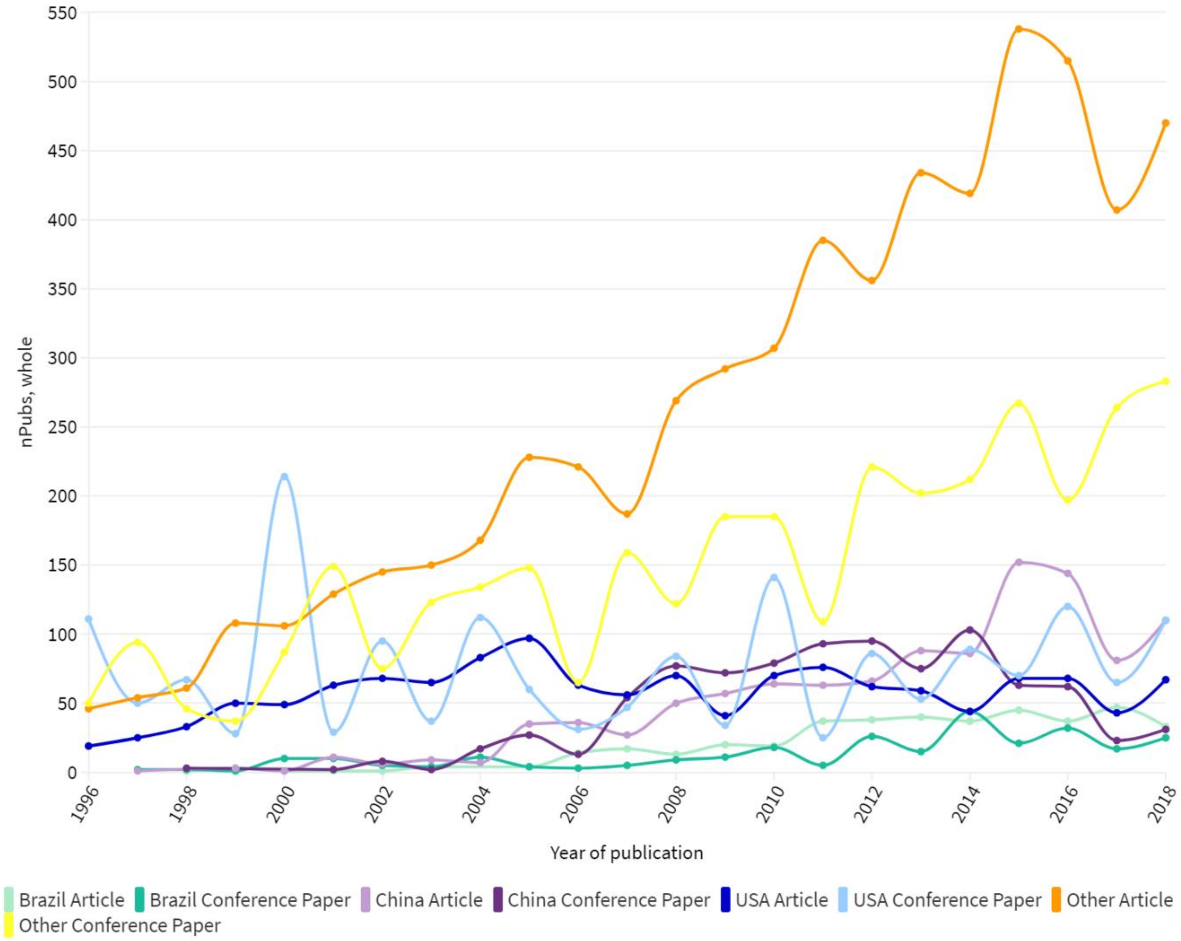


Figure 3 *Publication trends of the top three countries, 1996 – 2018*

The role of conference papers in precision agriculture research

Peer-reviewed publication outputs form the basis of traditional quantitative measures of academic endeavour and the research community continues to see peer review as fundamental to scholarly communication (Johnson, Watkinson, & Mabe, 2018). However, professional conferences can play a key role in configuring emerging fields of enquiry and act as venues for sensemaking (Garud, 2008). Conferences are especially important in terms of technology development because ‘failures’ are rarely published in academic journals. Oester, Cigliano, Hind-Ozan, and Parsons (2017) contend:

> Many times, techniques and projects that did not work out as planned are not published in scientific literature. Due to the limited space in journals and the suspected potential effect on the researcher’s reputation by publicizing work deemed as “failures,” malfunctions in research design are not widely publicized. Speaking to other researchers personally at conferences is where these conversations of failed attempts occur. Learning what has not worked for other researchers can help scientists save precious resources,

> including time and money. These conversations can also help researchers redesign unsuccessful projects (pp. 3-4).

Perhaps this is why the PA research community embraced conferences with fervour; the bibliometric analysis shows 6,348 precision agriculture-related conference papers were delivered at 1,129 different conferences between 1996 and 2018. Furthermore, there is evidence that the field of PA *emerged* from conferences. The first dedicated PA conference was held in 1990 (M. A. Oliver et al., 2013), where several papers now considered seminal were presented. The number of conference papers and field-specific conferences has steadily increased since. The major conferences are the *International Conference for Precision Agriculture* (and its predecessors) and *European Conference for Precision Agriculture*, which meet biennially on alternate years, see Table 4.

Analysis of the dataset reveals that PA papers are often presented at remote sensing conferences, i.e. SPIE (The International Society for Optical Engineering) and IGARSS (International Geoscience and Remote Sensing Symposium). The growing number of papers presented at conferences beyond the traditional boundaries of agricultural engineering and PA provides an opportunity to communicate PA research to the broader scientific community. It is important, however, not to assume that all research published beyond the domain is being conducted by traditional PA researchers. Rather, it is likely that specialists from other fields whose skills are needed will perform some of the research to address the challenges of the new data ecosystems.

Table 4 *Statistics of leading conferences delivering precision agriculture papers, 1996 - 2018*

Conference provider	***n*pubs (PA)**	***n*citations (all pubs)**	***n*citations per paper**
International Conference on Precision Agriculture	2,050	10,575	5
European Conference on Precision Agriculture	1,097	9,013	8
IEEE Conference (combined)	647	4,742	7
ASABE Conferences (combined)	507	2,287	5
IFIP Conferences (combined)	392	196	<1
SPIE Conferences (combined)	246	1,073	4
IGARSS Conferences (combined)	131	669	5
ISPRS Conferences (combined)	91	790	9

- IEEE - Institute of Electrical and Electronics Engineers
- SPIE - The International Society for Optical Engineering
- ASABE - ASAE/ ASABE Annual Meeting
- IFIP - International Federation for Information Processing

- IGARSS - International Geoscience and Remote Sensing Symposium
- ISPRS - International Archives of the Photogrammetry, Remote Sensing and Spatial Information Sciences

Analysis of the dataset indicates that authors from most of the top ten publishing countries initially published more conference papers than articles, which is not unusual given the early popularity of conference papers shown earlier in Figure 1. Since 2013 however, researchers from most countries have produced slightly more articles than conference papers, although the publication types have generally tracked together for the last decade. There are two exceptions, the publishing powerhouses China, and the USA. Until 2011, Chinese researchers published more conference papers than peer-reviewed articles. There has since been a dramatic fall in conference papers compared to articles, suggesting a change of research focus, and maturing of the Chinese research market, see Figure 3. The absence of major conferences in Asia is also likely to be a factor; the failure of the influential *International Conference for Precision Agriculture* (ICPA) to move beyond the confines of North America make conferences a less attractive proposition for researchers from outside the region, such as Brazil and China. This contrasts with the heavy presence of American researchers at domestic and overseas conferences compared to their article output. Figure 3 also shows a biennial spike in American outputs, which can be traced to over-representation at ICPA.

In 2016, American first authors presented 43% of ICPA papers, triple their global share of journal publications (13%). In contrast, first authors from China, now the largest article publishing country, produced only 5% of papers at the ICPA 2018. In the 2018 ICPA held in Canada, North American authors accounted for more than 45% of conference papers, compared to 15% from all Asian countries combined.

The under-representation of Asian authors appears to be limited to the conference platform; for peer-reviewed article output, Asian authors have a similar share of global output (30%) to North American authors (26%). European authors are also underrepresented at ICPA, although not to the same extent, with 21% of ICPA papers in 2018 compared to their 26% share of global article publications. Interestingly, Australasian authors are over-represented at the ICPA conferences; in 2018, Australasians presented 7% of the ICPA papers compared to 4% of global article output that year. The highly regarded *European Conference on Precision Agriculture* (ECPA) is also highly influential within PA research, especially for European researchers. Despite the associated travel distances, American authors still have a significant presence at ECPA, delivering 14% of the papers at the 2017 conference.

The important revelation that the diversity of authors presenting at key conferences is not representative of the research community should be of concern. As noted above, there is compelling evidence that researchers are more likely to communicate ‘failures’ at conferences than publish them in the scientific journal platform (Oester et al., 2017). Conferences also provide a valuable pathway to academic impact for early career researchers (de Leon & McQuillin, 2020). Thus, this research identifies a challenge for the field, which is to find ways to make conferences more inclusive for international researchers to support the global advancement of the field.

Regrettably, the bibliometric analysis also shows that proceedings for some major PA conferences are not readily available online, even on paid platforms. Online availability of scientific literature, especially free access, offers substantial benefits to science and society (S. Lawrence, 2001). Decades of research in science studies have shown that scientific progress is embedded in social practices (Latour & Woolgar, 1986; Pickering, 1992), which can be hidden when findings are merely announced from a stage (Jacobs & McFarlane, 2005). Thus, the legacy of written proceedings is important for ongoing scrutiny of investigations and the advancement of scholarship in the field. At a minimum, professional research bodies such as the International Society of Precision Agriculture should aim to make research easy to access by publishing indexes of proceedings in the public domain.

Trends in institutional affiliations

First authors from 2,971 different organisations published PA outputs between 1996 and 2018 (all publications). Not surprisingly, authors from national research organisations from large countries feature prominently in publication statistics. The United States Department of Agriculture (USDA) and the Key Laboratories/Chinese Academy of Sciences (combined) had the most first authorships at 597 and 516 publications respectively. Authors from eight universities have produced over 100 publications since 1996 (Table 5).

The China Agricultural University is by far the highest ranked university in terms of total first authorships with 214 first authorships and 2,377 citations. Somewhat surprisingly, the highest ranked American university, the University of California, is only the fourth ranked university, with the University of São Paulo and Wageningen University placing second and third. The broad diversity of research institutions that underpin the research community is illustrated by organisations from nine different countries featuring in the top twenty organisations.

Table 5 *Leading research organisations - all PA publications, top twenty countries, 1996-2018*

Organisation	**Country**	**Ranking**	***n*Pubs**	**Citations**
United States Department of Agriculture (USDA)	USA	1	597	20,236
Key Laboratories/ Chinese Academy of Science	PRC	2	516	5,399
China Agricultural University	PRC	3	214	2,377
University of São Paulo	BRA	4	146	1,432
Wageningen University	NLD	5	140	5,682
University of California	USA	6	138	5,574
University of Nebraska	USA	7	134	5,740
Spanish Council for Scientific Research (CSIC)	ESP	8	123	4,093
Leibniz Centre for Agricultural Landscape Research	DEU	9	119	2,334
University of Florida	USA	10	116	2,866
University of Bonn	DEU	11	113	3,686
Zhejiang University	PRC	12	113	1,431
University of Kentucky	USA	13	104	1,404
Agriculture and Agri-Food Canada	CAN	14	98	1,669
University of Georgia	USA	15	95	2,244
University of Minnesota	USA	16	94	2,448
Institut National de la Recherche Agronomique	FRA	17	88	2,140
Commonwealth Scientific and Industrial Research Organisation (CSIRO)	AUS	18	87	2,174
Iowa State University	USA	19	87	2,538
National Engineering Research Center for Information Technology in Agriculture	PRC	20	82	534

Overall, universities dwarf government organisations in PA publishing. Over 65% of all first authorships come from universities and other academic institutions compared to the 21% of publications from government bodies. The publishing gap between universities and governments, companies and other research organisations is widening; in 2006 university-based authors produced 62% of all publications, in 2018 that figure had risen to 81% of all PA publications, see Figure 4. This is interesting because we know that there is significant investment being ploughed into agricultural technology companies, including start-ups, yet there is little record of their research efforts in the academic publishing domain (Ryan, 2019).

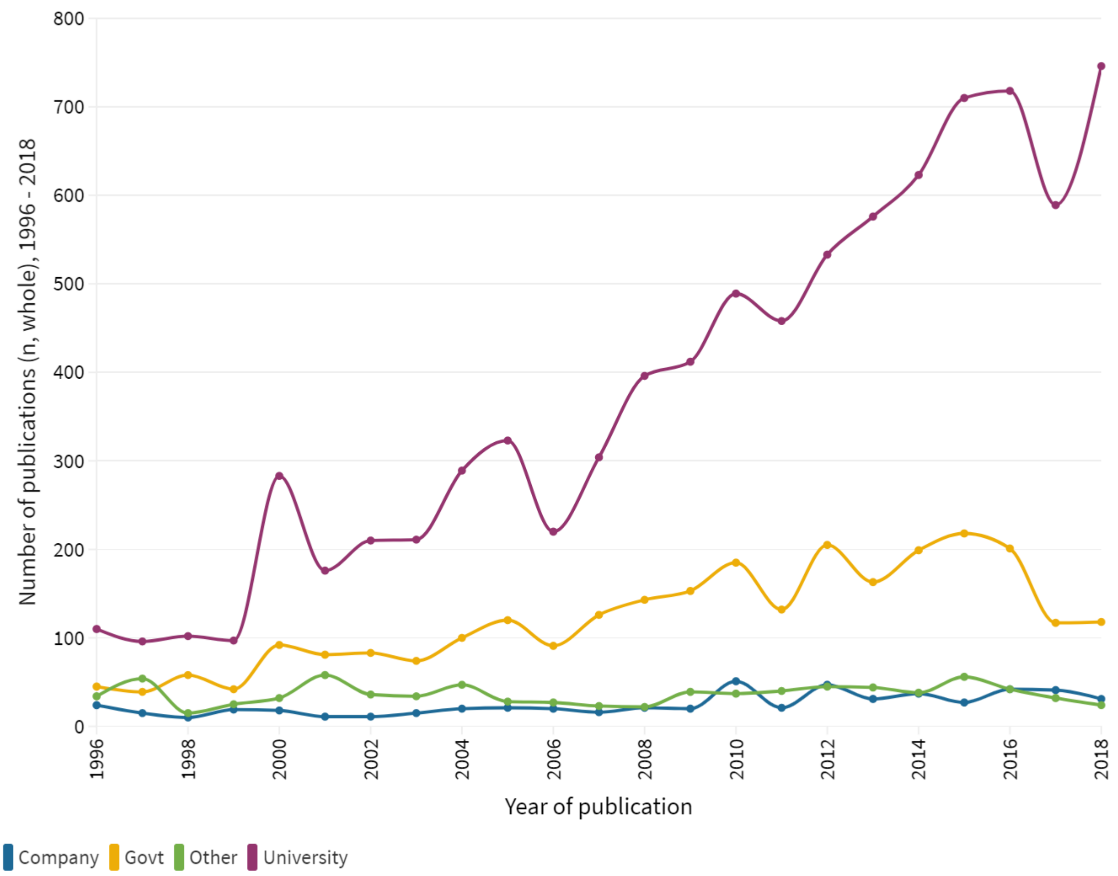


Figure 4 *The rising influence of universities in PA research, 1996 – 2018*

Trends in hyperspectral sensing and imaging (HSI) publishing

Later in this thesis, the focus narrows to investigate how the characteristics of hyperspectral sensing and imaging (HSI) assemblages affect precision agriculture research and development. Hence, it is useful to pause and explore publication outputs of researchers utilising HSI technologies in more detail. Lelong, Pinet, and Poilvé (1998) and Staenz, Szeredi, and Schwarz (1998) published the first PA articles in the HSI subfield and since then 300 articles have been published in 89 different journals. The analyses reveal some interesting differences in the publishing habits of HSI researchers compared to the wider a PA community. For example, while the *Precision Agriculture* journal and *Nongye Gongcheng Xuebao* still rank highest, as expected, most of the remaining top ten journals are specialist remote sensing publications. Interestingly, the top ten journals account for over 61% of all publications in the subfield, indicating that researchers may be targeting a select few journals, or that the research is presently not being applied far beyond the remote sensing domain. A concern for HSI researchers will be the low mean citations achieved for their papers, which are significantly less than for the wider PA field, even when considering age of publication. This suggests that despite its high profile, HSI research may have a limited research audience.

Table 6 *Statistics of major periodicals that publish documents in PA, subfield HSI, 1998-2018*

Journal title	**Year founded**	***n* citations**	**Mean citations**	**SJR index (2018)**	**HSI *n*pubs**
Precision Agriculture	1999	2,075	20	0.729	59
Nongye Gongcheng Xuebao/Transactions of the Chinese Society of Agricultural Engineering	1985	148	8	0.422	23
Guang Pu Xue Yu Guang Pu Fen Xi/Spectroscopy and Spectral Analysis	1991	127	6	0.215	19
Remote Sensing	1992	640	6	1.43	18
Transactions of the ASABE[8]	1907	648	5	0.400	15
Remote Sensing of Environment	1969	3638	4	3.208	13
Computers and Electronics in Agriculture	1985	796	4	0.950	11
IEEE Journal of Selected Topics in Applied Earth Observations and Remote Sensing	2008	202	3	1.508	10
International Journal of Remote Sensing	1980	173	3	0.833	9
ISPRS Journal of Photogrammetry and Remote Sensing	1989	318	2	2.979	7

The bibliometric analysis yields valuable insight into the scientific community who employ hyperspectral sensing and imaging tools and techniques in their research. First, the number of countries involved in hyperspectral research is substantially limited compared to the wider field of enquiry. Like the wider field, China and the USA have been the major contributors to scientific output for the period 1998-2018. However, China's influence in hyperspectral research is enormous, accounting for over a third of global research publications (1998-2018), compared to the USA's 18%, Table 7. Importantly, temporal analysis also reveals a widening research gap between China and many other countries. For example, in the five years to 2018 China has produced over 46% of peer-reviewed articles, compared to the USA's 10% and Germany's 9%.

Table 7 *Leading first author of all PA articles, subfield of HSI, top six countries*

Country	**First authorships (npubs)**	**Citations**
Canada	19	3,760
Germany	32	2,262
China	101	1,931
United States of America	55	1,807
Spain	11	519
Italy	10	289

---

[8] From 2005, "Transactions of the ASAE" changed its name to "Transactions of the ASABE".

Table 7 also reveals that while Canada and Germany are relatively small players in the global PA community, their researchers are significant contributors to HSI research. Authors from Canada have easily achieved the most citations for HSI papers in PA despite only publishing 19 papers. Canadian researchers, largely representing Agriculture and Agri-Food Canada, and McGill University, pioneered the use of HSI tools in PA experiments in the early 2000s and produced some of the most influential early papers in the field. Canadian author Dr Driss Haboudane from Agriculture and Agri-Food Canada has achieved an 2,548 citations from his two first authorships produced while at the University of Quebec, making him easily the most cited author in the HSI subfield (Haboudane, Miller, Pattey, Zarco-Tejada, & Strachan, 2004; Haboudane, Miller, Tremblay, Zarco-Tejada, & Dextraze, 2002). Dr Chaoyang Wu (China's Institute of Geographic Sciences and Natural Resources) is the next most-cited author, with 383 citations for his 2008 paper relating to the estimation of chlorophyll content using hyperspectral vegetation indices (C. Wu, Niu, Tang, & Huang, 2008), which he produced while at the University of Toronto. The most prolific first author of HSI indexed articles is the USDA's Dr Chenghai Yang, who has authored seven articles and is co-author on a further three scholarly articles.

An examination of the dataset reveals scant collaboration between different countries in HSI publishing to date, although there is evidence of collaboration between research institutions within countries. China's research effort is led by the Key laboratories/ Chinese Academies of Sciences (*n*pubs = 18) and an array of Universities including Zhejiang University (*n*pubs = 11) and China Agricultural University (*n*pubs = 8). In 2018, Chinese researchers authored over half of all HSI articles in PA (first author), firmly placing Chinese researchers at the forefront of hyperspectral research. Contributions from German researchers have mainly centred on the University of Bonn, who have the highest article output of any university worldwide despite only entering the field in 2008.

Overall, universities play a key role in HSI research, with authors from universities publishing 213 articles and authors from government research organisations producing seventy-one. The leading universities are the previously mentioned University of Bonn (Germany) and Zhejiang University (China), with researchers from both having produced eleven indexed articles. McGill University is third with six articles, although the most recent publication was in 2006, suggesting their researchers are not actively researching HSI technologies today. Government research organisations such as China's Key Laboratories and Chinese Academies of Science (*n*pubs = 20) and the USA's United States Department of Agriculture (USDA) (*n*pubs = 14) are also important centres for PA research, especially in the HSI area.

Technology-related bibliometric indicators

A key research question posed in this thesis is "What are the characteristics of new precision agriculture technologies?" To provide a robust response to this question, the bibliometric dataset was examined to identify shifts in the use of data-intensive technologies and methods (M. Z. Cushnahan, Wood, Yule, & Wilson, 2016). Each record in the bibliometric dataset was coded to the technologies and crops employed in the research, a time-consuming but necessary task to discover trends in the technologies employed by researchers and the crops to which they apply their techniques. These data were scrutinised and reveal a broad trend towards utilising remote sensing technologies and techniques.

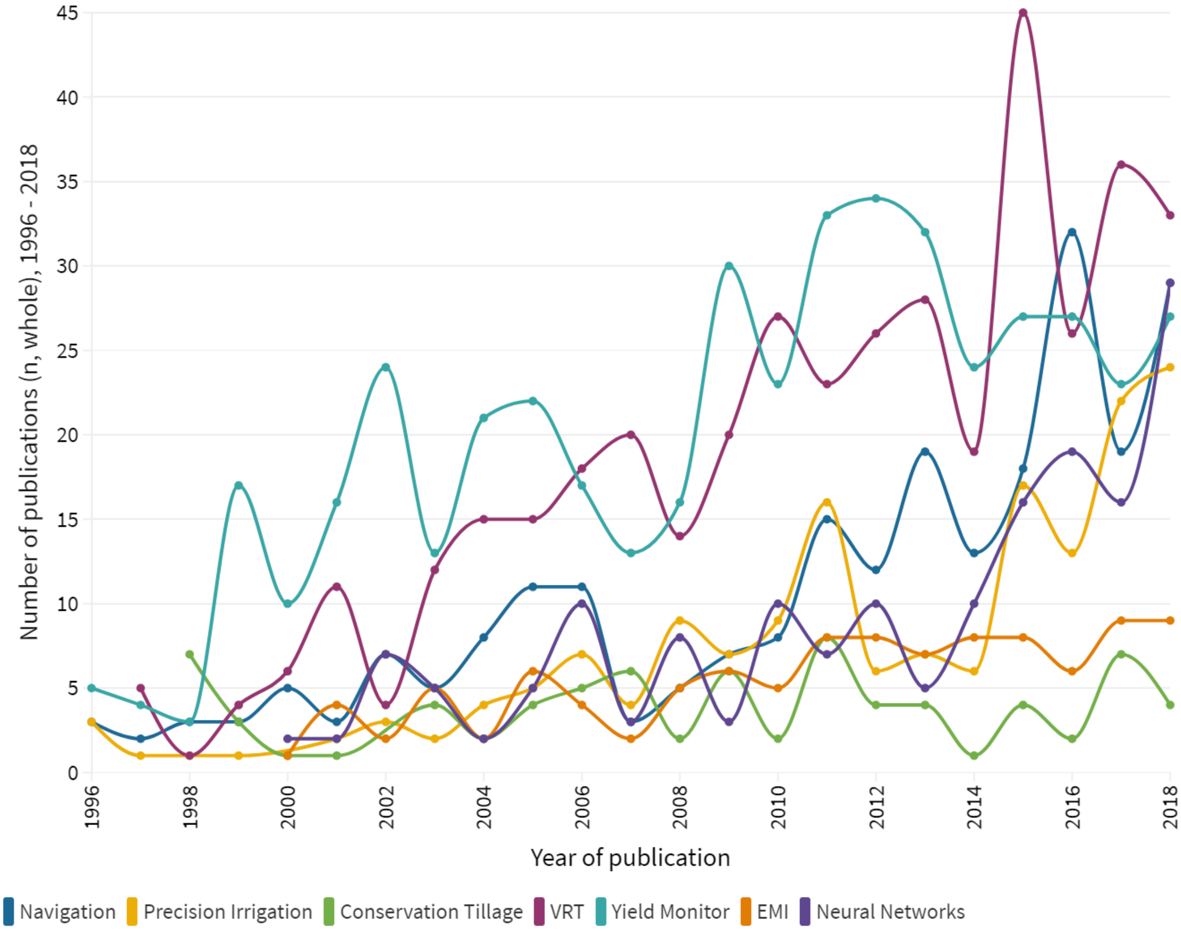


Figure 5 *Publication trends of early (first generation) PA technologies 1996-2018*

Importantly, this analysis supports the notion that few inherently 'new' technologies are emerging from contemporary PA laboratories. Many researchers of the 1990s and early 2000s focused on developing hardware, often known as first generation technologies, such as yield monitors, no-tillage machinery, and variable rate application technologies (VRT; see Figure 5). However, the emphasis of much recent research has moved away from inventing new technologies. Instead, the examination of scientific papers in the dataset reveals a strong reliance on technologies developed in other sectors. Fundamentally, it seems that the role of PA researchers has moved to assembling and adapting a community of technologies. The tinkering can take different forms. The first involves adapting existing technologies to new agricultural contexts. A second common approach is to employ an existing device to collect data – in this case it is how data are processed and the statistical treatments applied that creates novelty. Indeed, while the field appears to be technology based, PA research frequently simply involves applying existing data treatments from existing technologies to a new crop. This data-centric approach provides fertile ground for researchers; the almost endless variety of crops available means that there is always an opportunity to create 'novel' research.

Figure 5 shows that while early researchers often employed yield monitors and vehicle navigation in crop-based research, researchers are increasingly utilising data-rich technologies such as remote sensing technologies, machine vision and robotics. Importantly, these technologies are often more data-rich than the direct measurement methods traditionally used in agricultural research, such as yield monitors and biomass measurements. This shift could be motivated by a number of factors, including reduced research budgets after the global financial crisis (Wolinsky, 2009), improvements in remote sensing technologies, rapid advances in data processing and storage, and the pressure on researchers to frequently produce novel publications to maintain academic standing. The later ethnographic chapters of this thesis provide richer theorisation and empirical analysis of some of the challenges and opportunities borne out of the shift to data-centric PA research. However, first I will discuss the rise of data-intensive technologies in more detail.

The rise of data-intensive technologies in precision agriculture

The bibliometric analyses confirm the recent findings of Kamilaris et al. (2017) that there has been a rapid expansion in the use of data-intensive technologies and the analytical techniques employed to interrogate the data they generate. A point of interest is how quickly data-rich technologies have grown. Figure 6 shows the explosion in the data-intensive technologies in PA research in the five years to 2018. Much of this growth was driven by Chinese researchers studying emerging remote sensing technologies that were facilitated by improvements in transistor technology, data storage, data process and in the case of UAVs, the advent of lightweight batteries and GPS controllers (Cracknell, 2018).

The data also confirm that the research communities are extending their borders beyond the laboratory, with the increasing field use of data-intensive real-time and near-real time technologies previously confined to laboratory settings. These technologies produce massive volumes of data at very high velocity, with data arriving early in the science process. While researchers may have previously waited weeks to get soil testing data back from experiments, data collection using the remote sensing technologies is almost instantaneous.

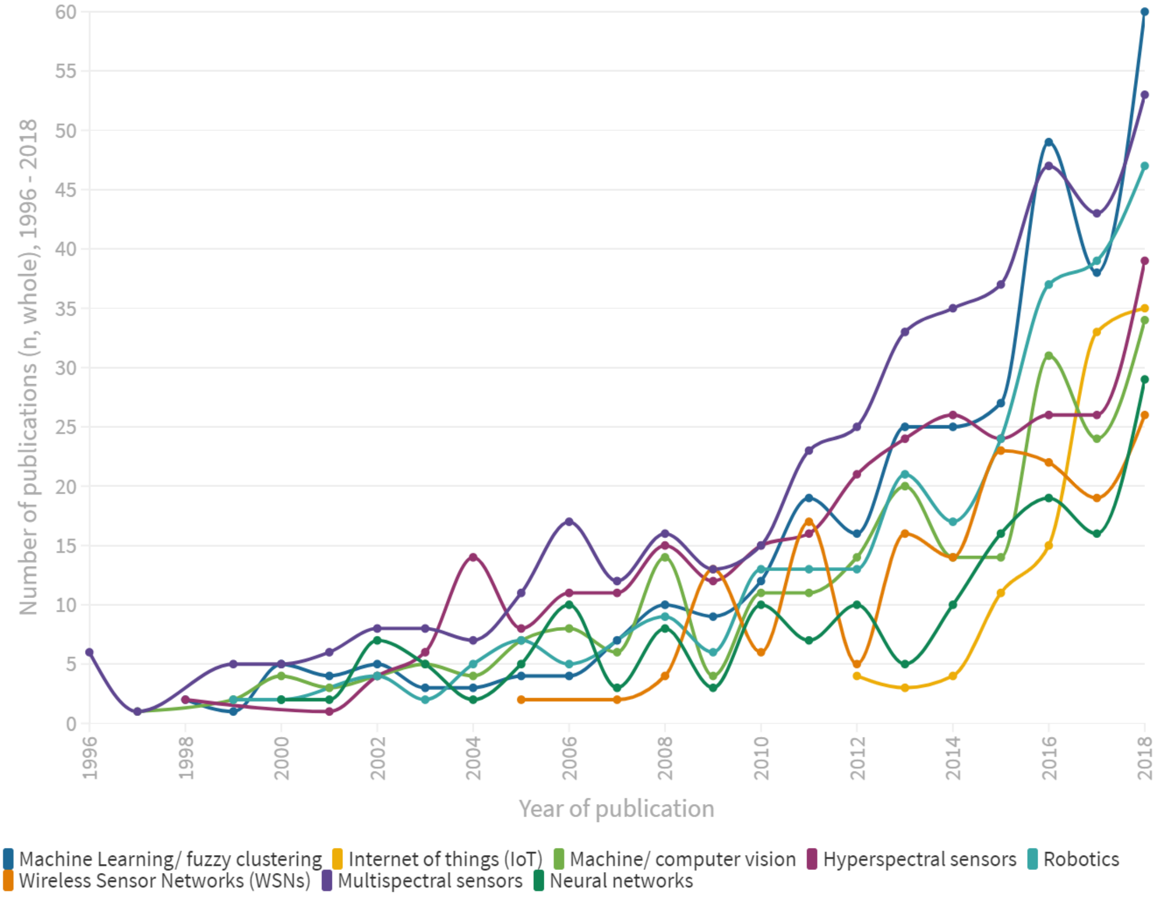


Figure 6 *Number of data-intensive technologies in PA publications, 1996-2018*

While not all data-rich tools meet the criteria of big data; volume, variety, velocity, veracity (Kitchin & McCardle, 2014), nascent advances in machine learning and near real-time hyperspectral imaging sensors such as AisaFENIX do. The results reveal the burgeoning popularity of these technologies in PA scholarship, which supports the assertion that big data have arrived. The implications of this are discussed in the following sections and chapters. Importantly, the data also signpost a shift from the dominance of direct measurement through yield mapping and wet chemistry soil analyses to the data-rich proxies of remote sensing and machine vision technologies in the past decade. Advanced sensing technologies such as multispectral and hyperspectral sensors, LiDAR, EMI technologies are increasingly becoming the focus of scientific publications, rising from seventeen (combined total) in 2000 to 153 in 2018.[9]

---

[9] Publications may be coded to more than one technology.

A key publishing pattern observed in the data is the focus of publications over the life cycle of a technology. Deeper analysis of the technology-related publications reveals that early research publications often focus on the development and testing of a new technology, followed by assessments of accuracy, then economic analysis, followed by use of the technology on different crops, and finally articles on the uptake rates of the technology by farmers. Generally, the number of publications then drops off as the technology becomes established, such as is the case with EMI technologies. Publication numbers on EMI technologies were fairly stable in the early 2000s but have declined in recent years. A similar trend may be emerging with wireless sensor networks (WSNs) as the publication numbers decreased in 2016, although their number rebounded in 2018, see Figure 7. These patterns may be temporary; it is possible that EMI and WSNs are less appealing as they become mainstream. The results suggest that researchers interested in studying emerging technologies may be switching to other data-intensive technologies such as robotics and machine vision, which have seen increased publications over a similar timeframe. An interesting anomaly to this pattern seems to be navigation technologies, which are making a comeback with the emergence of robotics.

In terms of data collection platforms, researchers traditionally collected ground-based (proximal) field measurements, although aerial photography was also used (Burger, 1996; Flowers, Weisz, & Heiniger, 2000; Franzen, Reitmeier, Giles, & Cattanach, 1999). An aim of this study is to gain insight into the characteristics of data collected by researchers and to examine how PA data have evolved as researchers capture and utilise data from different platforms. Of note is the spike in proximal remote sensing since 2010, which on closer examination appears to be their use as ground-reference data for other remote sensing technologies.

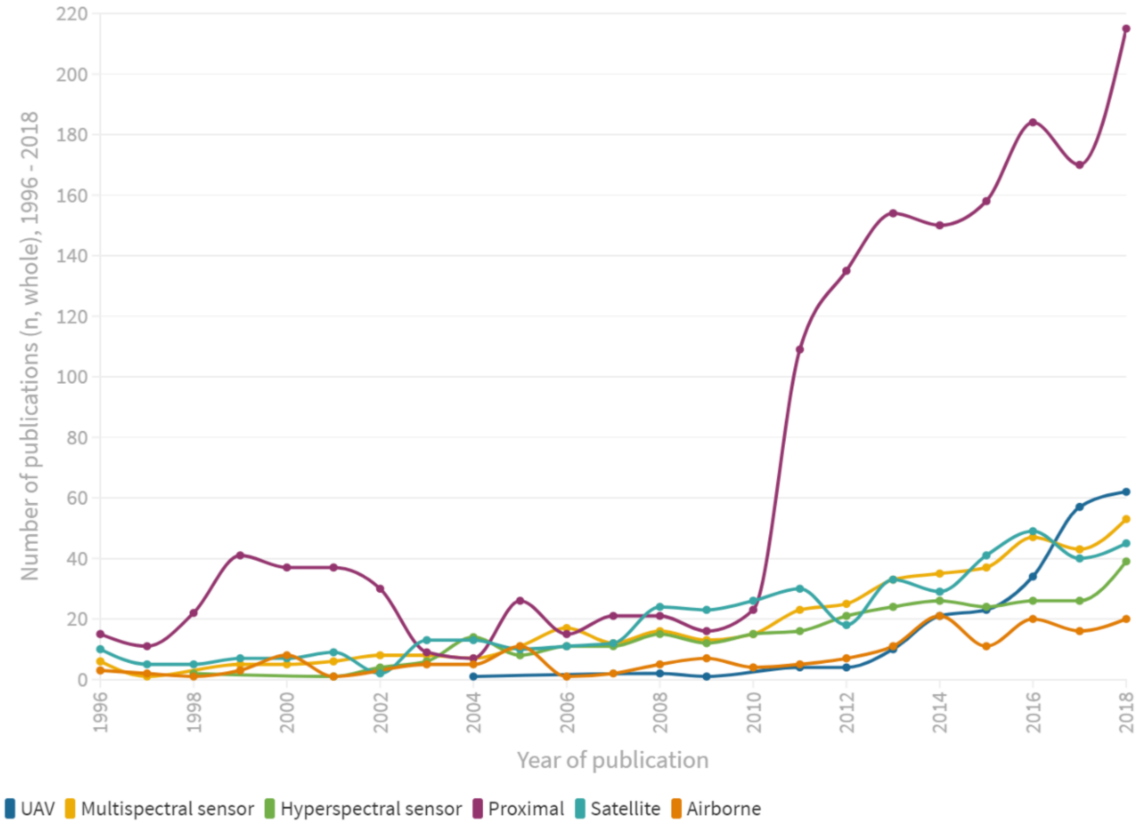


Figure 7 *Publication trends in remote sensing data collection platforms, 1996-2018*

The analysis also marks the growing use of satellites since the early 2000s. It also reveals that the use of sensors mounted on UAVs rocketed from 2013, largely driven by the rapidly falling cost of UAV units (Matese et al., 2015; Toth & Jóźków, 2016). Interestingly, proximal data collection has spiked since 2010, reflecting a trend towards data assemblages that employ ground-reference data to build relationships with aerially acquired data. Unlike all the other remote sensing platforms that have become more widely used in recent years, the use of airborne sensing platforms remains low, likely because the capital cost of entry remains high (Toth & Jóźków, 2016).

Key findings: an insight into contemporary precision agriculture data

The findings of this chapter deepen our understanding of the global distribution of PA research, which is rapidly changing in terms of who and where research is being performed. The key conclusion from the data is that PA research is performed by an ever-widening pool of academics from a broadening sphere of expertise. Once dominated by researchers from the USA and the crops grown there, the bibliometric indicators suggest that PA is now a truly global field of study. The results may suggest a shift in researchers' collaboration practices or signpost an overall decline in influence of American researchers.

Another key observation is the apparent evolution of the scientific division of labour within scientific communities. Precision agriculture research is increasingly published in domains beyond the traditional realm of the field, particularly into optical engineering, remote sensing, data processing and data science. This may be linked to the arrival of the data-intensive technologies observed in the study, especially the advanced remote sensing technologies and the analytical approaches needed to refine value from the data produced.

The analysis uncovers a developing sphere of PA application. I discovered how the data ecosystem of PA has changed due to the popularity of certain technologies being used by PA researchers. The arrival of new data-rich technologies and advances in data analytics have prompted debate around what these may mean for PA research and development (M. Z. Cushnahan et al., 2017; Mintert et al., 2016). By comparing the technologies and techniques employed by researchers over time, I conclude that many modern current PA researchers are working in a completely different data economy to their predecessors. Importantly, this work shows that data shifts are occurring beyond the arrival of big data, with the emergence of a complex new data ecosystem for PA research. This includes the new role that small data play in PA research to both serve as a calibration tool for big data tools, but also new analytical techniques create the possibility of aggregating small data over time and space, making them valuable beyond their initial purpose.

The dataset also revealed trends in the data characteristics of the enabling technologies employed by researchers. I discovered that while data-rich technologies such as hyperspectral imaging sensors and machine learning existed in 1996, the bibliometric indicators show they were not widely used for PA research then. The evidence presented signposts an important trend towards these technologies in the five years to 2018. The results suggest many researchers are refocusing their scientific effort on indirect measurement using predictive models. Consequently, the scientific division of labour may be shifting from those employed in domain expertise to data analytics. To benefit from these advances, researchers and funders working in the field should understand the possible consequences of the new data economy on their science processes and the scientific division of labour required. The results of this study will help those engaged in PA research to better prepare for the opportunities and challenges ahead.

Much of the commentary on the current state of PA focuses on the arrival of big data technologies such as hyperspectral sensors (N. Rao, 2017), but most fail to recognise that small data still play a major role in research.[10] The results of this study confirm there have been dramatic changes in the technical characteristics and role of data in PA research. They suggest that to produce a deeper, clearer view of our world researchers will need to employ a range of data analytics that can accommodate assemblages of traditional (small) data sources and big data sources. The possibility of processing and analysing data with dramatically different characteristics could have significant ramifications for research design, the scientific division of labour, data analysis and refining value for end users of the research. For example, in the past decade, there have been more studies utilising both ground-based and airborne measurement, which adds challenges in terms of data fusion and the increased importance of highly accurate geolocation of sampling sites. This is not surprising given that a key challenge for PA researchers using remote sensing is the difficulty of mapping crops with accuracy using only high spatial resolution or only high spectral resolution remote sensing data (M. Wu et al., 2017). Analysis of the publication data indicates that proximal direct measurement or remote sensing is often used to provide data of high spatial resolution to verify and validate measurements and algorithms from other remote sensing platforms. This is most commonly performed by using proximal sampling to validate aerially acquired data such as the 2008 study by N. R. Rao (2008). However, recent research by M. Wu et al. (2017) combines hyperspectral and high-resolution remote sensing data from satellite and aerially acquired data respectively, to improve the overall accuracy of their maps.

In addition, there is evidence of researchers using hyperspectral sensors to discriminate spectral bands of interest, which can then later be applied in the field using cheaper multispectral sensors (Bagheri, 2017). This advance is important as the capital outlay required for hyperspectral imaging sensors is cost-prohibitive in terms of commercialising services for on-farm applications.

Besides cost barriers, many researchers use proximal data collection platforms to avoid confounding variables associated with aerially acquired hyperspectral imagery, such as the light exposure and atmospheric conditions, and the yaw of the data collection device. Some researchers describe their ground-based research as a step to transitioning to aerial, UAV or satellite data collection platforms in the future including Gao et al. (2016), Goldshleger, Chudnovsky, and Ben-Binyamin (2013) and Li, Mistele, Hu, Chen, and Schmidhalter (2014).

The progression in airborne remote sensing research is highly relevant to this study as data collected from different platforms brings inherent data processing challenges, such as the need for atmospheric corrections and image stitching with airborne platforms. Other data-related

---

[10] In this thesis I employ Kitchin and Lauriault (2015)'s description of small data, which states that small data are characterised by their generally limited volume, non-continuous collection, narrow variety, and are usually generated to answer specific questions.

challenges still to be overcome to make the aerial acquisition of hyperspectral data attractive include addressing limitations on data collection due to cloudiness and the restricted/reduced periods when data can be collected due to the angle of the sun (Adão et al., 2017). Like other big data technologies, the data density of these technologies creates both new opportunities and challenges for science-making teams (Pullanagari et al., 2012; Von Bueren & Yule, 2013). Importantly, different scenarios of big data call on tools and methods that can be drastically different at times. M. Z. Cushnahan et al. (2016) call this dramatic change in data characteristics, a new *'data economy'*. These data-rich technologies are positioned at the centre of the *Pioneering to Precision* case study investigated in the latter part of this thesis.

The growing influence of remote sensing technologies and data analytics also has implications for the scientific division of labour within science teams. However, while the bibliometric analysis sheds little light on how PA research is changing, it also shows how it is not changing. The lack of farmer focus is of concern. According to Crane (2014) there has been a strong emphasis on farmers at the beginning and end of the research process, but little analysis of the farmer and researcher interaction and the effect it may have on the outcome of the science process. The disconnect between researchers and the presumed end-users, farmers, is evidenced by titles or abstracts for only eleven percent of the indexed articles even mentioning farmers or smallholders. This suggests that some research may be performed solely with an academic audience in mind. Furthermore, while many publications in the database mention the poor uptake of PA technology, in 2018, fewer than 0.1% of papers in the dataset, focus on technology uptake. Publications that offer strategies or approaches to manage research with changing data ecosystems are rare. Given precision agriculture's accepted failures in realising widespread technology uptake, it is surprising there is so little scholarship in this field.

A close examination of the bibliometric dataset establishes that most studies focus on technical innovations, neglecting end-users and those creating the science behind them. This supports Crane (2014), who asserts that a significant knowledge gap exists in understanding how science-making teams within PA research operate. The limited literature investigating the possible impact of emerging data assemblages on PA research includes the work by M. Z. Cushnahan et al. (2017), who propose that a new scientific approach may be needed to efficiently refine value from big data PA projects. Similarly, Tremblay (2017) argues that PA must advance beyond the methods and statistical principles elaborated by Fisher, which have traditionally served researchers well. Tremblay asserts that contemporary PA research is being inhibited by strict adherence to Fisher's method, which has been followed so extensively that it is now largely defining the problems. Tremblay recommends that contemporary experimental design and analytical techniques should be capable of learning from machine and sensor data, that is, they must rely upon observation farm production data along with data from controlled experiments (Tremblay, 2017). These findings raise questions about how the arrival of big data technologies will impact science-making teams and the methods they employ to refine value from PA data.

Conclusion

A question central to this study is “How do big data influence science-making in precision agriculture?” This chapter evidences a growing body of knowledge published under the banner of PA. The bibliometric analysis also reveals that big data technologies have indeed arrived in PA and we find that emerging technologies such as hyperspectral sensing and imaging are being increasingly employed by PA researchers. The bibliometric analysis shows that researchers are increasingly working with complex assemblages of technologies and data. In a somewhat unexpected finding, the bibliometric analysis identified a recent trend toward ‘small data’ technologies, which appear to be employed to validate data from nascent technologies.

While the bibliometric analysis has been useful in identifying some key trends in precision agriculture research and the tools employed by researchers, it clarifies little about what is written about *science-making* in precision agriculture. Thus, while the bibliometric analysis is useful in terms of defining how PA research is growing and changing, it provides no sense of the internal dynamics within science teams nor the science method used to refine value from data. Hence, it is difficult to gain a deep understanding of the science-making journey of PA researchers, and the consequences and challenges that advances in technology and data analytics may bring. Furthermore, this research gap prevents us from crafting solutions to these challenges and exploiting these opportunities. Consequently, I needed to employ a different approach to advance our understanding of the influence that the arrival of these technological advances may have on science-making. In the next chapter, I take a more traditional approach to explore the literature relating to PA research.

Chapter 3: A brief history of the evolution of precision agriculture research

Introduction

In this chapter I chart the evolution of agricultural science into the modern field of PA to develop an understanding of the field's scientific practices and the data ecosystems they produce. I examine the methods and practices developed to create knowledge from small data and chart the evolution of technological assemblages that generate data more cheaply from an increasing range of farm environments. This ultimately raises ontological questions as to what is 'data' in this field, what defines it, where data starts, and noise begins.

In Chapter 2 I discovered that data-rich tools, such as hyperspectral sensing and imaging are being increasingly employed by PA researchers in their studies. However, until recently, most PA researchers came from backgrounds in agronomy, soil science, agricultural engineering, and geo-statistics. Experts in their own field, it is unlikely many of these academics would also have the specialist expertise required to process or analyse data from complex remote sensing data. Hence, the move to data-rich technologies and tools may signal a possible shift in the scientific division of labour within research teams to include more personnel with remote sensing and data analytics skills.

While Chapter 2 revealed an explosion in the volume of PA research since the 1990s, most studies focus on technical innovations with little reflection or attention placed on user uptake or the innovation processes behind them (Akpo, Crane, Vissoh, & Tossou, 2014; Crane, 2014; Sonka & IFAMR, 2014). In this chapter I attempt to situate contemporary PA research by exploring the evolution of science-making in the field both globally and within New Zealand.

The 'revolutionary years': agricultural progress between 1500 - 1850

As a sub-field of agricultural science, PA research traces its origins to rationalism and the Age of Enlightenment. There is general agreement that between 1500 and 1850, agrarian systems in the United Kingdom and Europe experienced a fundamental technological transformation from low-intensity agricultural systems based on fishing and fowling to a high-intensity system based on arable crops (B. M. Campbell & Overton, 1993).

When reviewing early historical commentary on scientific progress in agriculture, one could be forgiven for assuming those working on the land played no role in this transformation. While farmers have long been aware of the benefits of record-keeping and understanding of soil and crop input requirements (Whelan, 2011), there is little evidence of this in the literature until the 1980s. Indeed, site-specific farming approaches have been informally used for thousands of years

(Nawar, Corstanje, Halcro, Mulla, & Mouazen, 2017), with farmers varying management activities based on yield or soil conditions to improve the efficiency of agricultural production. Despite this, there is little literature on the role of farmers as agrarian innovators. Instead the heroes of progress were the gentry who recorded and published their observations of farming practices (Ernle & Prothero, 1936), and who occasionally applied an inductive reasoning approach to experimental trials. In the 1960s, historians first questioned the legitimacy of the 'British Agricultural Revolution', with some arguing that it is simply a myth (Kerridge, 1969). This scholarly debate was fuelled by the revelation that many technologies previously credited to revolution could be reliably traced to the thirteenth century (B. M. Campbell & Overton, 1993). Similarly, some historians suggest that this later period was a time of the refinement of technologies rather than a time of invention (E. L. Jones, 1965).

So how did this reimagination of historical events develop? It appears the early proponents of the agricultural revolution such as Ernle and Prothero (1936) and Toynbee and Milner (1894), relied entirely on literary sources for their views. This led to the role of elite actors being greatly exaggerated in their accounts (Kerridge, 1969). For example, Viscount Townshend (1674-1738) is widely credited with pioneering investigations of the improving effects of the four-crop rotation system (Timmer, 1969). However, these practices were really 'borrowed' from Dutch and Flemish Farmers (Klinkenborg, 2012). Nor did Townshend introduce the first turnips to Britain from Hanover as is often reported; records show they were already growing on his estate when he was a boy (Overton, 1996). Another acclaimed innovator of the day, Jethro Tull, did not invent the seed drill, rather he refined an existing technology and promoted its use amongst other landlords and farmers. Indeed, Overton (1996) credits much of the productivity gains of the day to the introduction of crop rotation systems, and importantly, the increased intensity of farming systems, supported by introduction of arable cropping onto the newly drained fenlands of Eastern England.

There is some criticism of the revisionist approach to agricultural history. Debate on the historical role of farmers in agricultural innovation took a dramatic turn in the 1990s, when historians proposed that there were actually two agricultural revolutions, the first led by farmers, and later a second revolution led by landlords (Allen, 1991). The acknowledgement of the first recognises that an agricultural revolution was well underway before 1750 in what historian Robert Allen describes as the 'yeoman's agricultural revolution' (Allen, 1992). Allen (1991) explains that most of the productivity growth was accomplished by small farmers in the open fields during the seventeenth century and that:

> … this revolution was marked by a doubling of corn yields; it raised England's national income, and the benefits were distributed widely. Small farmers who held their land on copyholds and beneficial leases gained as land values increased (p. 916).

Allen also recognises the ‘second revolution’, which aligns with the classic agricultural revolution (Dilley, 1991). Rather than being driven by the brilliance of aristocratic innovation and experiments, Allen (1992, p. 15) asserts that “‘the landlords’ revolution consisted of enclosure and farm amalgamation, and ultimately led to a dramatic increase in inequality in British society”. Importantly, our often-romanticised view of the ‘Agricultural Revolution’ fails to account for the hardships borne under the guise of technological utopianism.

So, what does this mean for the progress of European agriculture prior to the mid-nineteenth century? The influence of empirical experiments on agricultural progress during this period has probably been greatly exaggerated. Historical analyses of the field rely heavily on written works and scant records, which are unlikely to offer an accurate depiction of progress in farming in that period. Historians writing about farmers usually provide little insight in terms of what data looked like and how they were used, if at all, to reach conclusions and influence decision-making. However, Chalmers (2013) suggests there was a progressive transformation of both scientific theory and in what were considered to be observable facts during this period. Scientific investigation was dominated by the Cartesian paradigm of positivism (rationalism) at this time, which posits that a reality exists that is driven by immutable laws (Pretty, 1994).

What I cannot decipher from these historical scientific accounts is the influence of farmers on the innovation process and the role of farmers in the diffusion of technologies. Shapin (2018) contends that innovation relied on a diverse array of cultural practices aimed at understanding, explaining, and controlling the natural world, each with different characteristics and each experiencing different modes of change (Shapin, 2018). What is generally agreed is that agronomic research in the eighteenth century concentrated on the physical changes of enclosure and embodied the economic changes of increased intensity of cultivation with a view to increasing overall productivity (F. M. L. Thompson, 1968) and reducing labour inputs (Allen, 1992; Ernle & Prothero, 1936). A notable drawback of this progress were the negative social impacts borne of increased labour efficiency, such as the depopulation of rural areas (Allen, 1992; Ernle & Prothero, 1936). What this work demonstrates is that change in agricultural technology has wider social reverberations. This raises considerable legal and socio-ethical questions. However, few studies explore such consequences in PA circles. Only recently have these issues started to receive the attention they deserve through works by Carolan (2017), Eastwood et al. (2019) and Kritikos (2017).

Progress of site-specific soil management in the 18th and 19th Century

Many early British and European agricultural experiments followed the naïve-inductive approach originally articulated by the eminent scholar and Lord Chancellor of England Francis Bacon (1561-1626) (Gower, 2012). Viscount Bacon’s hypothesis-driven experimental designs were highly influential, which involved the construction of scenarios and the collection of accumulated observations in the hope of distilling an answer (Whelan & McBratney, 2000). Like today, researchers of the day sought to discover, predict, and control natural phenomena (Gower, 2012; Pretty, 1994).

Many researchers continued to apply the observation-led naïve inductivist approach through the latter half of the eighteenth century and early 19th century. As described by Nola and Irzik (2005), the naïve inductivist approach involves gathering facts, often through observations independently of any theoretical considerations. These facts are then used to infer hypotheses, laws, or theories. Much of the focus of agricultural progress during this period concentrated on field observations to inform theories relating to crop nutrient management. For example, agriculturist Johann Friedrich Mayer, a Protestant pastor, conducted experiments in his garden and like his British counterparts observed farm work on nearby farms in an effort to improve the livelihoods of the Hohenlohe peasants in his parish (Mayer, 1768). Importantly, Mayer was a pioneer of research extension. He published his findings on the use of crushed gypsum as fertiliser in his 1768 book *Lehre vom Gyps als vorzueglich guten Dung zu allen Erd-Gewächsen auf Aeckern und Wiesen, Hopfen- und Weinbergen* (“Doctrine of Gypsum exquisitely as good manure to all natural plants to the fields and meadows, hops and vineyards”). In addition to a commentary on his experiments with gypsum, Mayer’s book contains testimonies from esteemed users and a question and answer section to support his work (Mayer, 1768). This is interesting because it is an early example of the use of testimonials, which did not become widespread until a century later (D. B. Jones & Tadajewski, 2016).

This work was closely followed by publications on crop rotation in 1769 (Mueller, 1977) and husbandry (Mayer, 1773). Mayer’s books were complemented by the regular publication of a journal of his findings, *Beyträge und Abhandlungen zur Aufnahme der Land- und Hauswirthschaft*, which he published until 1786. The experimental approach spread to France, when in 1783, Comte d'Angiviller, acting for Louis XVI of France, placed Abbé Tessier in charge of installing and conducting the first experimental research program at the farm attached to Château de Rambouillet (Gillispie, 2004).

Scientific method in agricultural science 1880-1950

Agricultural research stations emerged in the USA after farmers lobbied for research into accurate analyses for the fertiliser or "artificial manure" they employed in the 1880s (Marcus, 1985). The lobbying resulted in the passing of The Hatch Act of 1887, which provided funds for the type of scientific research that brought about an agricultural revolution in the USA (Hillison, 1996). In terms of scientific method, the emergence of these experimental stations signalled an important shift from the largely observational approach to quantitative methods.

In the early part of the twentieth century, agricultural scientists began to make progress on addressing problems with the validity of the naïve inductivist experimental approach by employing statistical methods relating to measuring the uncertainty of observations developed earlier by statisticians such as Bayes, Helmert and Laplace (Andersen & Hepburn, 2016; Fisher, 1925). These researchers preferred the reductionist approach, which involves breaking down components of a complex world into discrete parts, analysing them, and then making predictions about them based on interpretations of these parts. In 1900, Pearson refined the method to measure the discrepancy between observation and hypothesis, known as c^2 (Fisher, 1925).

Early quantitative research was often hampered by field heterogeneity. Scientists investigating plot yields in the 1920s found that variability across trial areas prevented the accurate assessment of the benefits of their treatments (Harris (1920) in Franzen and Mulla (2015)). These challenges fuelled the rapid advances in the statistical approaches employed by agricultural scientists, particularly the works of Sir Ronald Aylmer Fisher (1890–1962). Fisher first introduced the term ‘variance’ and proposed its formal analysis in his 1918 article *The Correlation Between Relatives on the Supposition of Mendelian Inheritance* (Fisher, 1918). Four years later, Fisher introduced the method of maximum likelihood in 1922 (Aldrich, 1997).

Fisher began his work on the foundation of experimental design at the Rothamsted Experiment Station in England in 1919 (Box, 1980). In the following seven years Fisher developed a series of statistical tools that would form the foundation for most small plot and even full field experiments (Franzen & Mulla, 2015). Fisher’s *The Design of Experiments* proved revolutionary for agricultural scientists. In this canonical publication, Fisher established principles of experimentation and statistical analysis that laid the foundation for agricultural science today (Fisher, 1937). Fisher’s approaches proved immediately popular with agricultural scientists and in 1929, Lindley and Bauer published the first known soil sampling recommendations for soil testing to address field heterogeneity (Franzen & Mulla, 2015). New Zealand’s agricultural scientists closely followed Fisher’s advances. In 1933, Bruce Levy (later Sir) and his colleague Ernie Madden published *The Point Method of Pasture Analysis*, which employed Fisher’s analytical techniques to test their hypotheses (Levy & Madden, 1933). Levy is credited internationally with having invented the point analysis method of determining pasture species composition, now widely used in ecology research (Matthew, 2019).

Fisher’s methods, such as the now standard analysis of variance (ANOVA), supported the drive of scientists towards hypothesis-driven science processes and enabled agronomists to remove the effects of site variation from field experimental results (Cook & Bramley, 1998). The drive to control variables was of critical importance to researchers and the arrival of ANOVA was a key advance. An ANOVA involves partitioning the sources of variation of data and comparing the relative sizes of these components to decide which sources of variation are important and which can substantially influence the precision and value of the output of an ANOVA. However, field homogeneity remained and remains problematic as the Fisherian approach essentially ignores the spatial component of experimental and analytical error. Consequently, scientists have moved to perform their experiments in highly controlled conditions where variables could be ‘fairly’ tested and where the emphasis was on the improvement of future crop performance (McBratney et al., 2005).

According to Cook and Bramley (1998), using Fisher’s techniques meant that the research community had effectively removed localised variation from their consideration until the arrival of precision agriculture. Furthermore, employing Fisher’s techniques means that a good hypothesis could only be judged after experimentation, based on the low probability that observations conform to the null hypothesis (Cook & Bramley, 1998). While it may have seemed reasonable for researchers to adopt a pragmatic view of fields or paddocks as homogenous management units, the arrival of GPS technologies meant researchers had to reconsider the principles underpinning their method (Cook & Bramley, 1998).

The classical science model in the post-war era: a solution to the problem of data scarcity

Gathering data continued to be a problem for researchers in the post-World War Two period. The collection of agricultural research data remained relatively slow and expensive, with experiments judiciously planned to produce statistically significant results with as little data as possible for pre-determined problems (Kell & Oliver, 2004). Data generation was usually achieved through some form of measurement and often only sparse data was generated towards the end of the science-making process. Data were largely captured on-farm, with the aim of influencing on-farm decision-making.

Using this model, agricultural scientists tightly control data by employing sampling techniques that limit the scope, temporality, size, and variety to achieve a derived output within the scope of the problem statement. The scientific principle of parsimony is a key driver for knowledge creation in the classical hypothetical-deductive science model. Importantly, much of the scientific effort focused on experimental design, i.e. planning and performance of the data collection stage, rather than data analysis. This approach to science uses the bare minimum of samples (often limited by budgets) to achieve a derived output within the scope of the problem statement. Data are typically acquired or gathered in a batch, where all the needed data are available at once (Fokoué, 2015). Researchers employing the classical science model generally devote significant effort to capture and define the levels of error, bias, uncertainty, and provenance of their scarce data (Kitchin & Lauriault, 2015).

Historically, a great deal has been achieved with this approach and it is interesting to note the leading role played by agriculture in the development of the generalisation statistics that ensure a ‘robust’ move from sample to the application world (Norton, 1978). While much has been achieved using this model, it is not without its pitfalls. Small samples can be unreliable; large samples have been expensive (Huberty, 2015). Importantly, despite researchers’ best efforts, samples might not be representative. Moreover, the model is not very scalable, and researchers usually need a new sample to address every question. Indeed, researchers have had to divine in advance their research questions and hope that they proved adequate.

A key consequence of the classical science model is that researchers often create static snapshots of controlled factors, delivering information on a single property. Data between different plots, sites and times has typically not been commensurable, providing little opportunity to aggregate data with a high level of accuracy, although the advent of geo-statistics has helped researchers extrapolate data beyond the original collection site. Another important implication of this model is that scientific effort and expertise typically focussed at the early stages of the science process, namely experimental design, and data collection. The challenges and benefits of the classical science process are discussed in depth in Chapter 6.

Data journeys in the classical model of science

The classical science model generally follows a linear process that includes the following steps:

1. A research question is posed
2. Hypothesis (framing of research question/ problem and desired solution)
3. Materials and Method are selected
4. Experiment is enacted
5. Data are collected (sampling)
6. Analysis – of sample and generalise (e.g. Kriging) to application world (the farm)
7. Conclusion – Researchers specify a solution, then move to the next problem.

The controlled nature of data collection means that scientific effort is concentrated on experimental design and control of variables; data are expected late in the science process. Consequently, as Figure 8 demonstrates, even when data are large, the late arrival of the data means the data journey is typically confined to a simplified, linear process.

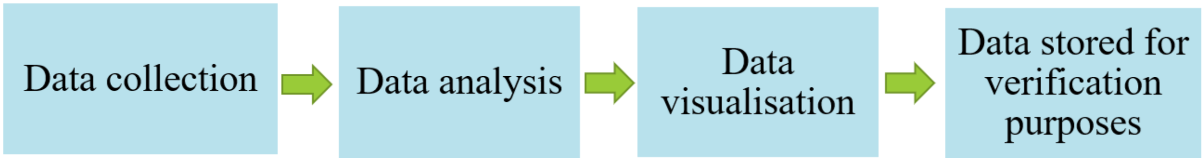


Figure 8 *Data journey in the classical science process*

Importantly, data in the classical science process is not usually designed to be repurposed. Consequently, little or no effort is paid to creating data structures that support future analyses of the data. This has consequences for replicability and verification of methods, which are detailed later in the thesis. Another important characteristic of this model is that data generated are typically immutable, based on direct measurements where little processing occurs prior to analysis and storage. For researchers this means there is usually little difference between the captured raw data and that which are analysed, making verification a simple and usually repeatable process.

The Green Revolution and the rise of geostatistical models

The success of the classical science model in modern agricultural science is defined by two key advances: The Green Revolution and the rise of geostatistical models. Contemporary PA research extension can be traced to the Green Revolution. Between the late 1950s and late 1960s, Norman Borlaug's Green Revolution generated a set of research technology transfer initiatives that transformed global agriculture. The Green Revolution is relevant to contemporary PA for three main reasons. First, it proved that rapid increases in crop yield could be achieved by applying fertiliser and chemical inputs. Secondly, it showed that farmers will adopt technological advances where benefit is demonstrated. Third, and arguably most importantly, while the livelihoods of many farmers were improved through substantial increases in output, there were *disastrous* environmental and societal consequences resulting from the increased mechanisation, widespread use of fertiliser, larger field and farm size, and the development of pesticides and herbicides (Clunies-Ross & Hildyard, 2013). Furthermore, the industrialisation of agriculture led to a dramatic reduction in labour input in agricultural operations that had a serious impact on rural communities. These impacts were widely justified by politicians, commentators and lobbyists as being necessary sacrifices provide food security for a rapidly growing population (Clunies-Ross & Hildyard, 2013). However, as early as the 1970s debate on whether some of these consequence were ethically defensible had begun (Paddock, 1970).

A key advance in the post-war period was the rise of geostatistical models. Contemporary PA is predicated on geo-referencing – assigning geographic coordinates to points on the earth's surface (Wolf & Wood, 1997). Of course, in the post-war era, the precise geolocation technologies farmers and researchers rely on today were unimaginable. However, the origins of many modern PA techniques and technologies lie in the post-World War Two period, when much progress was made in accounting for field heterogeneity using geo-statistics. During this time, researchers searched for solutions to the pressing problem of how to predict soil and plant characteristics from sparse ground data (Franzen & Mulla, 2015; M. A. Oliver et al., 2013). Finally, the 1960s saw the emergence of a new subfield of statistics, geo-statistics, which largely focussed on responding to the spatial variability of soil nutrients by extrapolating wet chemistry results across a field or farm using theoretical models (Franzen & Mulla, 2015; M. A. Oliver et al., 2013). Importantly for statisticians of the time, geo-statistics extended the field beyond the limited realm of traditional statistics.

New statistical tools soon emerged, notably the introduction of a geostatistical method known as ‘Kriging’, which was introduced to PA by the Canadian researcher Matheron. Matheron (1963) expanded on the empirical ideas of the pioneering geo-statistician Danie G. Krige to advance the notion that neighbouring samples could be used to improve prediction, and put them into the theoretical framework of the regionalised variable theory that underpins geo-statistics today (Franzen & Mulla, 2015; M. A. Oliver et al., 2013).

Geo-statistics proved popular in PA research for several reasons. First, Kriging provided researchers with a method to estimate or “interpolate” values from unsampled areas. This allows areas to be mapped with small sample sets; in some cases, a minimum sample set of only thirty observations would suffice (Franzen & Mulla, 2015; Gotway, Ferguson, Hergert, & Peterson, 1996). Secondly, interpolated values are unbiased. According to Marchant, Daley, and Webster (2013), they have minimum variance and that variance can also be estimated and so express the uncertainty of the interpolated predictions. So, while interpolation is very precise, it is often inaccurate due to the high level of spatial variability in many situations. Yet, interpolation is very appealing to researchers with limited means of direct measurement as it is better than nothing.

While the requisite statistical analyses and theory were developed in the 1950s (Gauch, 2012), these methods were not successfully applied by researchers to model crop yield until the 1980s (J. W. Jones et al., 2017). J. W. Jones et al. (2017) argue there were several reasons for the gap. First, the limitations of computer processing and storage made data capture and analyses expensive. Secondly, the authors suggest that while farmers may have wanted this information, it was only when these developments were eventually fuelled by various needs at national and international levels that innovation in this area occurred. They suggest that these advances were also supported by innovations in modelling approaches by the agricultural economics community.

However, Gauch, a pioneer in crop modelling, rejects the notion that limitations in electronic processing hindered developments in modelling crop yields until the early 1980s (Gauch, 2012).

Instead, he suggests the key barrier to progression was not computer processing, but deficient knowledge of the principles of scientific method, in particular the principle of parsimony (Gauch, 2012). Gauch (2012) posits that his knowledge of parsimony led him to believe that statistical modelling could increase accuracy without collecting more data. It does appear, however, that progress in agricultural data modelling was supported at least by the development of advanced application technologies such as variable-rate application technologies (Roberts, 2007; Schueller & Bae, 1987), and the economic analyses that accompanied them.

It is likely the arrival of variable rate application technologies were key drivers of the demand for models to inform application decisions. Indeed, the extension of Fisherian methods arising from the arrival of geo-statistics undoubtedly supported the parsimonious approach to data acquisition by drawing relationships from small data collected in the field. Furthermore, while the advent of geo-statistics opened the doors of the relatively well-funded agricultural research world to statisticians, the scientific method, and division of labour within the field largely remained the same.

The arrival of precision agriculture

The field of PA soon emerged from these advances in crop modelling. This progress was led by American researchers investigating large-scale, relatively homogenous monocultures. The developments, such as geostatistical models, were primarily new techniques to account for incomplete or low-resolution data. On close inspection, researchers generally continued with the science-making structures inherited from their predecessors. The data economy, composition of science teams, and scientific method mostly remained intact in what some researchers described as a new research paradigm for soil science (Hudson, 1992). It was still strongly influenced by the concepts of parsimony and efficiency, as data were still sparse and arrived late in the science process (Gauch, 2012). Those employed in theoretical PA research required similar skills to empirical research, save for the increased influence of geo-statisticians needed to generate and interpret the models (M. A. Oliver et al., 2013).

Responding to variability

Farmers have always known their fields were variable. However, without tools to measure or address this variability, variability was usually considered ‘noise’ (Cook & Bramley, 1998). Because farmers generally managed on the basis of homogeneity (Bramley, 2009), the scientific community focussed on improving agricultural production through the introduction of blanket applications of pesticides, herbicides and fertiliser to improve productivity. The modern era of PA was a response to the academic recognition that land is variable, but the key catalyst was the rapid advances in electronics, which permitted enabling technologies such as global positioning systems (GPS), geographical information systems (GIS), crop yield monitors, remote sensing and variable-rate application technology to emerge (Bramley, 2009; Roberts, 2007; Schueller & Bae, 1987). Importantly, these technologies also supported the advancement of the industrialisation of agriculture, with goals of reducing labour inputs, centralising intelligence, increasing yields in monocultures, and supporting large-scale mechanisation of agricultural operations. This is possibly a key reason that academics working in the field seem to have been so well resourced in recent decades compared to other agricultural sciences.

While many initial studies focused on yield monitoring (Gauch, 1988; J. Lamb et al., 1995; Schueller & Bae, 1987) and variable rate applications (Fleming, Westfall, Wiens, & Brodahl, 2000; Mulla, 1993; Mulla, Bhatti, Hammond, & Benson, 1992), over time the PA concept has come to represent an extension of industrial farming into farmer-less agriculture, with an increased focus on autonomous vehicles, centralised intelligence and wireless sensor networks controlling operations such as irrigation, fertigation and animal feeding regimes (Franzen & Mulla, 2015).

The data economy of agricultural research became more complex in 1983 when the US Military’s Global Satellite Navigation System, also known as GPS, became accessible for civilian use. In 1990, fearing adversaries might use GPS to their advantage, the US Military purposefully introduced random signal distortion (Sullivan, 2012), also known as Selective Availability (SA). Random signal distortion inflated the confidence interval substantially, so researchers had to employ their own expensive decryption hardware and subscribe to the services of a Differential Global Positioning System (DGPS) signal service provider, which could filter out roughly ninety-five percent of the error (Wolf & Wood, 1997).

Despite the challenges in responding to SA, the new GPS technology facilitated the geolocation of agricultural data beyond the use of stereo-plotters (Larsen, Nielsen, & Tyler, 1994; Young & Carter, 2013). This meant that yields could be recorded automatically with GPS recorders (M. A. Oliver et al., 2013; Whelan, 2011). Yield monitors were the first novel feature of PA and remain the most widely used and researched technology in this class today (Cook & Bramley, 1998). While yield monitoring for site-specific management was being performed prior to the advent of GPS systems (Gauch, 1988); data collection was slow and expensive. Moreover, except under research conditions, the geolocation of data points was imprecise, with data reference points often measured from fixed landscape features.

Another benefit of accurate geo-location of agricultural data meant that for the first time, farmers and agricultural scientists could reliably return to a precise position to gather data year-on-year. In the 1990s, the uptake of GPS technologies blossomed into applications previously unimaginable such as livestock monitoring and variable rate fertiliser application. However, the cost of DGPS was still a barrier for many practical applications (Adrados, Girard, Gendner, & Janeau, 2002). Selective availability was deactivated in May 2000 (Adrados et al., 2002), improving accuracy tenfold overnight (Sullivan, 2012), and generally removing the need for DGPS for ground-based operations. This free, accurate technology unlocked vast opportunities for researchers in precision agriculture.

The arrival of GPS-enabled yield monitors meant that yield data could be grouped into management zones. This allowed farmers to variably manage their crop inputs, such as seed rates and fertiliser applications (M. A. Oliver et al., 2013). Uptake of these technologies was slow in the early years of PA for many reasons (Swinton & Lowenberg-DeBoer, 2001). While the expense of the early units and their poor reliability were undeniably factors, the absence of efficient and effective variable-rate application technologies until the late 1990s was also likely to have influenced farmer uptake of the technologies in the early years (Griffin et al., 2004).

The first yield maps produced by Searcy, Schueller, Bae, Borgelt, and Stout (1989) showed the effect of compaction from farm machinery on yield (M. A. Oliver, 2010). However, unless coupled with improved decision-making the technology was redundant (Lowenberg-DeBoer & Boehlje, 1996). For example, early researchers such as Cook and Bramley (1998). recognised that while yield maps intensify observations, benefit only follows uptake and application by farmers. This requires interpretation and evaluation by the decision maker. Thus, the two distinct branches of PA research emerged: one developing process control technologies, and the second focusing effort on developing decision support tools.

The first wave of PA research largely focussed on machine and process control technologies, including equipment capable of varying application at different scales such as the footprint of a one-irrigation sprinkler or a fertiliser top-dressing aircraft (Hedley, 2015). Early precision sprayers, such as the early model C-DAX sprayer, emerged during the mid-1990s. The development of these technologies was predicated on the advances in GPS, so it was not until SA was removed that PA really took off. Most early process control technologies were simply an extension of existing enabling technologies developed for other applications. These developments include improvements in steering, fertiliser placement, seeding, and focused on measuring crop outputs, such as yield, and on positioning farm machinery such as tractors, fertiliser spreaders and spray units more accurately in the farm landscape (Bramley & Trengove, 2013; McBratney et al., 2005; M. A. Oliver et al., 2013; Robertson et al., 2012; Whelan, 2011).

The second branch of modern era research involved the study of technologies that supported farmer's decision-making, often called decision support tools or 'information intensive' tools. Like research on process control technologies, research on decision support solutions rapidly expanded after the release of GPS systems. The removal of SA allowed researchers to accurately determine and record the position of an appropriate receiver continuously (Neményi, Mesterhazi, Pecze, & Stépán, 2003). This advance meant that for the first time geo-referenced yield measurements could be automatically recorded (M. A. Oliver et al., 2013; Whelan, 2011). Automatically geo-referenced measurements allowed farmers to map inherent field spatial variability, including factors affecting production such as drainage, landscape effects, soil structure, texture, bulk density, and nutrient status (Corwin, 2013).

Also called 'Information-intensive' technologies, decision support tools refer to strategies and technologies used to influence farmers' decision-making. The solutions rely on farm and field level data to influence farmers' decisions about input application and cropping practices. Unlike the process control technologies, these tools are usually focused on data and information. Much of the decision support research has centred on measuring crop outputs such as yield, and technologies are designed to inform decision-making to facilitate the site-specific application of treatments (M. A. Oliver et al., 2013). These maps allowed farmers to manage field crops spatially, optimising the use of inputs such as fertiliser, plant population and pesticides according to field location (M. A. Oliver et al., 2013; Robertson et al., 2007; Schueller & Bae, 1987). Importantly, the adjustments to the management and control of inputs were predominantly directed toward improving *future* crops rather than on-the-go management changes, except for variable rate fertiliser applications (Robert & Stafford, 1999). This is important because the farmers had to wait until their next crop to reap the rewards of the technology, possibly risking production losses in the meantime.

Nevertheless, the early soil nutrient and yield maps were a major advance and they proved popular with arable farmers in the USA. By the late 1990s, approximately 33–45% of PA retailers in the USA were offering the GPS informed soil sampling (Erickson & Widmar, 2015; Mintert et al., 2016). Before this, data visualisation was not widely used in agricultural management, other than for maps of soil characteristics, and there is little evidence of its use in the literature.

Remote sensing tools: prediction from reflection

The early focus on future crops was a serious limitation. Agricultural productivity can change within a short time frame, so there was a need for PA data to be timely. In the mid-2000s, multi-spectral sensing tools, such as the GreenSeeker™ active canopy sensor (Trimble Navigation Limited, Sunnyvale, California, USA) became more widely used (Rutto & Arnall, 2004). Remote sensing techniques also became more popular, providing aerial and satellite images of crops during the growing season. Both of these tools shifted the temporal emphasis of measurement to the immediate growing phase, allowing growers to respond with near real-time adjustments to the growing environment (Rutto & Arnall, 2004; Whelan, 2011). Contemporaneously, remote sensing technologies became more widely available to PA researchers. These new technologies brought with them their own challenges, with researchers having less control of the data collection process, and in extreme cases working with exhaust data. However, what this assemblage of new technologies did offer was a fertile ground of previously unresearched technologies. For the academy, the late 1990s and early 2000s presented an unprecedented explosion of opportunities to undertake research under the banner of precision agriculture.

Precision agriculture in New Zealand

Agricultural science research has a long, distinguished history in New Zealand. The primary industries produce over half the country's merchandise exports, which in turn account for around three-quarters of total exports of goods and services (Kaye-Blake, 2014). So, while research spending is low by developed country standards, much of the country's research effort has focussed on agriculture (OECD, 2007). Indeed, in the 1990s, New Zealand researchers were producing three times as many papers relative to its size than the world percentage in agriculture (Goldfinch, Dale, & DeRouen, 2003). This research effort extended to include PA in the late 1990s, when universities, particularly Lincoln University and Massey University, and Crown Research Institutes began to invest in PA research (Craighead & Yule, 2001; Yule, 2016), often supported by newly-established remote sensing departments.

The development of PA tools in New Zealand has followed international trends. The technologies employed are oriented to waste minimisation, promising to mitigate farm chemical pollution (Brown et al., 2015; M. A. Oliver et al., 2013), enhance production (D. W. Lamb, Frazier, & Adams, 2008) and deliver economic benefits to the sector (Schimmelpfennig & Ebel, 2016; Swinton & Lowenberg-DeBoer, 1998). However, despite two decades of research effort, like their overseas counterparts, the uptake of PA technologies has been poor (Eastwood, Jago, Edwards, & Burke, 2015; Yule & Wood, 2015), unless there is a compelling value proposition to do so (Jago, Eastwood, Kerrisk, & Yule, 2013).

From the studies available, I suggest that in some regions and sectors, the use of guidance-related technologies such as auto-guidance in tractors is ubiquitous. Other auto-guidance and auto-control technologies, such as GPS-guided precision sprayers and GPS-controlled row planters have also enjoyed high uptake in the cropping sector (Erickson & Widmar, 2015; Mulla, 2013). However, information-related technologies, such as soil sensing, crop monitoring and yield monitoring have not enjoyed the same level of acceptance by farmers in New Zealand as they have globally (Erickson & Widmar, 2015).

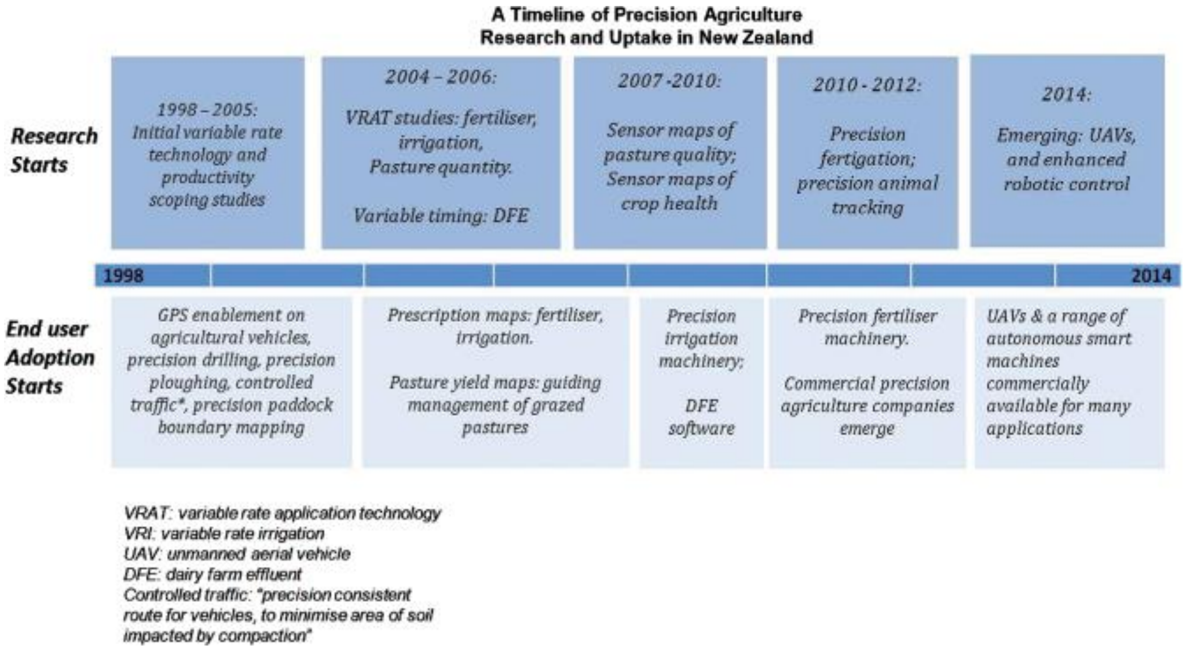


Figure 9 *Timeline of PA research and uptake in New Zealand. From Hedley, (2015 p.14)*

Some local researchers embraced the arrival of PA, with several important innovations emerging from the country's diverse research community, see Figure 9. Researchers originally focussed their efforts on monocultures on the relatively flat terrain of the wider Canterbury region, which were conducive to an industrial approach to farming. Early research includes that performed by (Craighead & Yule, 2001), who investigated the profitability of variable rate fertiliser application in local arable systems. Soon departing from the norm, New Zealand researchers carved out a valuable niche, employing innovations such as animal tracking, proximal pasture sensing, and variable rate application technologies (VRAT) for irrigation and fertiliser application (Hedley, 2015; H. G. Lawrence, 2007; Murray, Yule, & Gillingham, 2007; Yule, Lawrence, Hedley, et al., 2010). These research focus areas were largely developed to service the country's lucrative pip fruit, viticulture, and dairy industries, as the overseas research efforts primarily targeted the monocultures such as corn and soybeans leaving a research gap for local researchers to progress. For example, Bollen, Praat, and Yule (2001) investigated the potential use of site-specific data techniques and information management systems in pip fruit and viticulture enterprises. The researchers identified opportunities to optimise physical and financial aspects of horticultural production systems, along with provision of reliable audit for product security (Bollen et al., 2001).

In terms of remote sensing, GPS-informed farm mapping is the most widely employed PA technology in New Zealand today. From a sociological perspective, GPS farm mapping is also one of the more interesting technologies, in terms of who the data are provided to and for what purpose. GPS-it founder Matt Flowerday started using GPS and GIS technologies in 1998, initially specialising in kiwifruit orchards. Today the company is the leading commercial provider of GPS mapping services in New Zealand and employs proximal, fixed wing aircraft and UAV platforms for remotely sensed data collection. Interestingly, much of the company's work is not primarily focussed on direct farmer benefit. Instead, the client is Zespri, who use the technology to monitor orchards and ensure they comply with their quota of specific varieties. This is interesting in that the primary purpose of the sensing service is not to provide the farmer with information to optimise production, instead it is used for regulatory purposes or to maximise returns to growers by controlling supply. Many pastoral farmers have access to GPS imagery through their fertiliser co-operatives. GPS-it has partnered with the large fertiliser co-operative Ballance Agri-Nutrients Limited to provide members with highly accurate GPS imagery to support planning and precision activities (Ballance Agri-Nutrients Ltd, 2019). The other major fertiliser co-operative in New Zealand also offers members GPS Farm Mapping services through their SmartMaps system (Ravensdown Fertiliser Co-operative Ltd, 2019b).

Despite the progress in other sectors, PA tools and techniques were not applied in the hill country farm environment on farms such as those at the centre of the P2P research until a study by (Murray et al., 2007). The primary reason for this is the steep and variable terrain of these landscapes. Hill country farms are significant contributors to the country's economy. In 2018-19 they generated NZ$5.3 billion from wool, sheep and cattle sale receipts at the farm gate, and a further $3.6 billion of value was added beyond the farm gate for export and the local market, a total of $8.9 billion (Beef + Lamb New Zealand, 2020a). However, society's expectations around food production transparency, environmental sustainability and climate change obligations are becoming increasingly influential at an on-farm level (Beef + Lamb New Zealand, 2020a). In response, researchers such as Murray began to investigate ways to improve their fertiliser practices. Hill country farmers have traditionally widely used conventional (wet chemistry) soil data to guide their fertiliser decisions (Murray et al., 2007). Despite the expense and time required to perform traditional soil testing, it has been the favoured approach because it measures the reservoir of available nutrients in the soil. Unfortunately, the hilly terrain, large scale and spatial heterogeneity (Kemp & Lopez, 2016) of these farms make it impossible to acquire spatially intensive conventional soil testing data to inform variable rate fertiliser applications for all but the flat to moderate slopes. In the research context, traditional plot experiments in the hill country farm context are generally incompatible with conventional statistical methods as it is too difficult to distinguish between sources of variation.

While tissue nutrient analyses are widely available, they are not widely used to guide fertiliser decisions in the pasture context in New Zealand. Tissue analyses measure the nutrient taken up by the plant, rather the amount of nutrient available in the soil. The uptake of nutrients by plants is highly variable (Morgan & Connolly, 2013). There are numerous limiting factors which mean the amount of nutrients in the plant does not necessarily accurately reflect the reservoir of available nutrients in the soil (Murray et al., 2007). In addition, wet chemistry tissue nutrient tests are typically more expensive in New Zealand (source: https://www.ravensdown.co.nz/), and sample collection more cumbersome to collect than soil samples. For example, tissue analyses methods recently developed for clover involve cutting plant tissue from an area and then separating the laminae only for analyses (Olykan, Lucas, & Moot, 2019).

The variable and steep terrain of hill country farms impacts both research and utilisation of PA tools. For instance, even the ATV-drawn C-DAX Pasturemeter described later in this chapter is unsafe to navigate on many hill country farm slopes. The inability to drive over hill country slopes also makes ground-based data collection very expensive and time-consuming. The researchers followed in this study concluded that for producers to improve production efficiency they must first be able to measure it. This prompted the researchers to explore new ways of measuring pasture and soil characteristics, which are discussed in greater detail later in this chapter.

Domestic research organisations

Today, New Zealand has a diverse range of organisations working in PA research. These include several Crown Research Institutes, industry-funded research groups and two universities. New Zealand has an active national association tasked with advancing PA. Established in 2012, the Precision Agriculture Association NZ Inc (PAANZ) connects participants in the sector to one common organisation. The association’s activities include promoting the uptake of PA agriculture technologies in land-based primary production systems, accessing funding for research and the development of technologies, building capability within the sector and promoting technology uptake through industry events, symposiums and field days (Barrowclough, 2013). In 2017, PAANZ hosted the successful PA17-The International Tri-Conference for Precision Agriculture 2017, held at Claudelands Conference and Exhibition Centre in Hamilton. The conference aimed to provide an opportunity for all members of the active research community to showcase their research; in total papers and posters were presented by over fifty New Zealand-based researchers (first or co-authors), giving an approximation of the size of the research community in the country (Nelson, 2017).

Variable rate application technologies in New Zealand

While remote sensing and measurement PA technologies have been slow to establish in New Zealand, some variable rate application tools have achieved commercial success. One such group of technologies are the variable rate irrigators (VRI). The success of New Zealand’s VRI technologies is an important example of multiple supporting factors being in place before a technology realises market success. First, the demand for VRI is influenced by the local political-economic context. Second, reliable application and decision support technologies became available around the same time. Sustainable water management is an increasingly contentious issue in New Zealand, with a 2019 poll revealing that 82% of respondents were very concerned about water quality (NZIER, 2019). Also, in some regions of New Zealand, especially Hawkes Bay and Canterbury, there is conflict over the allocation and management of freshwater resources (Lomax, Memon, & Painter, 2010). Despite these pressures, irrigation has boomed, with a 94% increase in irrigated land from 384,000 hectares in 2002 to 747,000 hectares in 2017 (Irrigation New Zealand, 2017; Statistics New Zealand, 2019). According to Statistics New Zealand, agricultural irrigation covers three percent of New Zealand's land area, with irrigation being largely concentrated in Canterbury (478,000 hectares, or 64% of irrigated land), and Otago (94,000 hectares, or 13%) (Statistics New Zealand, 2019). The tension between farmer demand for irrigation and increased scrutiny of agricultural practices in the domestic media has driven demand for research aimed at tackling problems associated with water quality and climate change. The demand for solutions to environmental problems has accelerated in the past decade, and in response much of the country's precision agriculture research effort has focussed on precision irrigation and the off-farm consequences of agricultural enterprises.

One of the pioneers in the precision irrigation sector was a company called WMC Technology (Where's My Cows).[11] The company was established in 2005 by Massey University Engineering student Stu Bradbury and his friend George Ricketts (Forbes, 2011; Lindsay Corporation, 2019). The company originally offered GPS farm-mapping services including livestock monitoring. Soon after, the pair developed their patented variable rate irrigation system – a GPS device attached to a centre-pivot irrigation system which was launched to market in 2007 (Forbes, 2011). In 2010 the US- based irrigation system supplier Lindsay International purchased WMC Technology and now distributes its variable rate system through its 400-odd dealers worldwide (Lindsay Corporation, 2019).

In 2009, Dr Carolyn Hedley published her thesis on employing proximal electromagnetic sensing methods to advance the monitoring and mapping of soils (Hedley, 2009). Dr Hedley's research demonstrated that EM mapping could facilitate soil inventory mapping for applications such as soil moisture and carbon mapping. This work showed that high resolution environmental monitoring and mapping techniques could inform precision management of natural resources at the farm scale. The development of on-the-go EM mapping enabled a step change in the pedological investigation of New Zealand soils (Hedley, 2009). The resulting ECa (soil conductivity) maps provided the foundation for decision support tools and commercial services mapping soils for variable rate irrigation systems that were being developed at the time, such as

---

[11] Now Lindsay FIELDMAP

the Lindsay Corporation Irrigators (Hedley, Bradbury, Ekanayake, Yule, & Carrick, 2010). The same year, increasing social concern about sustainable water management prompted the local regulatory body Environment Canterbury (ECAN) to set a series of regional targets in the *Canterbury Water Management Strategy (CWMS)*.

ECAN's previous attempts to address concerns had encountered significant barriers, so a collaborative approach was proposed in the CWMS (Lomax et al., 2010). Researchers collaborated with farmers to address serious concerns in finding the technologies that were useful in reducing drainage and accompanying nutrient leaching losses (Bradbury, Mackenzie, Mackenzie, & Hedley, 2013). This confluence of technological advance with social and regulatory pressure created the impetus needed to drive technology uptake. These technologies, supported by the delivery of decision support models created by (Hedley et al., 2010), have been a success story for PA in New Zealand. VRIs are widely used to reduce water use, and one local study indicated water savings of between 8–27% (Hedley & Yule, 2012). EM surveying is conducted throughout New Zealand, with farmers able to access commercial services anywhere in the country (AgriOptics New Zealand Ltd, 2019).

Another important area of research in New Zealand has revolved around addressing the negative impact of agricultural and horticultural operations on ground water quality. Historically New Zealand's farmers, and agricultural policy makers, have focussed on maximising food production within the framework of the family farm. More recently, societal calls to increase the uptake of environmental practices in farming have grown louder (Knook & Turner, 2020). Consequently, farmers are under increasing pressure to adopt environmental practices that align with shifting societal and political expectations of what constitutes good farming practices to maintain their licence to operate (P. Edwards et al., 2019). These pressures have driven demand for research investigating solutions to minimise the environmental footprint of the primary industries in New Zealand.

To date, researchers have had some success in developing variable rate application technologies that aerially apply fertiliser (aerial topdressing) in a bid to minimise fertiliser wastage and avoid sensitive areas. Aerial topdressing is the only practical method of applying fertiliser to the country's medium to steep hill-country pasture and is the preferred application method on this class of land (Grafton, Yule, & Lockhart, 2010). Before commercial aerial topdressing was introduced in 1949, fertiliser was broadcast by hand either on foot or from horseback, which was labour intensive and expensive (Grafton et al., 2010).

While aerial topdressing has important practical advantages in the hill country context, blanket treatments have been the norm and the uniformity of applications has been poor (Yule, Grafton, & Pullanagari, 2014). These blanket aerial treatments have failed to exploit the range of soils, microclimates and the different production systems found on hill country farms (Gillingham, 1973; Gillingham & During, 1973; Yule et al., 2014). Blanket treatments assume that fertiliser response will be uniform despite the variation in output in terms of dry matter grown. Aerial differential rate application technology has the potential to improve the efficiency of the fertiliser applied by only applying it when soil nutrient reserves are low and where growing conditions are conducive to plant growth. Hitherto the terrain on hill country farms has limited the use of most traditional PA technologies to areas where the terrain allows GPS-guided ground vehicles to make pasture measurements and apply fertiliser variably (Mulla, 2013).

Nascent research in aerial variable rate fertiliser application published by Chok, Grafton, Yule, and White (2016) is world leading. This technology is particularly relevant to the New Zealand situation because it responds to challenges relating to the precise application of fertiliser on steep terrain. Importantly, the technology was quickly commercialised through the country's largest aerial fertiliser applicator topdressing company, Aerowork Aerial Spreading, under the brand name 'IntelliSpread' (Ravensdown Fertiliser Co-operative Ltd, 2017).[12] Ravensdown reports that employing their differential rate application system 'resulted in 18% of farming areas of IntelliSpread customers being avoided because they are environmentally sensitive or unproductive' (Ravensdown Fertiliser Co-operative Ltd, 2019a, p. 1). Interviews with company representatives reveal the early adopters have predominantly been hill country farmers, where ground-based application options are limited by the steep terrain. Importantly, this advance in application technologies has driven demand for accurate species and nutrient mapping of hill country farms to inform decisions on variable rate fertiliser application.

Evolution of pasture measurement in New Zealand

In most cases, limited research budgets have prevented New Zealand's PA researchers developing new hardware. Instead, they have developed novel applications of existing technologies. An important exception is the development of the C-DAX Pasture Meter (C-DAX Ltd., Palmerston North, New Zealand), for mapping pasture yield. In New Zealand, pasture measurements are an important metric for researchers and farmers used to establish the mass of dry matter produced, available as animal feed, or retained for feeding out or sale (Gray et al., 2003). Unlike many other agricultural technologies, considerable effort has gone into the extension of formal feed planning and monitoring systems to pastoral farmers (Gray et al., 2003), leading to good uptake rates and demand for measurement tools, especially in the dairy sector.

Traditionally, manual pasture cuts were the main method of accurately assessing pasture yield (Thomson, 1983). Rising plate meters (RPMs) were widely researched in the 1980s, with

---

[12] Aerowork Aerial Spreading is a wholly owned subsidiary of Ravensdown Fertiliser Co-operative Ltd.

standardised calibration formulations being released in 2001 based on a ten-year, multi-site trial (Thomson et al., 2001). This was a radical advance on the slow existing method and prompted researchers to explore opportunities to produce a commercially viable rising plate meter for dairy farmers. In 1999, Massey University researchers used GPS enabled RPMs to measure yield over a grid to improve reliability of spatial information.

In 2002, the local dairy co-operative Fonterra partnered with the Australian government research body Commonwealth Scientific and Industrial Research Organisation's (CSIRO) to investigate the potential use of satellite technologies to monitor dairy pastures (Clarke, Litherland, Mata, & Burling-Claridge, 2006). The intent was that the "Pastures from Space™" programme would deliver farmers real-time satellite-based pasture growth rate information to inform grazing decisions at the paddock scale. By 2006, the programme had delivered several research outputs, but significant technical barriers remained preventing its widespread commercialisation (Clarke et al., 2006). One insurmountable barrier was cloud cover, which meant meteorological conditions often prevented data capture (Mata et al., 2007). This meant that data sometimes had to be extrapolated from previous flights because satellite images were unavailable (Mata et al., 2007). This problem has only been overcome very recently, with the advent of satellite constellations capable of daily image capture in New Zealand (Livestock Improvement Corporation, 2017). Consequently, in commercial terms, the project was effectively in hiatus until late 2017, while other technologies stepped into the void.

According to Clarke et al. (2006), by 2006 around 20% of New Zealand's dairy farms were using feed budgets to manage their grazing systems, although (Nuthall and Bishop-Hurley (1999) in Gray et al. (2003)) suggest this number was as high as forty percent across all pastoral farmers. Many of these farmers were using RPMs (Clarke et al., 2006).[13] RPM sampling is time-consuming and monotonous, so it is little surprise that when the C-DAX Pasture Meter came to market in 2006 it became the most widely adopted PA tool in the country. Originally developed in collaboration with Massey University, the C-DAX Pasture Meter was independently calibrated by AgResearch and Dairy NZ (Yule, Lawrence, Hedley, et al., 2010). The aim of the device is to inform farmers' plans for rotational grazing systems by quantifying variation in pasture dry matter mass between paddocks. The GPS enabled C-DAX Pasture Meter device is towed through paddocks measuring pasture height 200 times a second at speeds of up to 20 km per hour (Rennie et al., 2009). A 500-metre pass with the C-DAX Pasture Meter collects 18,500 measurements compared to a typical 250 readings using an RPM. Data collection is performed and analysed by farmers via the C-DAX interface (Yule, Lawrence, & Murray, 2010). Delivering accurate real

---

[13] By 2006, the RPM gave reliable estimates of pasture cover when at least fifty samples were taken per field in fields with an average herbage mass between 1000- 4000 kg dry mass per hectare (Lile et al., 2001).

time data, the C-DAX Pasture Meter helps farmers improve the efficiency of their rotational grazing systems by indicating which paddocks they should graze next, and for how long (Rennie et al., 2009). When it debuted in 2006, this technology embodied a major shift in the intensity of pasture sampling for New Zealand researchers and farmers.

Case studies examining the benefits of employing feed budgets informed by the C-DAX Pasture Meter have shown significant improvements in pasture utilisation and milk production (Yule, Lawrence, Hedley, et al., 2010). Farmers employing this approach recorded improved yield through better residual management by identifying poorly performing paddocks or areas (Yule, Lawrence, Hedley, et al., 2010). In 2015, over 3,000 New Zealand dairy farmers used the C-DAX Pasture Meter to inform their grazing decisions (Yule, 2015).[14] While the C-DAX Pasture Meter delivers an excellent solution for measuring pasture production on New Zealand's relatively flat terrain, the sensor which is towed behind an all-terrain vehicle cannot safely traverse the country's four million hectares of pasture in the medium to steep terrain classes (Yule et al., 2014). Furthermore, like many sensing technologies developed for PA, the C-DAX Pasture Meter uses a multispectral sensor, which has a limited spectral range. Thus, while it is an excellent tool for measuring pasture dry matter, other applications are limited by the lack of spectral detail. In order to overcome these limitations local researchers switched their focus to airborne sensing options that are not limited by ground-based data capture.

Despite the tremendous advances in differential rate application systems over the past decade, the soil and pasture mapping systems needed to inform fertiliser decisions have lagged behind (Back, Yu, Kim, Chung, & Lee, 2014) and adoption has been slow. To address this research gap, local researchers expanded into advanced pasture mapping research. Early research used proximal multispectral sensors, with the researchers experimenting with active sensors produced by Crop Circle, Topcon, and Trimble. Most of these tools provide an index such as Normalised Difference Vegetation Index (NDVI) and Simple Ratio (SR). Like many PA tools, these sensors were developed for measuring monocultures encountered in arable and cropping systems, rather than the diverse pasture systems commonly found in this country. Researchers reported achieving 'acceptable results' in monocrops such as wheat and maize, however results in pasture were more variable (Yule et al., 2014; Yule, Lawrence, & Murray, 2010). The methods employed using the active sensors are based on light reflectance and were less reliable than estimating pasture mass using pasture height measurements.

In the past decade New Zealand researchers have employed remote sensing tools such as the ASD Field Spec® Pro proximal spectroradiometer (formerly Analytical Spectral Devices Inc., now PanAnalytical, Boulder, CO, USA), hereafter called ASD, and the AisaFENIX (Specim Spectral

---

[14] C-DAX were approached for more recent sales numbers but declined to release estimates.

Imaging Ltd, Oulu, Finland) hyperspectral imaging tool (Pullanagari, Yule, King, Dalley, & Dynes, 2011; Pullanagari et al., 2012; Von Bueren & Yule, 2013; Yule, 2011; Yule, Lawrence, Hedley, et al., 2010). A feature of these technologies is the characteristics of the spectral data they generate (Yule, 2015), which are spectrally and spatially intensive. They generate high volumes of data layered into their spectral bands, delivered as Digital Numbers (DNs) in near-real time. The introduction of data-rich innovations such as hyperspectral imaging has promised to improve decision-making (McAfee, Brynjolfsson, Davenport, Patil, & Barton, 2012), whilst also helping to optimise existing agricultural applications such as fertiliser, seed, and pesticide placement.

A pioneering study published by Sanches (2009) quantified pasture characteristics employing the ASD attached to a Canopy Pasture Probe. This work demonstrated that the ASD can deliver non-destructive in-field sward measurements. However, it also identified challenges for the use of spectroscopy to predict pasture quality parameters in situ, such as surface wetness (Sanches, 2009). Nevertheless, for the first time, the study offered a solution to measuring pasture production in terms of quantity *and* quality in the hill country environment. A later study by Yule et al. (2014) measured metabolisable energy, organic matter digestibility and crude protein, which was estimated by multiplying total nitrogen analysed by a factor of 6.35. Grafton, Willis, McVeagh, and Yule (2016) advanced these findings by repurposing the original data and estimating Normalised Difference Vegetative Indices (NDVI) as a measure of pasture quality using Landsat 8 Imager (NASA). While these findings are noteworthy from a scientific perspective, data collection using the ASD is laborious and expensive, so does not offer an economically viable solution for farmers. Thus, while the ASD has incredible spatial and spectral resolution, it offers little progress in terms of the speed of data collection.

Pioneering to Precision - taking advanced pasture mapping to the air

With the proximally-acquired data collection methods proving to be a commercial dead-end, hill country researchers shifted their focus to the air, initially investigating both UAV (citation withheld) and fixed wing aircraft (citation withheld) as possible data collection platforms for sensors to measure pasture and soil on the steep hill country terrain. Three key drivers led New Zealand researchers to pursue airborne remote sensing technologies. As previously mentioned, satellite-based solutions were not entirely suitable due to on-going issues with clouds and despite their high spectral and spatial resolution, proximal hyperspectral sensors were impractical for measuring hill country terrain. Secondly, the rapid advances in the data processing, storage and analytical technologies associated with the rise of big data meant that the speed and cost of data-intensive technologies had dramatically reduced in the years leading up to the start of the project (Wolfert et al., 2017). Finally, farmers wanted to capitalise on the commercialisation of variable rate airborne fertiliser applicators by mapping their farms at a much higher resolution than ever before. This demand led to the Ravensdown Fertiliser Co-operative, who had already developed the airborne variable rate application technology, to look for methods to develop maps that would inform their application system, and hopefully drive demand for their application service. Thus, in 2014 Ravensdown engaged in a research arrangement with New Zealand's Department of Primary Industries, called the *Pioneering to Precision Primary Growth Partnership Project* (P2P) to investigate broad-scale pasture mapping solutions.

Uptake of precision agriculture technologies

The success of the green revolution supports the notion that farmers will adopt technologies when they have a use for them. Many studies posit that progress of the PA field has been slower than expected in terms of user uptake (Bewley & Russell, 2010; Eastwood, 2008; D. M. Lambert, Paudel, & Larson, 2015). However, given the dearth of PA publications identified in Chapter 2, it is surprisingly difficult to get an accurate picture of trends in the global uptake of PA technologies. This is for two main reasons. First, there is surprisingly little comprehensive, longitudinal research available on the uptake of the technologies. Secondly, the research that is available uses different groupings of technologies and tends to be very region or sector specific. Chapter 2 identifies a small number of publications focusing on technology uptake, although disappointingly, they accounted for fewer than 0.1% of all scientific outputs in the dataset.

While most studies in PA focus on the development of technologies, the surprisingly poor progress made in user uptake rates has not been entirely ignored by the academic community. Even as the field bourgeoned in the 1990s, questions were being asked about the consistencies between PA research and the political and economic requirements of an industrialising agriculture industry (Wolf & Wood, 1997). A review of the literature reveals that early papers focused on variable rate fertilization and variable spraying technology.

Some researchers consider the farmer uptake issues to be related to farmer characteristics (Daberkow & McBride, 2003; Daberkow, McBride, Robert, Rust, & Larson, 2000; D. M. Lambert et al., 2015; Swinton & Lowenberg-DeBoer, 2001; Winstead et al., 2010). For example, Adrian, Norwood, and Mask (2005) found that farmers' confidence in using PA affected their intention to adopt PA technologies. They also found that the farmers' perceptions of net benefit affected the intention to use PA technologies Adrian et al. (2005).

As early as the 1980s, some researchers were concerned by the apparent technology bias of some production-oriented researchers. For example, Bertram Farmer (1986) describes peers who consider technological change a simple matter of diffusion:

> [The] apparent refusal of farmers to stretch out their hands with gratitude and to accept any or all of the content of the proffered 'package of practices' (for frequently only part is accepted) is seen as resistance to change attributable to ignorance or stupidity and to be overcome by massive extension exercise (p. 184).

Occasionally researchers enduring the partial or wholesale rejection of their technologies cast farmers as luddites or uneducated (Gemtos, Fountas, & Aggelopoulou, 2011; Lindblom, 2014). Others suggest that deficiencies in the development process and technology diffusion are to blame (D. W. Lamb et al., 2008). Importantly, there is evidence that some farmers have dis-adopted technologies (Cook, Adams, & Bramley, 2000). Walton et al. (2010) comment that only a few studies (Barham, Foltz, Jackson-Smith, & Moon, 2004; Carletto, de Janvry, & Sadoulet, 1999; Foltz & Chang, 2002; Walton et al., 2008) provide an analysis of adoption and subsequent *abandonment* of agricultural technologies stating “Technology adoption in precision agriculture has received considerable attention, while abandonment has received little” (p. 135).

Similarly, D. W. Lamb et al. (2008) question whether technology development is driven by technology push rather than market demand:

> While these technologies have been shown to provide production and environmental benefits, widespread adoption has been slow. In many cases, new technologies have been produced through developer push rather than user pull. Insufficient attention is paid to well-known technology adoption paradigms and as a consequence, the adoption of precision agriculture technologies is not as great as it could and should be (p. 4).

Despite huge investment in the development of data-rich PA technologies, uptake rates appear to be following their predecessors in terms of slow user uptake. A potential impediment to adoption of data-rich innovations is the lack of analytic capacity to create substantial new knowledge from data (Bennett, 2015; Yule, 2015). High quality independent advice and validated agronomic models are needed to support farmers’ decision-making (Bennett, 2015; Zarco-Tejada, Hubbard, & Loudjani, 2014). Sonka and IFAMR (2014) highlight the need to link data from operations to scientific research to effectively capture value from data, especially the biological data needed to increase agricultural production. Precision agriculture data are of little value unless they can be transformed into knowledge (Huberty, 2015); uptake of PA technologies is unlikely to be widespread unless farmers can understand how to integrate information into their own farming system (Bennett, 2015; Yule & Wood, 2015).

Some researchers predict a dramatic upswing in the use of many existing PA technologies (Schrijver, 2016), which may be optimistic in the short-term if studies continue to show little overall economic benefit of adopting early (Olson & Elisabeth, 2003) and contemporary (Erickson & Widmar, 2015) PA technologies. The relative success of yield monitors, variable rate fertiliser applicators and precision irrigation systems, indicates that if a technology is cost-effective and 'fit for purpose', farmers will use it (Hedley, 2015; Lissaman, Casey, & Rowarth, 2013). So why do some technologies succeed in the PA market and others don't? While there is little argument that farmer-related characteristics (Daberkow & McBride, 2003; Daberkow et al., 2000; D. M. Lambert et al., 2015; Swinton & Lowenberg-DeBoer, 2001; Winstead et al., 2010) and economic considerations (Schmoldt, 2001) play a role in the adoption of PA technology, the influence of social factors and the science-making processes on uptake rates is not widely researched. Instead, the focus of most adoption studies has been on the 'end' of the innovation process, i.e. the diffusion of the technology. Hoffmann (2011) states that the stage prior to the beginning of an innovation's diffusion (especially those events that affect the nature of diffusion later on) has been almost entirely ignored in past research. It is apparent that the 'appropriateness' of technologies to farm communities, agribusiness and the environment is a potentially critical, yet under-researched factor (Eastwood, Chapman, & Paine, 2012; Nowak, 1997).

Some researchers suggest that researchers and technology developers have preconceptions of farmers, their needs and their motives (Fujisaka, 1994; Yule & Wood, 2015), which may lead to the creation of technology that is not 'fit for purpose' (Douthwaite et al., 2001). Other authors point out that some technologies are simply inaccurate or unable to do the job that they were designed for. For example, Back et al. (2014) describe sensors used for measuring application rate on variable-rate fertiliser applicators as being inaccurate and incapable of generating high-resolution maps of the application rate (Back et al., 2014).

Several authors blame some of PA's uptake problems on large-scale, fully-integrated systems that are designed to be most effective when fully adopted because they hinder the farmer's ability to adapt technologies to meet their needs (McBratney et al., 2005; Yule & Wood, 2015). In contrast, Yule and Wood (2015) suggest that farmers will experiment with enabling technologies to see how they can help their business, therefore researchers should create technologies that allow farmers to innovate by improvising (Yule & Wood, 2015). In a similar vein, Douthwaite et al. (2001) state that instead of assuming that a new technology is 'finished' when it leaves the research institute, science-making teams should release their product onto the market and then co-develop in partnership with the end-users as they adapt the technology to meet their needs (Douthwaite et al., 2001). Yule and Wood (2015) suggest that some agricultural scientists take a negative view of the improvisational approach that farmers tend to take to innovation. They argue that for farmers, improvisation takes place within actual farming processes and they suggest that

PA technology be offered as a suite of enabling technologies which farmers could pick and choose from (Yule & Wood, 2015).

Chapter 2 indicated that most papers focused on the positive benefits of new technologies, whereas few outlined the variance in performance and financial, ethical, or social risks of using the technology. Whilst performance issues with technologies may have been under-reported by technology developers, agricultural anthropologists were already starting to question technology bias and the negative effect of the prejudice of those creating agricultural technologies as early as the 1980s (Chibnik, 1981; Farmer, 1986). Chibnik (1981) challenges economists, technology producers and policy makers on their assumptions of farmers' 'risk aversion' and financial situation being a barrier to their technology adoption. Chibnik's study also points to a lack of information on the risk associated with using the technology being presented to the farmer as the major impediment rather than risk aversion inherent in the farmer's behaviour stating that "To make intelligent decisions in such circumstances, farmers must inexpensively learn something about the likely consequences of adopting the new practice"(Chibnik, 1981, p. 4).

The negative influence of the poor performance of earlier PA technologies on the uptake of technologies that followed was reported by Larson, Roberts, English, Cochran, and Wilson (2005), who suggest the unreliability of cotton yield monitors in the late 1990s impeded the wider adoption of later PA technologies in that sector. Fujisaka (1994) also acknowledges that a pro-technology bias potentially exists in agricultural science "Understandably, researchers and extension workers have not been keen to highlight adoption failures" (p. 410).

This is a key concept because if previous poor performance is unknown, its negative effect on future uptake of technology cannot be quantified. A transparent assessment of why technologies have not been adopted or have been abandoned also provides valuable information on how to improve the development and diffusion of subsequent technologies (Sveiby, Gripenberg, & Segercrantz, 2012). While few papers acknowledge the downsides or failings of PA technology, several papers note the harm that variability or ambiguity around risks and benefits can have on the likelihood that farmers will adopt a PA technology (Isik & Khanna, 2003; Roumasset, 1976; Schoengold & Sunding, 2014).

Conclusion

This chapter supports the findings of the previous chapter that data-rich technologies and associated analytical techniques have arrived in the PA realm. Indeed, some PA researchers have hailed these technologies as a panacea to wicked problems (M. A. Oliver et al., 2013).

The analysis in Chapter 2 and this chapter were designed to support investigation of the overarching research question by answering the following questions:

1) What are the key trends in scientific publishing in precision agriculture?
2) Is the nature of data generated by precision agriculture technologies being researched changing?
3) Is the sphere of application of precision agriculture research changing?

There are useful learnings from these chapters. The field of PA research owes its existence to the advances in the enabling technology, GPS. The arrival of accurate GPS-based guidance technology meant that farm vehicles could be navigated more accurately than if operated by humans, and equally importantly, data collected could be geo-located much more accurately than ever before. I discovered much of the research effort in the 1990s and 2000s focused on enabling technologies that facilitate process control. However, I found that in recent decades many researchers have shifted from using direct measurement tools to investigating remote sensing technologies. Specifically, in the New Zealand context, remote sensing technologies have become of interest to researchers working with the challenging terrain of the country's hill country farms. This chapter shows that local researchers have looked to employ aerially acquired data to overcome the difficulties of ground-based data collection in the hill country environment. I discovered that aerially acquired projects have not always gone smoothly, with New Zealand's frequent cloud limiting the use of satellite-based sensors. Recent progress in airborne variable rate fertiliser application methods has prompted researchers to revisit wide-scale airborne remote sensing, with the P2P project emerging as a possible solution to measuring pasture and soil nutrient levels in the hill country landscape.

Chapters 2 and 3 also attempted to explore evidence of what these technologies may mean for researchers in the academy. I found that while researchers in the early decades of PA research mostly employed single technologies, PA researchers working with remote sensing technologies are increasingly employing assemblages of technologies in their research, including big data technologies. However, I also discovered there is a dearth of reflection on, and publishing of, detailed scientific methods and practices, including failed methodological iterations. This includes spare discourse around what data is and what defines it in PA research. Consequently, despite the large volume of academic outputs in the field of PA, many questions on how big data technologies and their assemblages will impact PA researchers and the methods they use to refine value from data remain. On exploring the evolution of science-making in the field of PA I discovered that while there are many excellent studies hypothesising on the future of PA, such as (Schrijver, 2016) and Mulla (2013), most academic literature addressing PA focuses on agricultural systems. I identified only a few peer-reviewed studies on the science-makers and their methods in this field. Indeed, I found that the literature is dominated by quantitative PA researchers publishing in their own field, with little scrutiny or analytical input from the social sciences and other scientific communities.

Until recently, there has been little discourse on the socio-ethical consequences or implications for science-making and knowledge production in the literature addressing PA (Kritikos, 2017). Furthermore, while PA is both a product and reinforcing element of the political economy of contemporary agriculture, there has been little sociological analysis of science-making in the field (Crane, 2014; Hoffmann, 2011). Indeed, there has been very little reflexivity amongst the research community and despite tens of thousands of scientific publications being produced since the 1990s, little is known about their actual scientific methods and approach to science. With all reasonable published avenues of investigation exhausted, I concluded I could only answer the questions posed in Chapter 1 of this thesis by visiting the laboratories themselves. This thesis accordingly aims to contribute to new knowledge of the influence of big data technologies through an ethnographic exploration of a working PA laboratory. Hence, in the following chapters I turn an ethnographic lens on the production of scientific knowledge in a PA laboratory working with data-rich technologies with a view to revealing insights into the nature of the researchers' world and the methods they employ to refine value from big data projects.

“Although it is a rare outcome, it is essential for us to visit the places where the papers are said to originate.”
Bruno Latour

Chapter 4: Research design and methods

Introduction

The previous two chapters confirmed that big data technologies and associated analytical techniques have arrived in the PA realm. We discovered that there has been very little reflexivity amongst the research community and despite tens of thousands of scientific publications being produced since the 1990s, little is known about the actual methods and practices of PA researchers. Hence, many questions on how these technologies may impact PA researchers and the methods they use to refine value from data remain. With all reasonable published avenues of investigation exhausted, the focus moved to visiting researchers at work to contribute to new knowledge of the influence of big data technologies through an ethnographic account of a working contemporary PA laboratory.

It is the contention within this thesis that the influence of the arrival of big data technologies on PA research can only be understood through the exploration of scientists working within a ‘real life’ context. Recent literature stresses the need for PA laboratories to be studied as cultural spaces within agricultural anthropology (Crane, 2014; Maat, 2011). Despite this progress, the previous two chapters demonstrate that agricultural researchers have rarely been the explicit focal point of detailed empirical studies (Crane, 2014), especially in the laboratory setting. Moreover, aside from commentary on user uptake and concerns regarding farmer data ownership, few studies have specifically explored the scientific practice of PA researchers. The PA research community has historically been concerned with technology development, with little or no discussion on the consequences of new technologies on their own science-making practices. Indeed, it would be fair to say that PA has not developed a critical sense of its own emergence.

Hence, to answer the following research questions, I embarked on a 30-month ethnographic study of the P2P research laboratory: “How do the characteristics of technological assemblages affect precision agriculture research and development?” To address this research question, I begin by exploring the following questions:

1. What are the characteristics of new precision agriculture technologies?
2. How do information-rich technologies promise something different for precision agriculture?
3. What challenges, risks and opportunities emerge from the assemblage of data rich technologies in precision agriculture?
4. How do big data influence science-making in precision agriculture?

In this chapter I describe the research approach taken to respond to these questions. First, I describe the grounded theory approach and then discuss the procedures followed for data collection and analysis. Following that, I discuss the role of reflective practice. I then conclude the chapter by sharing some of the ethical problems that arose in the practice of my research. This includes discussion of the conceptual frameworks that I drew upon to assist me and provide a review of ethical considerations made to protect the interests of the study participants.

Background

A fundamental theoretical contribution of Science and Technology Studies (STS) is that beyond revealing insights into the nature of the world, science is itself a socially embedded spectrum of actors and institutions that have cultures of practice (Crane, 2014; Latour & Woolgar, 1986). Like local knowledge, scientific knowledge is the product of social processes that occur in particular cultural contexts (Agrawal, 1995). Yet, Chapters 2 and 3 of this thesis reveal there has been very little reflexivity amongst the PA research community in the empirical literature. Indeed, PA research laboratories are substantially under-theorised as cultural spaces within agricultural anthropology (Crane, 2014). Latour (1987) emphasises the importance of studying scientists at work in their natural laboratory setting:

> Although it is a rare outcome, it is essential for us to visit the places where the papers are said to originate." Latour continues: "We have no choice, however, if we want to apply our first rule of method: if the scientists we shadow go inside laboratories, then we too have to go there, no matter how difficult the journey (p. 63).

Thus, to obtain detailed accounts of the lived experience of PA researchers and their work, a qualitative methodology was adopted for this thesis. In particular, the methods employed draw heavily from work produced by STS scholars such as Bruno Latour and Steve Woolgar (Latour & Woolgar, 1986), and Karin Knorr Cetina (Knorr Cetina, 2001; Knorr, 1981), who used fieldwork methods to address similar problems to those posed in this thesis (Knorr Cetina, 2001; Latour & Woolgar, 1986; Sismondo, 2011).

This chapter critically explores the methodological approach adopted. Each phase of the research is explained, detailing how the fieldwork was consequently shaped by events within the object of inquiry. I then engage with those issues of ethics integral to the fieldwork, describing the challenges encountered and how they were navigated in the course of the research.

Research approach

The methods employed in this thesis draw from more generalised social studies on the role of laboratories in science to guide this research (Maat, 2011; Rhoads, Wilson, Urban, & Herricks, 1999). The findings in this thesis are based on observations, interviews, and artefacts obtained during ethnographic fieldwork conducted by the P2P researchers between May 2015 and October 2017.

Ethnography is considered the methodology of choice for actor-network theory informed researchers (Nimmo, 2011). Through immersion you observe a lot, including interpersonal dynamics. While they are critical to how any social group operates, for this study, the determination was always made in terms of value for the thesis argument as it emerged.

I acknowledge that if the study was conducted from another sociomaterial approach, such as cultural-historical activity theory or practice theory, the research would look very different. Hammersley (2018) suggests that most ethnographic work shares most or all the following features:

- Relatively long-term data collection processes
- Taking place in naturally occurring settings
- Relying on participant observation, or personal engagement more generally
- Employing a range of types of data
- Aimed at documenting what actually goes on
- Emphasises the significance of the meanings people give to objects, including themselves, in the course of their activities, in other words culture
- Holistic in focus. (Hammersley, 2018)

Armed with research questions initially relating to why the uptake of PA is disappointing, it was apparent that a robust response to the research questions would require more sophisticated data than direct questioning alone (Bauman et al., 2002; Becker et al., 2002; Christensen, Johnson, Turner, & Christensen, 2011; Glesne, 2016; Merriam & Tisdell, 2015). My interest in the details of scientific activity led to an STS based approach that would help me describe and explain the researchers seemingly messy and complex actions (Blommaert & Jie, 2010) by observing them in their natural habitats.

Science, technology, and society (STS) scholars study scientific knowledge and practice as cultural phenomena (Maat, 2011). A key objective of this study was to gain an understanding of the science-making process within PA teams, especially those working in HSI laboratories. An important STS concept emerging from Bruno Latour’s “The Pasteurization of France” is how science created in a laboratory experiment requires physical and social adjustment to ensure what works in the laboratory also works in the real world, and the obverse (Latour, 1993; Maat, 2011). This is potentially very relevant to the case study being explored in this study because the researchers were trying to develop a highly advanced hyperspectral imaging technology for a 'low tech' sector.

It also makes theoretical and ethnographic sense to survey the processes of a contemporary PA laboratory from a single vantage point, that of a participating observer (DeWalt, 1998; Knorr Cetina, 2001). Participant observation in fieldwork traces back nearly a century to the diaries of Malinowski, written in 1922 (Malinowski, 1967) and unsurprisingly how this work is carried out varies hugely. Lofland (1976) in Van Maanen (2011) describes ethnographic methods as "sprawling, diffuse, undefined and diverse" and this study proved to be no different. Ethnographic methods generally demand the ethnographer's presence in the culture of study (Geertz, 1974). In his 1987 work *Science in Action: How to follow scientists and engineers through society*, Latour (1987) stresses the need for work beyond interviews and surveys, following the researchers into their place of work.

Influence of actor-network theory

Actor-network theory was developed by the Paris group of science and technologies between 1978 and 1982 (Latour, 1996; Law, 2009), led by Latour and Woolgar (1986) and Callon (1986). ANT is a sociological investigative perspective from which to understand the processes by which scientific disputes become resolved, the introduction of new ideas become accepted, and by which new tools and methods are introduced and adopted by a group (Brewer, Gajendran, & Hilaire, 2013). In terms of this thesis, the actor-network consists of, and links together, both technical and nontechnical elements.

ANT talks about the heterogeneous nature of actor networks (Hanseth & Monteiro, 1998) and links the human and nonhuman actors in a network, who are called "actants" in ANT. ANT was selected as an approach for this study, and I deployed its powerful tools where useful to explore issues at the socio-technical interface. In practice, ANT provided the tools to untangle to the messiness of the PA laboratory. Importantly, it provided a systematic way to reveal the extent and influence of the messy practices and process that are usually left out of the "heroic" accounts of scientific and technological achievements (Brewer et al., 2013).

The deliberate focus on exploring, following, and documenting the active work of things did not preclude paying attention to humans, but rather reoriented my inquiry to consider myriad ways human and non-human elements come together. A key challenge was to know when a human, machine, or when even data was an actor. Hence, ANT provided a form of methodological sensitivity rather than a highly defined set of statements about how the world should be. So, while this thesis is not about ANT, following the objectives of ANT allowed me to include not only the P2P researchers in the ethnographic account, but also to follow non-human actants to characterise their interactions. The result is an ethnography that is open to the role played by non-human agencies such as technologies and data. This is important because while earlier I critiqued the lack of reflective concern for humans, the solution is not simply to bring back the humans, because in doing so you also bring into play the various agencies that they contend with in their practices.

The catalyst for this decision was an early observation in my study which revealed that the P2P researchers were struggling to follow the linear project plan laid out for them, coupled with the discovery that the published accounts of the P2P researchers' science methods were depicted as linear when the reality was much more complex and iterative. In Chapters 2 and 3 I discovered that PA literature's focus on technology is achieved by reducing technologies to a compliant instrumentality for the transformation of nature. However, when I visited the laboratory, I saw that technology and nature are not so simply reduced and play a more active part in scientific practice than the PA literature imagines. Employing an ANT informed approach appealed as it foregrounds the agency of non-humans within practice. In essence, the approach allowed me to remain open to the possibility that non-humans add something worth studying. Hence, operating from an ANT perspective allowed me to examine how novel tools such as the AisaFENIX (and the data they produce) impact on how science making in the contemporary PA laboratory actually happens.

A multi-site longitudinal study

The laboratory studies performed by Latour and Woolgar (1986) demonstrate that being close to localised scientific activity can provide a preferential view from which to understand the work of researchers and other actants. I too, attempted to capitalise on this privileged situation by getting to know the P2P researchers well enough to discover patterns of everyday life, grounding discussions and assertions on what they do and what they *really* think is going on, as opposed to gathering official and published versions of processes and events. By selecting a single object of inquiry, the P2P project team, this study also aligns with the tradition of laboratory studies (Knorr Cetina, 2001; Latour & Woolgar, 1986; Lynch & Amann, 1996), with the focus of the work being the observation and documentation of the researchers both within their laboratory and 'in the wild' of fieldwork. The longitudinal nature of the study allowed me to gather data on the researchers, their tools, and their data *in situ* as it emerged in a natural and unforced way.

The study evolved into a multi-site ethnography, an approach where I follow people, instruments, data, connections and relationships across space (Falzon, 2016). I also follow actants and practices beyond the spatial boundaries of the traditional laboratory. I found that through immersion you observe a lot, including interpersonal dynamics. While they are critical to how any social group operates, the determination was always made in terms of value for the thesis argument as it emerged. Of central importance in this respect is that ANT provides a lens through which to view the role of technological assemblages in shaping social processes (and possibly vice-versa). While a single case study is unable to provide a universal view of the complete PA research ecosystem, it does allow the me to detect patterns and themes in observations that may allow the findings to be applied more broadly.

Grounded theory approach

Social science research is about understanding social worlds. Thus, this thesis aims to provide interpretive understanding of the studied world, not just descriptions of actants and their practices. When considering possible research methods, I sought an option that would enable me to flexibly follow the varied social, cultural, and political contexts inhabited by researchers. Thus, to support the fieldwork and data analyses for this study, a grounded methodology based on that of Charmaz and Mitchell (2001) was selected. Their grounded approach builds concepts that assume the existence of multiple realities, and importantly, acknowledges the mutual creation of knowledge by researchers and other research actors (Charmaz & Mitchell, 2001). In this study I applied tools of the grounded theory approach with the aim of facilitating a search for potential actors and their interactions and to provide evidence of specific translations of innovation in the case presented.

This systemic, yet flexible approach allowed a variety of research methods to be employed, including direct observation, semi-iterative interviews, analysis of research artefacts, reflection meetings and end-user feedback sessions, which is ideal for explorative studies such as this.

In accordance with Charmaz and Mitchell (2001), the basic strategies employed for this study were:

1. simultaneous data-collection and analysis.
2. pursuit of emergent themes through early data analysis.
3. discovery of basic social processes within the data.
4. inductive construction of abstract categories that explain and synthesise these processes.
5. integration of categories into a theoretical framework that specifies causes, conditions, and consequences of the process(es).

This study followed an iterative process of data collection (Glaser, Strauss, & Strutzel, 1968) which enabled me to sequentially identify the most significant issues in the field of study. The strategy of collecting and analysing data simultaneously proved pivotal to the course of this study as it allowed me to focus on developing concepts about the data and to gather further data that fleshed out the nascent concepts. For example, in the early months of my research, a theme emerged from the early data analyses indicating that the researchers were struggling to follow the linear science process prescribed in the project set-up. This piqued my interest and it became apparent that the data generated by the project were unfamiliar and challenging to the group. The flexibility of the grounded theory method allowed me to adapt my line of inquiry to ask deeper, more relevant questions informed by real-time observations, rather than strictly following a predetermined research path. The trajectory of my research shifted to focus on the emerging themes that my supervisors and I thought merited deeper analyses. Hence, the focus of the study moved to the researchers themselves and the data characteristics of the nascent technological assemblages they employ.

Actants

A wide range of actants shape the social processes described in this 30-month study. Taking centre stage in the study's network is a high-flying hyperspectral sensor called the AisaFENIX and the data it generated. Other key actants include the AisaFENIX's sidekick, the ground based ASD, and the researchers who hoped to build an understanding between these tools and their assemblages, and the biochemical reality of their targets. This community of people and machines is described in detail in Chapter 5 of this thesis.

Importantly, the role of actants in my study was not stable - objects, and even individuals, were not static, pre-formed substances but rather surfaced through a series of negotiations between an ever-evolving assemblage of actors. Some human actors were followed for very short periods of time, whereas others participated for the duration of the project. When I initially began my study, I had envisioned that humans would take the centre stage of the project and would form my main data sources. In Year 2 of the study, a spin-out company was formed by two members of the P2P research team. This venture was supported by the University's commercial arm. The emergence of the spin-off company, led by my supervisor at the time, was not unexpected in this type of research. While this may have looked outwardly as a moment of revolution, it was factored into the project. Importantly however, the development impacted on access to some participants. The availability of participants that were involved with the spin-out company was noticeably impacted. The mood within the research team also moved markedly, with some of the researchers being 'left out' of the development. In terms of the shape of this thesis, the moment served as a pivot point for the study as it influenced who the relevant actors were. Undoubtedly, this decision also influenced my choice to observe inanimate objects (data and instruments), as well as humans as the research progressed.

Another influence on the decision to involve non-humans was the response when the researchers identified issues with their tools during reflections. By looking at the data through an ANT lens, the study evolved to include non-human actors, including the data being generated by an assemblage of tools. In time, it became apparent that my task was to create a coherent account of the complex assemblage of tools and data, without separating the natural from the social, economic, and political (Nimmo, 2011) without the distinctions which normally constitute the structure of a social scientific explanation (MacLeod et al., 2019).

Like Nimmo (2011), I employed ANT to move beyond human ontologies. Rather than treating tools and data as intermediaries, these non-human actants played mediating yet contingent roles in science-making. Table 8 provides a brief overview of the key actants followed for this ethnography. While this thesis is not about ANT, a challenge was to know when a machine, or when data was an actor. As in The Zimbabwe Bush Pump example (De Laet & Mol, 2000), the machines in this study (and their data) have characteristics that other actors have to consider. The ethnographic study revealed that the machines and their data had characteristics that impacted science-making, which is the main emphasis of this thesis. Hence, the *data* produced by the tools became a key focus of this study. In Chapter 6 I provide an ontological description of the data wrangled by the P2P researchers originating from the P2P's community of machines.

Table 8 *Description of key actants in the Pioneering to Precision project ethnography*

Organisation	**Name and Role**	**Number of participants**
Non-human	AisaFENIX hyperspectral sensor	2
	ASD Field Spec Pro® hyperspectral sensor	
	Data produced by tools	many
Massey University	Research Director/ Lead researcher	1
	Researcher	4
	Research technician	2
	Post-graduate students working on P2P-related projects	4
	Project administrator	2
	Commercialisation and Intellectual Property Manager	1
	Casual Field Assistants	3
	Overseas intern (short-term)	1
USDA	External Science Advisor	1
Ravensdown	Technical Development Manager	1
	Key Account Manager	2
	Agri Manager	1
PGP Participant Farmers	A diverse group of farmers whose farms are test farms for the project, i.e. their farms were surveyed with the AisaFENIX. The farms are listed in Table 11. The participant farmers include experienced farm managers and farm owners with expertise in agribusiness, indigenous farm management, rural valuation, and rural banking.	9

Data collection

At the outset of this study, it was decided that to achieve deep understanding, I, like many ethnographers before me (MacLeod et al., 2019), must engage in observation over a sustained period of time. So, for over the thirty months I was embedded with the P2P researchers, which enabled me to gather a vast collection of research artefacts relevant to the daily activities of the researchers, which are listed in Table 9. By employing ANT-oriented observation, I was equally concerned with non-humans, and focused on understanding the processual nature of the P2P researchers' world.

The volume, scale and diversity of data collected was vast. By employing ANT-oriented observation, I was equally concerned with non-humans, and focused on understanding the processual nature of the P2P researchers' world. Primary data sources included my research diaries (including notes on observations), interview transcripts and videos. The artefacts collected were intended not simply to provide ethnographic evidence of my presence in the research environment, but I used them judiciously to contextualise and support the findings and assertions of this thesis. In the tradition of Collier and Collier (1986), images and data visualisations collected during the study were also occasionally used as visual stimuli in interviews and discussions ('photo elicitation') with informants to support discussion.

I recorded over forty hours of video footage, audio recordings and other artefacts, including computer screenshots that inform and support this study's findings on the P2P researchers' personal perceptions of the science-making process. Audio recordings captured on a digital recorder were the primary means of gathering data from interviews and meetings. In most instances the recorder was unobtrusive, but I was mindful of collecting data that may compromise the respondent. On several occasions, I ceased recording when the conversation strayed beyond the realm of this study, or if information was commercially sensitive. In lieu of taking notes, I kept my interview protocol on my lap and checked that all questions were covered and circled questions that yielded especially useful data for later review.

Photographs of whiteboard graphics and brief videos were captured as necessary, with permission. While the video footage was not included in the final thesis, it proved particularly useful in building an understanding of how machines moved, how computer processes were sequenced and how the scientists interacted with the instruments. The video footage was also helpful when analysing how the scientists *described* their processes (which was not always linear), and when observing them demonstrate how they perform their work on computers. As previously mentioned, while capturing the scientists in action for this study occasionally involved laboratory benches and whirring machines such as in the earlier studies by Latour (Latour, 1987), much of the P2P researchers' work was performed on computers. Videoing the researchers employing multiple computer monitors to display and analyse their data was useful in understanding the non-linear approach to their science-making.

In addition, key descriptors including interview venue, time, participants' names, tone of the interview, points to follow up and reflective comments were documented in my research diary. Interviews were transcribed verbatim by me and coded using NVivo^{TM} Version 12. Other artefacts such as drawings, data visualisations, emails and photographs were stored electronically and linked to the relevant interview in the NVivo™ files. Identification data were recorded on a separate tracking sheet in a secure location as suggested by Glesne (2016) as a record of the interviews completed. The data were analysed using the data management software tool NVivo™.

Table 9 *Data sources for the ethnographic study of the P2P research team, 2015-2019*

Method	**Details of procedure**	**Rationale**
Research Diary	Written throughout the research period. Especially during data generation stages	Capture thoughts, ideas and interesting bits and pieces that 'might be useful' at a later stage. Provides a record of my journey.
Field observations (participant observations)	Observations of P2P science team, end-user feedback sessions and field activities while embedded in the P2P science team.	Learn about the 'science-making process' to enable me to enter meaningful discussion. Observe the science team & explore how they interact, share knowledge, gather knowledge. Collect qualitative data over the 30 months.
Interviews and dialogic reflection	Transcripts generated from audio recordings of semi-structured interviews with the PGP science team (frequency varied) and the project funder.	To collect qualitative data over the research period.
End-user informal discussion	Conversations with end-users, e.g. farmers and Ravensdown representatives.	To observe how P2P researchers seek feedback from the end users.
Formal end-user feedback session	Observations of the formal end-user feedback session.	Observe how the researchers interact with end-users.
Informal discussion with researchers	Conversations with researchers throughout the research period, many taking the form of dialogic reflection.	Provide an insight into the scientific methods & processes being used by scientists. To develop an understanding of HSI, GPS, GIS, and other PA technologies. Observe how the P2P researchers interact with scientists from other disciplines to identify new opportunities to refine value from hyperspectral data.
Photographs	Photographs of P2P researchers performing laboratory and field research. Record whiteboard graphics.	Provide a record of events. Used for data analysis (participant involvement, behaviours, trends). Recall of prior events, topics, themes, situations, and conditions.
Digital audio recording	Record conversations with participants in semi-structured interviews, feedback sessions, meetings, field observations.	Provide a full record of conversations that I can transcribe and use as data for this research. Available for reflective phases of research and a method to record events.

Table 9: Data sources for the ethnographic study of the P2P research team, 2015-2019 continued.

Method	**Details of procedure**	**Rationale**
Video recordings	• Video record of conversations with participants during semi-structured interviews, feedback sessions & meetings. • Video footage of *PGP* & related fieldwork. • Video footage of data processes being described. • Video footage of P2P researchers going about their everyday work. • Video footage of apparatus (AisaFENIX), being set up & in operation.	Develop deeper understanding of the inner workings of the P2P laboratory, with a focus on science-making & scientific division of labour.
Team meetings	Observations of & participation in regular P2P team meetings (includes notes taken during meetings in my research diaries).	
Data visualisations[15]	Data Maps	Develop a deeper understanding of how data maps such as data cubes are used in the data analysis process.
	Diagnostic Maps	Provide historical context. To provide data/ information on the different iterations of maps being produced as outputs of P2P research.
	Prescriptive Maps	
	Application Maps	Understand iterations of application maps associated with research.
Research Artefacts	Emails, media releases, blogs, newspaper articles, website content (including marketing information), technical manuals, personal communications, presentations (including conference and marketing presentations).	Provide an insight into the progress that each of the stakeholders believes they are making at each stage of the project.
	Research reports produced by P2P researchers, MPI & co-funder including project milestone reports.	
	Government documents and papers, including reports by the Auditor General.	

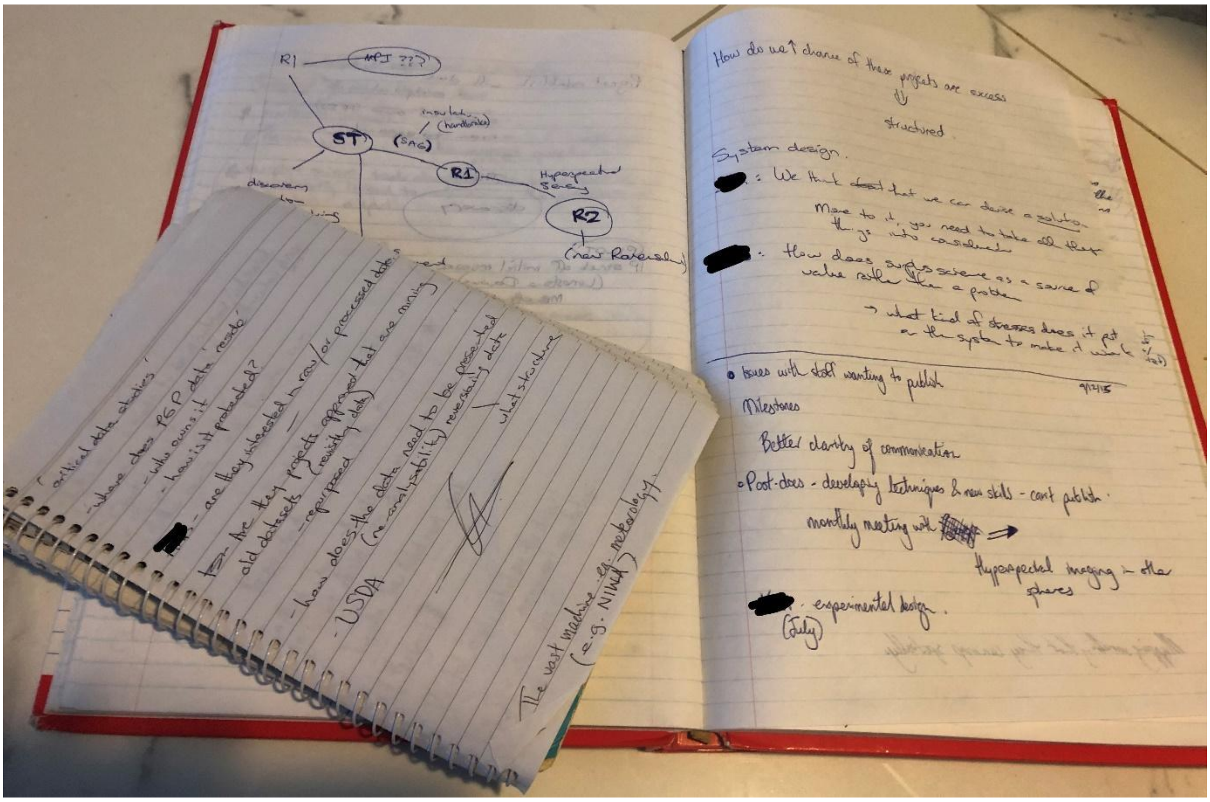
---

[15] These include data visualisations created by the P2P Science team to communicate trial results within the science team, to other researchers and to end users. Versions over time were analysed (including historic maps produced by external projects) to provide data/ information on the different iterations of maps being produced as research outputs.

Research diaries

I first started keeping my hand-written research diary as a little challenge to myself having turned digital native in the years prior to starting my doctoral research. My first research diary was a large hardcover red book where I recorded the everyday workings of the laboratory. It was here that detailed how I felt, what I touched, smelled, thought, and saw. I kept extensive notes, paying attention to formal meetings and events, and carrying on informal conversations with the researchers.

My diaries contain my records of observations of the daily activities of various actors in the laboratory, in the field, and in planning and reflection meetings. They include a vast body of comments and information acquired from direct observations and informal conversations, see Figure 10. After going through my fourth hard copy diary I reverted to my digital world, finding that digital notes were more easily searchable. The diaries proved to be the most important source of data for this thesis are my research diaries.



R1 — MPI ???

ST (SAG)

R1 — Hyperspectral Sensing

R2 (new Ravensdale)

(critical data studies)

- where does P&P data reside

- who owns it

- how is it protected?

[redacted] - are they interested in raw/or processed data?

~~[illegible]~~ - Are they projects approved that are mining old datasets (revisiting data)

- repurposed

- how does the data need to be presented (re-analysability) revisiting data

what structure

- USDA

The vast machine e.g. meteorology (e.g. NIWA)

How do we ↑ chance of these projects are excess

⇓

structured

System design.

[redacted]: We think ~~[illegible]~~ that we can devise a solution.
More to it, you need to take all these things into consideration

[redacted]: How does [illegible] science as a source of value rather than a problem

→ what kind of stresses does it put on the system to make it work

- Issues with staff wanting to publish 9/12/15

Milestones

Better clarity of communication

- Post-docs - developing techniques & new skills - can't publish.

monthly meeting with [redacted] ⇒

Hyperspectral imaging - other spheres

[redacted] - experimental design.
(July)

Figure 10 *My research diaries have proven to be a valuable research resource*

Official documentation

A large body of documents relating to this study on Government websites, such as the Ministry for Primary Industries (MPI) website (https://www.mpi.govt.nz/) were curated and electronically stored to ensure their availability for this study. These documents include:

- Ravensdown Quarterly Progress Reports, see Appendix A.
- Ministry for Primary Industries: Managing the Primary Growth Partnership (L. Provost, 2015)
- Ministerial media releases

The official documentation provides a glimpse into the interactions between the Government funder, MPI and Ravensdown. The quarterly progress reports in particular show the progress of the project in terms of outcomes and budget spend. They also record key decisions, such as the six-month extension to the project, granted as a result of Covid-19 disruption. The reports also reveal the progress the research has made in terms of implementation. For example, the November 2020 report reveals that a total of seven farms had used the AisaFENIX data to inform the application decisions in conjunction with the variable rate airborne application system, Intellispread.

Direct observations

Following and documenting the daily activities of the researchers was not straightforward. Unlike studies such as those enacted by Latour and Woolgar (1986) and Knorr Cetina (2001), the laboratory bench in the contemporary PA laboratory is no longer a physical space. Instead, almost all the laboratory work for the project was executed on computers, making direct observation difficult, but not impossible. Observing and documenting the researchers at work relied on screen captures, video recordings, data visualisations and explanations of the processes by the researchers. The exception was the demonstration of the AisaFENIX as a laboratory tool, which allowed direct observation of the researchers using the hyperspectral sensor in a laboratory setting.

In practical terms, the fieldwork provided a richer picture of the social and technical factors at play in this project, especially in terms of some recalcitrant informants. The reluctance of several of the researchers to participate in academic discourse and the practice of critical reflection within the group presented challenges (and interesting insights) for direct data collection for this study. The absence of opportunities for collegial reflective practice for the emerging researchers and students also limited the opportunity for data collection. However, being embedded in the fieldwork teams was a unique opportunity to get close to the actants, providing a privileged vantage point to observe and document work processes.

The fieldwork also provided me with a hands-on opportunity to follow the technologies involved in the project and to comprehend the scale and scope of the P2P project. Other than the regular team meetings, the P2P team's field work was also the only time the post-graduate students interacted with the University's paid permanent and casual staff. Hence, observing the fieldwork provided a better understanding of the cultural whole, by making sense out of what was seen, heard, and talked about when participants from across the group's social order were together. Assisting in the P2P project fieldwork also presented the opportunity to me to meet participant farmers and understand their motivations and thoughts on the project.

Semi-structured interviews

Semi-structured interviews established the foundation for the data collection and analysis. In total, 26 iterative, semi-structured face-to-face interviews were performed with members of the laboratory, project funders, end users and other parties relevant to the project. As a newcomer to ethnography prior to embarking on this journey, the initial stages of the study were guided by Associate Professor Brennon Wood. Early interviews were performed in collaboration with Dr Wood, an experienced interviewer, which added valuable insight to the interviews and served as an apprenticeship as a reflecting practitioner.

Like most academics, the P2P researchers were very busy, and interviews were typically performed when and where the researchers were available.[16] An example of the initial contact questions can be found in Appendix B. Finding suitable locations for recordings was complicated by some respondents' hesitance to arrange appointments because they were on call for field work if the weather was favourable. Thus, some interviews were recorded during lunch breaks in testing environments such as cafeterias. Most respondents worked in communal office spaces, such as the PhD communal office darkly named the 'Morgue', a silent, dark cold room where some feared upsetting their colleagues if they broke the silence of the room, even for a short time. This made recording processes on computers difficult outside of normal working hours.

Ideally, ethnographic interviews are less formal and for this study I was guided by the strategies of Hermanowicz (2002). The duration of the interviews (typically sixty to ninety minutes) was of sufficient length to allow complex issues to emerge. Some of my early interviews could be likened to speed dating.; short encounters where I aimed to get to know the participant and gain an insight into their role in the P2P project. I took a scatter-gun approach, interviewing as many members of the team as possible.

Negotiating access to the P2P researchers was relatively simple as the P2P team's Director was my supervisor at the time. However, finding the time, place, and mood to conduct sensitive and

---

[16] An example of the initial contact questions can be found in Appendix B.

meaningful interviews was deceptively difficult. Obtaining interviews from some paid staff members and students was a challenge. Some researchers appeared somewhat suspicious of my activities. This doubt appeared to stem from a general unease regarding social scientists among the group, who proudly described themselves as quantitative researchers. One researcher rather amusingly questioned if I would be trying to expose their personal foibles. Others appeared concerned that I may reveal their academic shortcomings. These fears largely dissipated over time and it turned out that the researchers' heavy workloads and siloed work practices would impact data collection far more than any hesitancy by the interviewees. During the 'speed-dating' phase of my interviews, I discovered that while the participants were not overly keen to talk about themselves, they were very interested in talking about what they '*do*'. From then on, I would start my subsequent interviews with the question "What are you working on at the moment". I later enquired about *how* they went about their work and *why*. Eventually, as the non-human actants became more important in my study (and as trust grew), interviews evolved to the participants *showing* me what they do. This led to great advances in the quality of data gathered around processes, data and challenges the researchers faced in the making of science.

Conversation, talk and chat

The interviews captured structured, basic data, the details of which occasionally supplied deep levels of meaning. However, I often came away from the interviews sensing that much more on the topic in question could be gathered from the respondent. The depth of detail collected from interviews varied significantly between respondents. Respondents with whom I had formed trust and a good rapport provided richly detailed pictures of their social world within the P2P team. Others remained guarded in their interviews, even late in the project, and in these cases I have drawn where possible from observations made by myself and others in the group.

Nonetheless, much of the primary data for this thesis emerged from quotidian conversation, talk and chat during the thirty months that I was embedded within the P2P team. As my time in the field progressed, my rapport with some P2P researchers grew, allowing for engagement and verbal expression that more closely resembled everyday talk. This important relational factor facilitated the gathering and analysis of more authentic data with greater depth of detail.

The primary data were used to illuminate the journey of hyperspectral data and their supporting data, the specific data practices that the researchers employed, the socio-cultural values that framed and were used to justify the researchers' data practices, and the institutional context in which the research was performed.

Farmer feedback discussion

Observations were made of a farmer feedback discussion held between the researchers and participant farmers to provide data to support analysis of the socio-cultural context and values of the researchers. These meetings served as an opportunity for the researchers and farmers involved in the project to view and give feedback on the progress of the research. The feedback meeting was held on 11 August 2015. Participants in the session included the P2P science team, farmers and managers of participant farms, fertiliser company representatives and a sociologist. In addition to the observations of the researchers, the session also provided a valuable opportunity to explore and clarify farmer participant views and canvas their opinions of the outputs produced for the PGP project. Data from interviews, meetings and other interactions were stored in and analysed using NVivo^{TM}.

Data analysis

I generally followed the data analysis procedure adapted from Creswell and Clark (2017), see Table 10. I also drew on the approach articulated by Charmaz and Mitchell (2001), which meant that data collection and analyses were conducted simultaneously. The nature of this research compelled me to reflect on my own truths and I concede it is ironic that the data journey is presented here as an ordered, highly linear model, when the reality is very different. In truth, there is a major nonlinear aspect of the method: the reflexive interplay of data collection, writing, and analysis. While the procedure roughly represents the steps taken, these steps were often taken backwards, iteratively, and sometimes were skipped entirely. This simplification of the data analysis also fails to convey any measure of the variation in workload involved, e.g. capturing, and preparing the data for analysis was a massive undertaking compared to exploring the data for this novel study. Having gathered most of the data myself and transcribing the interviews personally, I knew my data well, making it a fairly easy and enjoyable task to complete, especially when aided by the specialist qualitative software NVivo^{TM}.

Table 10 *Data analysis procedures*

Data Analysis Procedures	**Qualitative**
Preparing the data for analysis	• Organise documents and visual data. • Transcribe interviews, meeting notes and field observation notes. • Prepare data for analysis with NVivo™ Version 12.
Exploring the data	• Read through data. • Write notes
Analysing the data	• Code the data. • Assign labels to the codes. • Group coded into themes and categories. • Use NVivo™ Version 12 to analyse data
Representing the data	• Represent findings in discussions of themes and categories. • Present visual models, figures, and tables.
Interpreting the data	• Assess how to address the research questions. • Compare the findings with the literature. • Reflect on the meaning of the findings. • State new questions based on the findings.
Validating the data and results	• Use validation standards, e.g. triangulation, participant validation etc. • Check accuracy of transcripts and field notes. • Check for reliability.

Adapted from Creswell and Clark (2017).

For the purposes of this study, I draw from Latour's 1988 definition of the laboratory "I will call this place the laboratory, which for now, simply means, as the name indicates, the place where scientists *work*" (Latour, 1987, p. 449). Thus, I followed the P2P researchers, their machines, and their data to the university-based laboratory, and into the field to observe their field data collection activities. The grounded theory approach meant that feedback from data collected during the field work was influential on the direction, approach and focus of the research. The early interviews followed a predetermined common structure designed to extract the information that I (thought I) needed. This was the approach that I had followed in my prior research extension career when needing to obtain specific information from researchers.

Analyses of this first round of data collection showed that while this approach was useful for examining predetermined outcomes, it was not ideal for exploratory research. To support an exploratory approach required a change in my approach to interviews. The recording sheets were duly discarded, and interviews were subsequently guided by themes rather than predetermined questions. I drew on the expertise of my supervisor, who joined me for some early interviews so I could observe his skills and hone my own interview technique. In adapting my approach I also called on the valuable advice in *The Great Interview: 25 strategies for studying people in bed* (Hermanowicz, 2002), except for 'don't date members of your own family', which in this circumstance could not be avoided.[17] Where possible, results were compared with field observation notes and triangulated with other sources of data where appropriate (Flick, 2004).

Events within the P2P project also shaped this evolution of the fieldwork. Fifteen months into my field work, it became apparent the Director and the University were considering creating a spin-off company to commercialise the hyperspectral sensing and imaging service through the establishment of a spin out company. The formation of the company created tension in the group, particularly for information sharing, which affected some P2P researchers' willingness to share their methods and thoughts with me. In response, I adapted my approach to seek the information from others, or to switch to direct observation rather than relying on interview data. Funding conditions had always restricted public sharing of the project; however, the transparency of the work notably decreased following the decision to commercialise the hyperspectral sensing and imaging service. This meant I had fewer formal academic publications to draw on and some of the outputs from the post-graduate students were embargoed. Nonetheless, there were still research outputs such as conference presentations and proceedings, and a few journal publications that were captured and curated for analysis. While commercialising the hyperspectral sensing and imaging service created challenges, this was offset by the unique opportunity to follow an evolving science team where the goals of a small section of the research group diverged from those of the others.

Role of reflective practice

Sociologically-informed reflection is a key aspect of critical analysis and is commonly used in health and education professions, though is applicable to all professions (N. Thompson & Pascal, 2012). Reflection refers to the process of thinking about the work we undertake at the time (reflection-in-action) or at a suitable opportunity thereafter (reflection-on-action) (N. Thompson & Pascal, 2012). Reflection and co-reflection with others is advocated as an important process for the success of transdisciplinary projects (Rijswijk, Bewsell, Small, & Blackett, 2015). Critical reflection can be useful for addressing gaps in scientific understanding or to identify potential risks associated with a new technology (Hasu, Leitner, Solitander, & Varblane, 2012; Schön, 1983). Unfortunately, the use of this approach is under-theorised in the PA sector, although there has recently been some excellent work on reflexive practice and evaluation completed in New Zealand (Fielke et al., 2017; Rijswijk et al., 2015). An interesting feature of the P2P project is that a social scientist, Associate Professor Brennon Wood assisted the P2P team in an attempt to establish a culture of scientific self-reflection. This is unusual in terms of PA teams and provided

---

[17] Hermanowicz (2002) suggests that researchers avoid interviewing people they know, except if a very particular kind of research design calls for it. In this case my husband was already part of the research group so the situation could not be avoided.

an opportunity for the researcher to explore the role of scientific self-reflection during a period of disruption (Schön, 1983; N. Thompson & Pascal, 2012).

Somewhat unexpectedly, various forms of reflective practice eventually informed many of the key findings of this thesis. Interviews with the Director evolved to increasingly centre on dialogic reflection. The Director reported that the dialogic reflection approach helped develop a deeper understanding of the science processes that they were employing and also helped him work through solutions to the challenges the researchers faced in the P2P project. Interestingly, however, there was a noticeable reluctance by the wider research group to participate in critical reflection in the community setting. The inability of the group to apply such intellectual processes inevitably impacted both data collection for this thesis and the progress of the P2P project itself. 'Team meetings", which could have provided a rich source of academic discourse and practice, largely operated as a task allocation measure, and attempts to generate academic discussion were regularly derailed by belittling behaviour, often instigated by the non-academics in the group. In Year 3 of the project, "Virginia W" (P2P Researcher) lamented the lost opportunity to perform meaningful critical reflective practice:

> I just feel that there is a lack of communication that's totally unnecessary. We should be having a collegial discussion meeting, which was about our research and doing what we're actually meant to be doing, not about a quick meeting to cover health and safety in a cramped room. There's a history of people not undertaking the task that they were given at the last meeting and there's no collegial discussion and I would think that they're afraid of collegial discussions. There's no excuse for it really. It pisses me off. (YR3)

Ethical considerations

On reflection, my quantitative science background left me ill-prepared to address some of the complex ethical problems that arose in the practice of my research. My previous research participants were generally turf plants, or soil samples, so it was evident from the outset of this study that the ethical considerations would be very different. However, at the outset I underestimated the extent to which ethical questions would influence the trajectory and outcomes of this study. During the planning phase of the study, my supervisor Associate Professor Wood advised me that ethical considerations would be of paramount importance in the planning and practice of this study. And so it eventuated that ethical tensions became part of the everyday practice of doing this ethnographic research (Guillemin & Gillam, 2004). The continual consideration of ethical problems (and issues of intellectual property) helped shape the thesis, such as a continual re-evaluation of which actors to follow. Indeed, ethics questions were reflected upon throughout the study in what emerged to be a 'wrinkly' project.

Emergence and response to ethical considerations

While the research proposal was guided by the University's universal ethical principles for research laid out in the University's *Code of Ethical Conduct for Research, Teaching and Evaluations Involving Human Participants* (Massey University, 2015), information on how these principles should be applied in practice was limited. So, in the early stages of my study I discussed the project's ethical implications with my supervisors and the University's Commercial Enterprises unit, who provided me with valuable guidance. The project proposal was also influenced by the European Commission Report Research Ethics in Ethnography/ Anthropology including the principles of 'doing good, not doing harm' and 'protecting the autonomy, wellbeing, safety and dignity of all participants' (Iphofen, 2013).

The project proposal was peer-reviewed, and a low-risk notification made to the Massey University Human Ethics committee in May 2015. The project was subsequently recorded on the University's Low Risk Database, see Appendix C. The continual evaluation of ethical questions (and issues of intellectual property) helped shape the thesis, such as a continual re-evaluation of which actors to follow. As the research took shape, my supervisors approved proposed measures to ensure the research was conducted in accordance with the University's ethics guidelines.

Key ethical issues that emerged during the study included how to respond to perceived (but not directly communicated) hesitance from some participants, how to manage special relationships (including how to balance the 'participant' and 'observer' roles), and in the latter stages of the study, the process of anonymisation.

Protecting participants

Much of the continual evaluation of ethical problems relating to this study centred on protecting the participants. Transparency around the observational role was paramount. From the outset of the project, no attempt was made to conceal the observational role. Care was taken to ensure all participants understood their rights and all participants were given the option to not be included in the study. Participants were provided with a Participant Information Sheet at least 24 hours before their interview and were asked to complete Participant Consent Forms, which have been stored for future reference (see Appendix D for an example of the Participant Information Sheet). Each participant interviewed provided written consent.

Another key consideration in protecting participants related to access controls. Two main categories of access control were considered:

1. Restricting access to raw data (storage). Raw data such as interview audio and video, photographs and transcripts were stored on a password-protected cloud-based storage platform. I was the only person with access to this resource library.

2. Thesis embargo (when can others access the information disclosed in the thesis). As my thesis was sponsored by the university, embargoing my thesis had not originally been a major consideration in terms of obligations to prevent the early disclosure of commercially sensitive information. Certainly, my preference was initially to make my work available to the public for scrutiny. However, I have decided to embargo this thesis until after the participants' embargoes have expired to protect their interests.

Managing participant hesitance

I engaged various human and non-human participants in all stages of this study – from deciding on scope and problem statement to actively collecting data through to the submission of the thesis. Part of the continual evaluation of ethical issues included considering the shifting roles of participants (including myself) and researchers. Without the participation of various actors, the research simply could not have happened. I was mindful however, that this reliance on the participants should not overwhelm ethical decisions around including participants.

Hence, it was important to regularly re-evaluate which actors to follow, not just in terms of capturing relevant data, but to address participant hesitance. In Year 1 of the study I detected some recalcitrance from three participants. I had entered the realm of the qualitative researcher, where social scientists are sometimes viewed with suspicion. However, before long, I had established myself within the group and all but one of the researchers showed enthusiasm for my research. While all participants agreed to take part in the study, my gut instinct was that one researcher was hesitant to participate in the study. In response, I chose not to continue following the researcher moving forward.

Anonymising data: a balancing act

The process of anonymisation proved to be a complex, and far from water-tight process, supporting the assertion of Saunders, Kitzinger, and Kitzinger (2015) that anonymising data is a balancing act between maximising protection of participants' identities and maintaining the value and integrity of the data (Saunders et al., 2015). The ethics of 'anonymity by default' was discussed with my supervisors, however it was determined that some 'harm' (Iphofen, 2013) may result from lack of anonymity .

In practice, this means that pseudonyms were used throughout the thesis; some participants even asked to choose their pseudonym. Interestingly, in one case, the participant selected a name not typically associated with their gender – this is noteworthy for any researchers looking to utilise this research for secondary analysis involving gender.

Once engaged in the anonymity process, I found that changing people's names and disguising the name of the research organisation were only first steps in a more nuanced process around managing identifying details. The process identified practical challenges, including how to conceal identities on a research project and location that is traceable. It was determined that anyone closely tied to the P2P research setting may be able to recognise participants, in what Tolich (2004) refers to as threats to 'internal confidentiality'. However, it was decided that these difficulties faced in anonymising did not justify its abandonment (Kelly, 2009), as future harm from naming participants cannot always be reliably predicted (Wiles, Crow, Heath, & Charles, 2008).

Once I decided to attempt to conceal the anonymity of the individual participants, I then considered whether this should also equate to blanket anonymisation, i.e. disguising every single identifying detail mentioned in the thesis. While the use of pseudonyms (for people or places) does not guarantee anonymity, extending beyond this approach proved to be a more complex decision than first envisioned. Other than the work of Saunders et al. (2015), the literature provides little advice on how to anonymise research like mine in practice. After consideration, I decided to also obscure the name of the research organisation studied. So, while the title of the *Pioneering to Precision* project (or its co-funder) has not been concealed for this research, the identities of individual participants were concealed.

In addition to concealing identities and in some cases gender, this thesis also deliberately omits some references where typically a citation would be included. In these cases, 'citation withheld' is noted to indicate that the work is not my own, however the identity of the author(s) has been concealed in a quest to provide robust anonymisation. I recognise that this approach is a compromise in that the participant does not receive recognition for their work they would have received if the work was acknowledged with a full citation. Interestingly, at the conclusion of the study, two participants wanted me to use their real names in the thesis. This was not entirely unexpected as other researchers, such as Saunders et al. (2015) have encountered participants who wished to be identified (Saunders et al., 2015). However, I admit that this challenged my preconception at the time that participants would take an 'anonymity by default' approach. In response, I discussed the options with each of the two participants and then with the agreement of both the participants and my supervisors, I reluctantly concealed the participants' identities as naming may compromise the anonymity of other participants in the group through their research and organisational links.

In choosing pseudonyms, I wanted to avoid revealing too much about the ethnic and cultural backgrounds of the participants in the multicultural P2P research team. While some researchers have avoided pseudonyms altogether such as Corden and Sainsbury (2006), such an approach seemed to me to be quite impersonal and would also make it harder to follow individual narratives.

Some aggregation of occupations was also performed in the description of the individual participants. The participants are largely described as either being a ‘post-graduate student’, or a ‘P2P researcher’, which indicates whether the participant was a university employee or student. This division is helpful in terms of unpacking some of the power relationships at play in the study.

Managing special relationships

Another challenge of being ‘immersed in the setting’ (Irwin, 2006) of the P2P project, was managing the special relationships with participants, including that with my husband, who was a doctoral student within the research group that I studied (and therefore a participant in the study). Irwin (2006) reflected on some of the debates regarding intimacy and exploitation of being intimately connected to research participants. Like Irwin, I found myself balancing the permissions and obligations of a special relationship that involved a spouse, although in the case of Irwin their relationship developed during the research period. In practice, the measures taken for this thesis included not disclosing the content of others’ interviews and my husband did not read or review this thesis prior to publication. Information regarding pseudonyms was also not shared.

> “A lot of people have tried it and not been successful. I just made the comment, 'You better look at those to see why they failed.” Dr Sam U, Advisor to the *Pioneering to Precision* PGP.

Chapter 5: Pioneering to Precision: data acquisition in the new data economy

Introduction

The recent arrival of big data has sparked debate on how this may impact PA. Some scholars predict emerging big data in PA research will spawn a new wave of technologies that will succeed where first generation technologies failed. This view assumes that all blame regarding the field’s failings lay with technology alone, with little consideration of the innovation systems they emerge from and the social worlds they seek to inhabit.

While there is a rich tradition of studying scientific practice inside laboratories in other fields (Knorr Cetina, 2001; Latour & Woolgar, 1986; Sismondo, 2011), PA laboratories remain largely unexplored spaces. The bibliometrics analyses in Chapter 3 revealed that PA researchers and their scientific practice have rarely been the focus of academic enquiry (Crane, 2014). Scientific practices within the PA research community are largely obscured and there has been little scrutiny of how PA laboratories operate (Crane, 2014). Consequently, a significant knowledge gap exists in understanding how science-making teams work within PA, making it difficult to accurately represent how the arrival of big data technologies may affect research in the field.

In Chapter 3 I concluded that I needed to follow the researchers in their work to understand them, and their practices. As a result of following my actors into the laboratory, I started to see the data practices and process as a key site of angst for them.[18] Consequently, I deliberately narrowed the focus of this chapter onto a team acquiring and analysing data for the *Pioneering to Precision* (P2P) project. These researchers employed nascent data-rich technologies in an attempt to provide farmers with unprecedented insight into soil nutrient distribution in the hill country farm context. I introduce the actors in this narrative, both human and machine, and find that the innovation process is driven by the main actor, a remote sensing tool called the AisaFENIX.

To help answer the research question “How does the arrival of big data influence science-making in precision agriculture?”, I describe the aims, chronology, and background of the P2P project to provide context on the world within which these actors operate. A key actor that emerges is the AisaFENIX – this chapter will examine how the research team’s focus on this device was framed by the unique scale, scope, and characteristics of the data it generates. I examine the relationship

---

[18] Which I later came to realise the researchers simply called ‘the office’.

the AisaFENIX has with its supporting actors and find that the success of the AisaFENIX relies on being part of a wider assemblage of technologies and actors. As a fellow actor in the P2P project, I also depart, as the researchers did, from the comfortable control of the laboratory to explore on-farm field data collection, where the true challenge for these researchers about how to account for the large number of variables that leaving the laboratory leaves unresolved becomes evident.

Background

In late 2014, I was invited to perform an ethnographic study of a PA laboratory within Massey University, New Zealand. Established in 2001, the now defunct research group resided within Massey University's School of Agriculture. Prior to my arrival, the group worked on various small research projects, notably co-developing the commercially successful C-DAX Pasture Meter sensor with the local company C-DAX Ltd (Callaghan Innovation, 2016).[19] The highly specialised group focussed on remote sensing technologies, although they had also dabbled in animal monitoring research. In 2014, the research group rapidly expanded in terms of research personnel in response to the establishment of a NZ$10.3m externally funded project *Pioneering to Precision*, and the arrival of a new hyperspectral sensor, the AisaFENIX.

The invitation to study the P2P research team was extended by the group's Director, who was interested in understanding factors that influence the uptake of PA technologies. He hoped an ethnographic study might shed light on constraints within the innovation process. The offer, accompanied by scholarships from the University, provided access to staff, meetings, scholarly papers, documents, and other artefacts within the laboratory. A working space was supplied among the doctoral students working on P2P, giving daily access to the heart of the research group. I would spend the next 30 months embedded in the P2P research team in an effort to explore its researchers, its gadgets, and its science-making.

Arrival

I arrived in the P2P laboratory in February 2015. Few ethnographers have followed working PA scientists into their laboratory. As such, there is little record of their world. Entering the laboratory, I was concerned with ontological matters. Who are these heroes tasked with addressing global food security concerns? What does their laboratory look, smell, feel and sound like? Perhaps getting to know these researchers and the spaces they inhabit will help explain the failings of a field and shine light on how these explorers may solve some of society's wicked problems.

---

[19] C-DAX Limited is a wholly owned subsidiary of Ravensdown Fertiliser Co-operative Ltd.

Like all ethnographers arriving in the field, I immediately had to account for myself before the actants I proposed to learn to know (Berreman, 2012). My husband Tommy had worked as a PhD student in the laboratory for over a year, so I had some understanding of the project prior to my arrival. Despite this, other than my husband and the Director, I only had a passing acquaintance with the researchers. As a student of Latour and Woolgar (1986), I was cognisant of the range of reactions that my arrival might produce. However, as a fellow professional scientist in the field of agronomy, I was optimistic that my arrival would not arouse too much suspicion and that my interest would be willingly accommodated by the working researchers. Furthermore, I was confident that my familiarity with some aspects of their scientific activity would eventually allow me to make sense of the observations I was to record. My first encounter with the working researchers, a staff meeting, starkly exposed my naivety, see Research Diary Excerpt.

[RESEARCH DIARY EXCERPT] 2015 (Year 1)

Initial Meeting

The meeting is held in a small, dimly lit room with dividers pulled back to accommodate the group. Several team members cordially acknowledge each other as the room swells beyond capacity. Extra chairs are quickly fetched as latecomers arrive. Those already sitting edge their chairs together to make space for their colleagues. Almost everyone appears to be in their twenties or early thirties. As the group grows impatient, conversations break out, revealing the diversity of the group through the multitude of accents.

When the Director ascertains that everyone important is present, he quietens the group before starting his address. The Director introduces me, suggesting my research may help address problems that PA researchers have in getting uptake from farmers. The team members politely introduce themselves and briefly outline their background and research focus. The composition of the group surprises me; there is not a plant or soil scientist among them. Instead, the scientists hail from diverse backgrounds in engineering, volcanology, animal behaviour, agricultural aviation, and remote sensing. Indeed, a doctoral student who is a golf course architect is not a surprising addition to this heterogeneous group.

The introductions reveal a common theme. Each scientist related their work to a device they called the "Fenix". The Director states that he wants to talk about some of the issues with the Fenix. As conversations between the researchers ensue, any remaining hopes of familiarity are dashed by my inability to understand much of what is being said. The scientists speak of problems processing the Fenix data. I jot down some of the terms I am unfamiliar with. *Orthorectification. Binning. Spectral smile correction. Zero values. BSQ files. Radiometrically corrected. Envy* (sic). *Mosaicking*. The list continues.

I felt like an alien. I didn't even speak their language.

This interaction laid bare my shortcomings as a newcomer to the laboratory. Despite fifteen years in the agricultural sciences, I was an outsider. I had anticipated I would feel at ease amongst these peers. Moreover, I thought that I would speak their language. After all, we are all researchers working in the agricultural sciences, aren’t we?

Pioneering to Precision: aspirations of a ‘revolutionary’ project

The project the researchers were referring to as the ‘PGP’ was a project called *Pioneering to Precision. Pioneering to Precision* was one of 22 projects initiated under the now superseded Primary Growth Partnership (PGP) Programme. In 2009, the newly elected National led government established the PGP programme to fulfil its policy to increase export earnings from primary industries (Battell, 2018).

The PGPs provided an important new funding mechanism for projects seeking to improve the efficiency of hill country pasture systems and respond to the increased pressure coming onto farmers, both financial, and to optimise fertiliser efficiency to reduce off-farm effects such as contamination of hill land waterways (Dodd et al., 2016). Hence, the overarching objective of the P2P was to enhance fertiliser practice on New Zealand hill country farms by utilising remote sensing technologies to assess the nutrient status of farms to determine where nutrients should be targeted (Ministry for Primary Industries, 2014; Office of the Auditor General, 2015; Ravensdown Fertiliser Co-operative Ltd, 2014).

While economic benefits were the main drivers of the P2P project, an interview with the co-funder’s representative “Steve N” suggests that farmers’ licence to operate also influenced research priorities:

> We have been conscious of that. It wasn’t necessarily the trigger for the PGP.
>
> Some of the things that we may do specifically may not add to a whole lot of increase in fertiliser sales. But it may protect or help our farmers in their social license to operate and in their ability to farm into the future.
>
> I think there’s still a lot of conversation to be had about... eventually the public perception. They’re going to have to start asking some deeper questions about what that actually means. And what I mean by that is that it is easy at the moment to isolate a sector, e.g. dairying etc but that conversation will mature and move on to, actually this is an urban and rural issue that has to be addressed.
>
> So, in the short run, yes I think there is a lot around public perception, but I think there’s a lot around public perception around its not only around water quality, water quality is the topical one right now around New Zealand but if you look at it in a global sense it's going to be around pasture-fed versus pasture inside versus from a lab or from a petri dish and I think that they are all going to be challenges for New Zealand farmers to address.

> So, if we look at that big picture that's the kind of sphere that Ravensdown operates in rather than just solely fertiliser sales. (YR4)

The project leveraged a recent advance in airborne differential rate application technologies, which operates an automated hopper door based on the GPS location of the aircraft (Chok et al., 2016; Morton, Stafford, Gillingham, Old, & Knowles, 2016). These technologies have several benefits, including improved pilot safety and more accurate fertiliser placement (Chok et al., 2016). However, the benefits of the application technology cannot be realised if the farmer does not know where the fertiliser needs to be applied (or not applied). The P2P project promised to inform the fertiliser decisions that would be delivered by the new automated hopper technology. The hope was that the algorithms generated by the project would eventually guide and inform fertiliser decision-making, and ultimately replace farmers' (and fertiliser sales representatives) decision-making.

Pioneering to Precision: project genesis

Fundamentally, the *Pioneering to Precision* Primary Growth Partnership project (hereafter called the 'P2P') was a quest to build an algorithm to calibrate an airborne hyperspectral sensor as a proxy for measuring plant tissue nutrient status, which in turn would be determined as a proxy for soil nutrient status on hill country farms (Ministry for Primary Industries, 2014).

Occupying over ten million hectares of land, hill country farming is an important primary industry in New Zealand (Cameron, 2016). Export earnings from the beef, lamb, venison and more recently manuka honey produced on hill country farms are significant contributors to the country's economy (Apiculture New Zealand, 2016; Beef + Lamb New Zealand, 2018). Beef, veal and sheep meat exports alone, which are mostly grown on hill country farms generated approximately NZ$8.4 billion in revenue in the 2019-20 season (Beef + Lamb New Zealand, 2020b). The value of the country's honey exports, mainly derived from manuka grown on hill country farms generated NZ$355 million for the economy in the 2018/19 year (Ministry for Primary Industries, 2019). The key drivers for the P2P project included not only the New Zealand Government's aspirations to grow these export receipts, but also their desire to address some of the volatility in the economic performance of the sector (Beef + Lamb New Zealand, 2018). Adding to these considerations was the fact that fertiliser spend is a major expenditure item for New Zealand's hill country farmers, who face uncertain returns on their investment given existing application techniques (Chok et al., 2016).

Researchers have proposed PA, particularly variable rate fertiliser treatments, as a means to improve economic and environmental sustainability of hill country farms (Chok et al., 2016; Grafton et al., 2010; Murray, 2007). The project was a 50:50 investment partnership between the

Ministry of Primary Industries and the farmer-owned fertiliser co-operative Ravensdown Fertiliser Co-op Ltd, with a financial commitment between the parties totalling $10.3m (Office of the Auditor General, 2015).[20] The programme promised significant potential economic benefits in the form of additional export earnings of $120 million per annum by 2030 and a net economic contribution of $734 million to New Zealand's economy from 2020 to 2050 (Ministry for Primary Industries, 2014).

New Zealand's hill country pasture systems are extremely diverse, with a wide range of production in terms of dry matter grown and the quality of what is grown in a range of microclimates (B. Zhang & Tillman, 2007; B. Zhang, Tillman, Gillingham, & Gray, 2009). The terrain of hill country farms is highly variable, even over short distances (Gillingham & During, 1973; M. G. Lambert & Roberts, 1978). Slope influences many factors affecting plant growth, including soil fertility and moisture levels, so hill country pastures have high spatial variability (Gillingham & During, 1973; Radcliffe, Dale, & Viggers, 1968). The interaction between slope and soil physical attributes also has consequences for productivity through water availability, which strongly influences botanical composition and pasture growth (Kemp & Lopez, 2016). Despite the botanical diversity of these pastures, fertiliser applied on hill country farms is generally targeted to improve the production of clovers; specifically, subterranean clover (*Trifolium subterraneum),* and white clover (*Trifolium repens*).

The nutrient status of hill country soils is highly diverse (Gillingham, 1973). Even so, farmers have defaulted to blanket fertiliser applications where airborne application is required due to technology constraints (Gillingham & During, 1973), with variable results. Differential rate application has been suggested as a way to help farmers respond to the variability of the environment, with a view to increasing farm profitability (Murray, 2007). However, historically researchers have struggled to provide farmers with accurate maps of soil nutrient status of hill country farms to guide these fertiliser decisions. Scientific progress has been hindered by the enormous number of variables that contribute to hill country pasture systems, as Figure 11 produced by a P2P post-graduate student (citation withheld) demonstrates.

---

[20] Due to Covid-19 disruption, the P2P project was extended by 6 months; by November 2020 the total investment between the Government and Ravensdown had reached NZ$10.7m

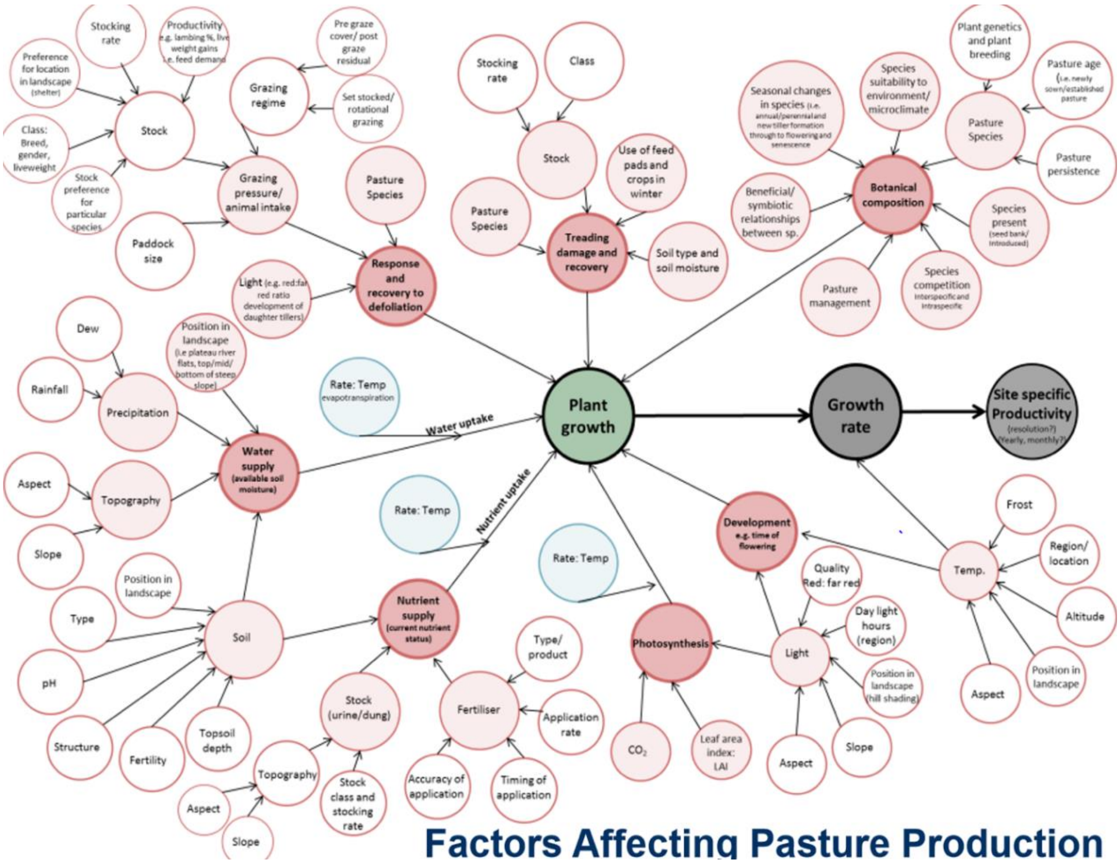


Figure 11 *Factors affecting pasture production. Source withheld*

Application technologies

Aerial topdressing is the only practical method of applying fertiliser to New Zealand’s medium to steep hill-country pasture land and aerial topdressing is the preferred application method on this class of land (Grafton, 2010). Indeed, approximately 40% (approximately 1.2 million tonnes per annum) of the country’s fertiliser is applied by fixed wing agricultural aircraft (Grafton, Yule, Davies, & Jones, 2009). Before commercial aerial topdressing was introduced in 1949, fertiliser was broadcast by hand either on foot or from horseback, which was labour intensive and expensive (Grafton, 2010). In Chapter 3 we learned that blanket aerial topdressing treatments have been the norm for hill country farms in New Zealand, but this approach has failed to exploit the range of microclimates and the different production systems found in these environments.

Differential rate application technologies have the potential to improve the efficiency of the fertiliser applied by only applying it when soil nutrient reserves are low and where growing conditions are conducive to plant growth. Hitherto the terrain on hill country farms has limited the use of most traditional PA technologies to areas where the terrain allows GPS-guided ground vehicles to make pasture measurements and apply variably fertiliser (Mulla, 2013). Recent advances in airborne variable rate application have improved their accuracy and these systems are now commercially available in New Zealand (Chok et al., 2016). The P2P project co-funder, Ravensdown, owns an aerial topdressing firm, Aeroworks. Ravensdown's representative "Steve N" explained that his company's motivation to fund the research was linked to their desire to match the resolution of application technologies to remote sensing advances:

> …the resolution it gives you is fantastic. However, you've still got to fly on the fertiliser, you've got to carry out the plan. You can't fly fertiliser to the resolution you get with remote sensing so therefore there is some progress that needs to be made. (YR4)

However, there is little benefit in the farmer using variable rate technologies unless the input decisions are informed by reliable information. Understanding the heterogeneity of hill country pastures would enable farmers to implement variable rate application where competitive and cost efficient (Schellberg, Hill, Gerhards, Rothmund, & Braun, 2008). Until recently, the absence of differential rate application technologies for hill country farms (Murray, 2007) has meant that there has been insufficient commercial demand for spatially intensive measurement to warrant large-scale research on the topic. In response to the commercial release of its airborne variable rate application system IntelliSpread (Ravensdown Fertiliser Co-operative Ltd, 2017), the fertiliser co-operative Ravensdown applied to the Government's PGP programme to support the P2P project, which would develop remote sensing technology suitable for use on hill country terrain.

Attempts to estimate pasture quality using satellite remote sensing platforms have been made in the past in New Zealand (Ausseil et al., 2011; Clarke et al., 2006). However, the limited available spatial and spectral resolution, coupled with the difficulties brought by cloud cover meant this technology had limited commercial success in New Zealand. An island nation with changeable weather, New Zealand is locally known by the Māori name Aotearoa which means 'long white cloud'. So, it should be no surprise that cloud is a major consideration for any future remote sensing technology proposed for the country, and as such is an important actant in the P2P project.

Globally, progress in spatially intensive crop measurement has been achieved using proximal remote sensing technologies, although this research has been largely conducted on relatively flat terrain with homogeneous crops such as corn and wheat (Franzen & Mulla, 2015). Remote sensing of pasture biochemical characteristics is relatively new and under-researched, although the C-DAX Pasture Meter has achieved some commercial success for measuring pasture mass (Yule, Lawrence, & Murray, 2010). Unfortunately, proximal technologies like the C-DAX Pasture Meter are unsuitable for commercial deployment in the hill country farm situation as the steep terrain prevents ground-based vehicles from navigating the steep terrain.

Prior to the P2P project some of the researchers on the team had used, with some success, a proximal ASD hyperspectral sensor to measure pasture quality parameters (citation withheld) The promise of the original proximal sensing project prompted the P2P researchers to perform an initial airborne trial of the AisaFENIX to describe pasture parameters in hill country farms (citation withheld). Both the preliminary proximal and airborne studies were run on a well-known New Zealand research farm (name withheld). This unique property offered researchers a rare opportunity to perform field trials in a hill country environment where many variables that are normally unaccounted for are well defined. The farm is as close to laboratory conditions as a field researcher could hope for. For example, decades of soil science research on the property have produced accurate, highly detailed soil maps of the farm (citation withheld). Even more important in the case of airborne hyperspectral sensing, detailed topographic data were available that included highly accurate Digital Elevation Model (DEM) data (citation withheld). Taking hyperspectral remote sensing to the sky would eliminate the terrain constraints of data collection using the proximal sensors such as the ASD but would introduce issues of cloud and topography, problems that were unresolved at the outset of the P2P project.

Centralising intelligence in airborne fertiliser topdressing operations

Decisions on fertiliser applications on New Zealand’s hill country farms remain primarily based on the farmer’s local knowledge. The farmer will typically sit down with a trusted fertiliser sales representative and map where and when fertiliser should be applied based on a wide range of factors. These factors could include a farmer’s knowledge of the farm (such as soil type and conditions, past and future crops, terrain, location of sensitive areas, areas with ongoing pest issues), the farm’s fertiliser history, and a small number of soil tests taken across the farm. The P2P project did not simply aim to improve the efficiency of soil nutrient measurements. If successful, it promised to centralise intelligence in new centres of algorithmic calculation.

Figure 12 represents the P2P project’s aspirational model, where a sensor will measure nutrient levels, the algorithm will determine the fertiliser needs, then the prescription is sent to the computer onboard the topdressing aircraft, which automatically opens and closes the hopper. All of this is achieved without human intervention. No expert judgement is needed. The expertise previously called on in the process is rendered redundant, possibly leading to a side lining of the farmer, sales representative, and pilot input in decision-making.

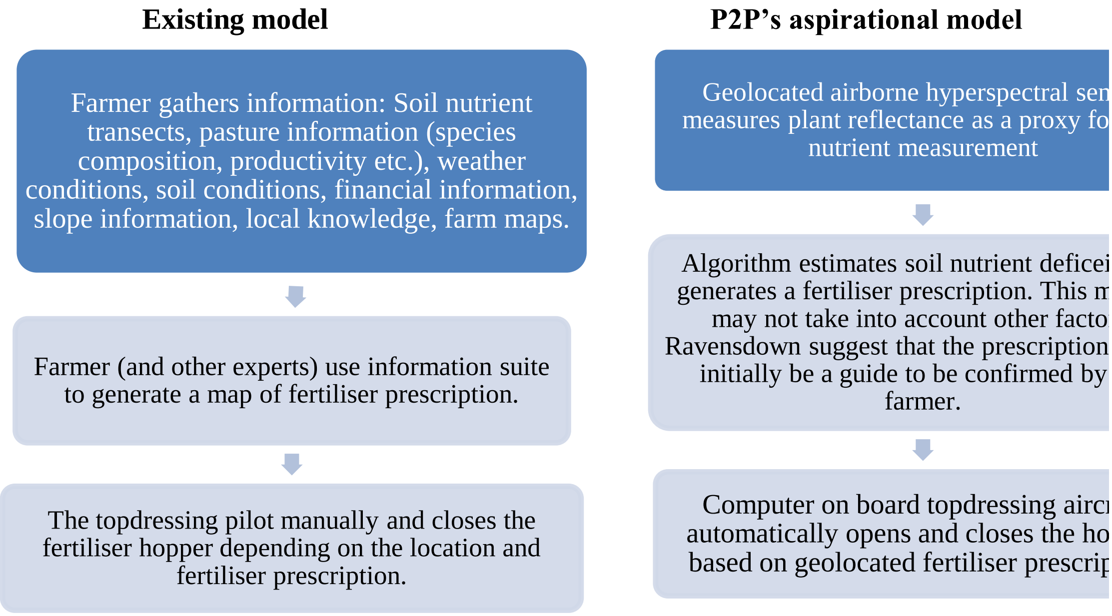


Figure 12 *Pioneering to Precision: A model to centralise intelligence in airborne fertiliser topdressing*

The case for removing the pilot from the decision-making process is compelling. A GPS-controlled hopper gate can prevent off-target fertiliser application into sensitive areas or neighbouring farms (Chok et al., 2016; Morton et al., 2016). New Zealand has approximately one hundred fixed wing aircraft applying fertiliser in the hill country environment. However, the industry's safety record is sobering with an average of almost two fatalities per annum (Grafton et al., 2009). Removing manual hopper operation from the pilot's in-flight duties allows the pilot to focus on the aircraft's operation, arguably making it safer for pilots to navigate one of the world's most dangerous aviation landscapes.

Replacing farmer-led decision-making, a key goal of many PA projects is far more fraught. Decision-making relating to fertiliser applications in New Zealand is often a cooperative effort between the farmer and an advisor, who is often a sales representative of the fertiliser company. While the removal of the farmer and the sales representative from decision-making processes might be an attractive outcome for the fertiliser companies from a cost point of view, it would not be without consequences.

From the researcher's perspective, the centralisation of intelligence may make decisions more 'correct', but Cayley (2009) suggests that a farmer who becomes reliant on standard methods loses their particular feel for the landscape. Rather than being a romantic or nostalgic judgement, Cayley holds that the relationship that a farmer has with their land influences stewardship of the land. Farmers who own and know their land are likely to do what is appropriate in the given circumstance, rather than focusing solely on profit, although it is likely that there would be a group of farmers for whom outsourcing this decision would be attractive. Ravensdown appeared to be cognisant of this risk in this case, with "Steve N" suggesting that initially at least, the centralisation of intelligence would simply be a validation of farmer decisions rather than a replacement for local knowledge:

> "It is giving them some confidence or giving them a view of their farm, or of their system or whatever your data is dealing with. For us, it is a farm. It is giving them more confidence. In reality, all it does is reinforce their view and give them confidence to make a decision that they already knew. And that's not to be underestimated. And I think that's where a lot of projects in this kind of area fall down."
>
> "For example, you may have… So, the [farmer] that we're working with, there is an area of his farm that he knew the general direction of it, but what the programme showed was that it was actually larger than what he thought. So, it is not something completely opposite to how they see the farm, but it fine tunes or adds value to that information that they already have in their hands." (YR4)

Widening the sphere of precision agriculture data capture and analysis

While there are many definitions of PA, it is generally agreed that the scope of PA application is site-specific, with the largest management unit being a zone within a single farm operation. Whilst a single measurement event may procure data from a large area, the application site is usually smaller zones. Consequently, a key opportunity for PA researchers is not just the improving sophistication of data analytics, or even the deluge of available data (Miller, 2010), Rather, it is the increasing availability of fine-grained temporal data on the location and biological features of site specific agricultural ecosystems. This is important, because until now, much of the emphasis of PA research has been on spatial interventions; the ability to return to the same spot again and again means we now need to think of PA as appropriate spatial *and* temporal interventions.

Prior to the arrival of big data and big data analytics, and the technologies that support them, PA data were principally produced and analysed within the confines of individual research projects. Where once data collection was directed at a plot, field or at most, the farm-level (Grafton, 2015), advances in data collection, storage and analysis mean that a new range of data points (Fraser, 2018) can be captured and applied at a local, regional and even global scale. The *Pioneering to Precision* PGP was designed to capitalise on these advances as P2P researcher "Virginia W" noted early in the project:

> Remote sensing technologies extend the precision agriculture concept by measuring at the larger scale, unbound by physical barriers such as fences farm boundaries, allowing data to be used both on-farm and beyond to inform, for example, judgements on sustainable agricultural policy. (YR1)

The P2P researchers were excited by the prospect that these data could allow an unprecedented view of farming systems from the "bottom-up". While massive databases have been available for decades, the fine-grained nature of these data is relatively new in PA, and the new techniques emerging for exploring these data are creating new opportunities to refine value from them. Importantly, the suite of remote sensing technologies employed in the P2P generate massive volumes of data, some of it in near real-time. The enormous volume of data results from the large size of the images being produced, which must be calibrated against large quantities of ground-reference data. The data's ontological characteristics, which are discussed in detail in the next chapter, also signpost a step change for PA that expands the scale and resolution of analysis and thus represent the emergence of a new data ecosystem for PA research. This is discussed later in this thesis.

Hence, remote sensing technologies were primarily selected for this study as they represent a suite of technologies that arguably generate the greatest volume of data in PA today (Kamilaris et al., 2017), at an unprecedented pace (Chi et al., 2016). Importantly, unlike first-generation PA technologies, where the application of data typically ends at the farm gate, data-rich PA allows the collection and application of data through the entire agricultural value chain, from cell to the final consumed product. This has implications for the drivers of value creation within the wider context.

Pioneering to Precision: initiation

The *Pioneering to Precision* PGP was framed around the notion that innovation is stable and predictable. The model of science employed was to develop a remedy, a techno-scientific industrial solution. Like all PGP programmes, the structure and monitoring of *Pioneering to Precision* was rigid and linear from the outset. Nathan Guy, Minister for Primary Industries (2013 – 2017) described the PGPs as "...the most highly monitored of the funding allocations within my portfolio" – (Nathan Guy, Minister for Primary Industries (2013-2017), personal communication).

MPI managed and oversaw all PGPs with advice and guidance provided by an independent Investment Advisory Panel (IAP), made up of primary industry and business experts. The PGPs had strong governance processes to give confidence and assurance in PGP investment and to ensure programmes had the best possible opportunity for success. For example, each PGP was governed by a programme steering group, which included at least two MPI representatives, and most had an independent chairperson. In the case of *Pioneering to Precision*, the steering group included two MPI representatives, and three representatives from Ravensdown. A Science advisory group was also formed to support the P2P programme steering group consisting of four Ravensdown representatives (including the Chair) and one independent researcher Dr Sam U, a hyperspectral sensing and imaging expert from the United States of America (organisation withheld).

Commissioned in October 2013, the P2P programme appointed two main science providers, Massey University, and the Crown-owned research organisation, AgResearch. The P2P researchers from Massey University, who are the focus of the case study for this thesis, were engaged to undertake fieldwork and develop nutrient models from aerially acquired hyperspectral imaging data. AgResearch's role was to provide specialist expertise in field data collection (botanical composition) and to lead the programme's research extension phase. In addition to the six-monthly milestone reports, a representative from each group (AgResearch and the P2P team) reported to the project's science advisory group.

Reporting, independent reviews and other measures also contributed to strong governance for PGP programmes and the PGP. Six-monthly milestones were set at the project outset and were monitored with quarterly progress reports submitted by Ravensdown, which were reviewed by the Investment Advisory Panel and MPI. The project's official timelines shown in Figure 13 were tight, and the project chronology omitted a major piece of critical work, learning how to *process* the data for analysis. The interviews revealed that the need to develop pre-processing techniques was not a surprise to the researchers, however the funders appeared to be ignorant of the massive challenge this would present. This oversight had implications for science-making in the project, both on timelines and for the researchers' workloads.

Timeline	2013	2014	2015	2016	2017	2018	2019	2020 - 2023	2024 - 2030	2030 – 2050+
	Project 1 Data collection & collation									
	Project 2 Algorithm development									
				Project 3 Software development & integration						
						Project 4 Technology Transfer				
Short term outcomes								Farmer adoption: 15% of farms		
Medium term outcomes								Farmer adoption: 40% of farms	Improvement in hill country productivity	
Long term outcomes								Farmer adoption: 98% of farms		Improvement in national economy

Figure 13 *Ravensdown's ambitious timeline for Pioneering to Precision PGP (P2P)*

Importantly, the timeline for the P2P project does not detail time allowed for data pre-processing and algorithm.

Meet the humans

I arrived at the P2P laboratory with preconceptions of roughly what the team's scientific division of labour would be. I knew the group specialised in pasture analysis and remote sensing. I anticipated the main actors to be field technicians, geo-statisticians, engineers, remote sensing specialists and domain specialists such as pasture agronomists and soil scientists, who would translate the results into useful information for farmers. Given the difficulties associated with user uptake in PA, I also expected to find research extension expertise within the group.

Again, my expectations were far off the mark. In an important departure from agricultural science tradition, the group contained no domain experts in agronomy, plant science or soil science. Nor were there any research extension experts. Instead, the primary actors were early career academics from the fields of remote sensing, statistics and engineering, and animal behaviour, some of whom expressed to me early in my study that they saw 'little value' in the endeavours of social scientists.[21]

---

[21] This view appears to be somewhat widely held in the precision agriculture sphere. A precision agriculture researcher at an international conference said I would be told that "there is no place for the social sciences in precision agriculture"

The P2P team: a community of people and machines

While the research team involved in the P2P, existed before 2014, the P2P project changed everything. With a financial commitment between the parties totalling $10.3m (Office of the Auditor General, 2015), the scale of the seven-year project (originally October 2013 – June 2020) was massive compared to the groups' previous projects. [22] The scale of the project and tight timelines meant the research needed to be expanded quickly. In response, a bright, young, trans-disciplinary community of humans and machines was formed. A complex fleet of machines was established, including those designed to capture data to feed into the new data ecosystem that was emerging. Many of the machines emanate from the remote sensing discipline. Other machines were engaged to take environmental measurements. Powerful computers were enlisted to process data and a large server was purchased to ensure that the borders of data storage were firmly erected to protect the P2P data stronghold.

While some human team members had some remote sensing expertise, the project was new territory for everyone; nobody had significant expertise in aerial data collection. Moreover, the PGP was by far the largest project the team had performed, in terms of budget, personnel and expectations. Interviews revealed the Director considered the P2P project as a final opportunity to create a world-leading technology before retirement. The young researchers saw the P2P as an opportunity to establish academic careers. An unsuccessful project outcome could have implications for their long-term success. Inevitably, the stakes were high for all actors involved. In terms of agricultural science, the door was firmly closed to researchers from the traditional agronomist and soil scientist fields. This was a conscious omission and had consequences for the project in terms of value creation.

Division of labour within the P2P team

The bibliometric analyses in Chapter 2 revealed the changing focus of much PA research toward assemblages of remote sensing and other ground-reference data. Observations of the P2P research team reveals that these shifts can have implications for the division of labour within science-making teams.

A key limitation of conventional precision agriculture laboratories is that they are typically enclosed spaces where data is strictly controlled by a specialist group of researchers tasked with refining value from predictable datasets. Observations of the P2P laboratory demonstrated that this limitation not only remains in the contemporary PA research laboratory, but the consequences

---

[22] Not all this budget was allocated to Massey University's P2P group, the exact details are not available due to commercial arrangements.

for science-making are exacerbated where researchers are working with novel technological assemblages, such as that found in the P2P project.

The P2P researchers were based at a university that has a formal policy of siloed physical spaces for research. Staff and students are not permitted to share offices. Small offices with closed doors mean the borders between students and staff, even between staff themselves stand firm. Post-graduate students rarely work alongside the university staff, even in a laboratory setting. The two groups rarely connect, except for monthly meetings, seasonal fieldwork and by email. In the early months of my ethnographic journey, I realised these physically enclosed spaces were not simply physical borders at NZCPA. In the absence of any physical laboratory to commune, students had little contact with researchers. Indeed, the researchers had little contact with each other. Consequently, the rigid siloes extended to the social structures and science making approach within the group.

The intractability of hierarchies

When I arrived in early 2015, the P2P project had started, and the group's hierarchy was firmly established. In an early interview, the Director described the organisational structure shown in Figure 14. This structure proved to be aspirational rather than actual. In practice, one may expect the line to be drawn between paid staff and students, however the social divisions were not so clear cut as in the P2P team. While the A Team mostly included University research and administrative staff, it also included a post-graduate student previously employed by the Director. The A Team socialised within and outside the work environment and dominated in status and political power. The B Team, mostly post-graduate students but included staff. Many of the B Team depended on the A Team to progress their studies or career, a matter not lost on some members of the higher caste.

The A Team had its own, complex internal hierarchy, with some actants tenuously positioned on the fringes, sometimes moving in and out of favour. Ideally, the relationship between lower and higher status academic castes is one of respect and paternalism. In practice, there was a great deal of tension in the relationship, which at the time of my arrival was held stable by the Director. Indeed, early in my tenure the Director reported he was working on eliminating the divisions, which he felt were harming the continuity of the group.

The divisions had by this stage become so entrenched that some members of the A Team actually referred to themselves as being members of the A Team. Moreover, members of the B Team also knew which group they found themselves. One student quipped: “A (Team) is for Assholes – B (Team) is to *be* there on time and *bring* your own lunch”. Despite the Director’s best efforts to quell the attitude from the A Team, these social divisions remained and continued to negatively impact the group’s performance both in the laboratory and in the field. This revolutionary science group had a new way of doing things, but established social problems remained.

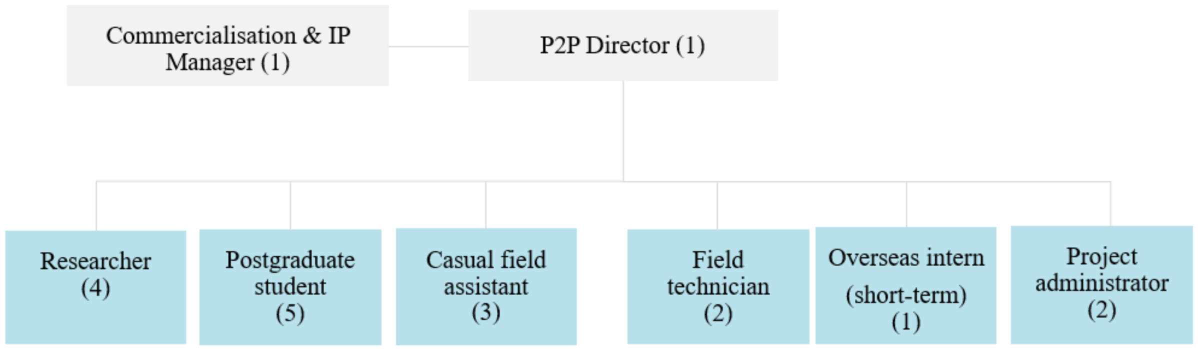


Figure 14 *The ‘official’ hierarchy of the P2P research team*

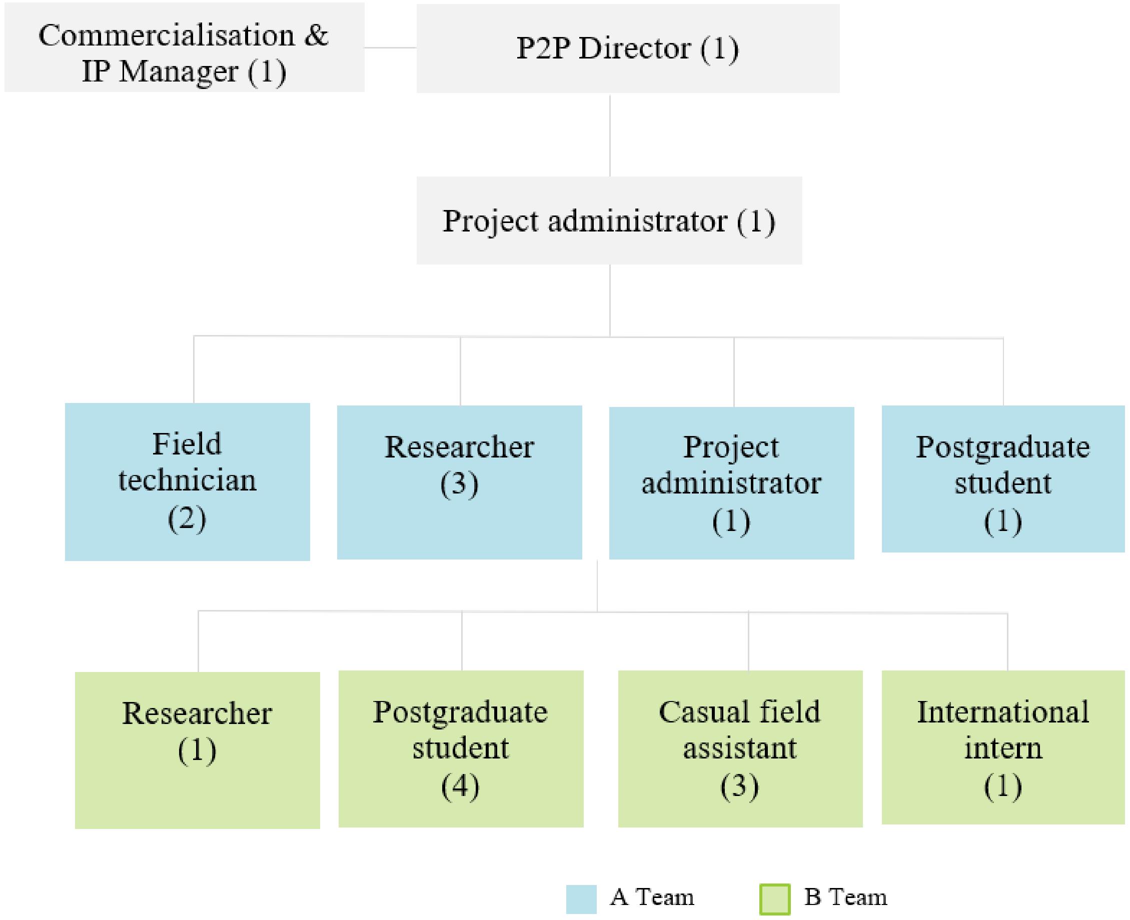


Figure 15 *The reality of the P2P research team hierarchy*

As the P2P investigated new frontiers in precision agriculture and remote sensing research, an existing workforce of scientists experienced in the field was simply not available. In response, the organisational structure of the P2P project is based around the apprenticeship model and as such is inevitably hierarchical, led by a leader who knows, with early career researchers who follow him, become like him. The team mostly consisted of a homogeneous group of early-career remote sensing specialists and students. While the construct of the P2P project was designed to analyse the data, by doing so, it also produced the researchers to analyse it through an apprenticeship model. My observations of and discussions with the students working on the project were mixed. Some of the students, particularly the student included in the A Team, enjoyed an experience similar to the exemplar provided by Hugo (2009) of how a masters student enters an intellectual community and is apprenticed into what it means to be a researcher. Other students felt that they benefited little from the apprenticeship model, as explained by “Joe A” (P2P post-graduate student) in an interview in Year 4 of the project:

> My experience was that I was brought on board and on Day 1 was collecting samples for the P2P. It had nothing to do with inducting me into the university. I had absolutely no assistance setting up workstations, and setting up the what would you call them… the accoutrements of work… I was basically told to go read up on the topic once that sampling was complete. And, the only meaningful interactions I had scheduled were other sampling events to collect data for the P2P, which I was informed would be available for me to use on my PhD and therefore I needed to be involved in the collection.
>
> Other researchers had special collection events and research trials where I and others were required to assist, but the courtesy didn’t seem to be extended to me. It felt like my primary function was to collect data for the project, not to do a PhD. Writing a PhD was my secondary function. After spending months collecting P2P data and being extolled as the only one who could collect it in some instances, I found out that data was of no use for my research.
>
> So, in hindsight, it felt like the PhDs were used as cheap labour, which allowed the project to be carried out at a lower cost than any other research institute would have been able to do it for. Once my P2P funding ended, I was left high and dry and had to finish things on my own. (YR4)

During fieldwork the researchers worked closely together for up to two weeks at a time, travelling, working, and living in close quarters. It was by following the researchers into the field that the fractures in the dynamics of the group became apparent. The social divisions between the P2P team were most prominent during the field work, where some of the B team felt they were simply being used as cheap (free) labour to collect data that did not have a lot of relevance to their studies.[23] Importantly, the Director rarely attended these events. There did not appear to be a clear chain of command or communication in his absence. I did not attend all the data collection events, however tales from the field offered by the students included suggestions of heated arguments, ignored safety concerns and unprofessional behaviour.

Relations appeared to be more cordial in the confines of the laboratory. In the laboratory the researchers mostly worked autonomously on their own piece of work. While I describe these fractures as siloes, it is a simplification of what I observed in the laboratory. The segmentation of the P2P project into isolated components meant each researcher was focussed on their own tasks. The lack of crossover meant there was little need or opportunity for the researchers to share their expertise with the students on the project. While some of the researchers, such as the researcher in charge of data processing made deliberate efforts to share their knowledge, others made little or no effort to interact with the students.

One senior researcher was concerned the largely inexperienced science team was working on individual tasks, making it difficult for the young researchers to cross-pollinate knowledge and for more senior researchers to help them. Furthermore, monitoring the progress of the tight timelines was difficult in the absence of regular collegial discussions according to "Virginia W" (P2P Researcher):

> I would say every three months we should perhaps meet and discuss what progress we have made and who needs a bit of help, who needs a bit of a kick up the arse, who's written what, where are we in light of our goals. Because we have to report every three months to Ravensdown so why can't we have a collegial discussion every three months before we put a bloody report in.
>
> I mean going around allocating tasks is fine if you are running a platoon of soldiers and you want some people go to flank the enemy, and want somebody to supply covering fire and all that but when everyone's doing their own thing doing in isolation to what other people are doing or knowing what other people are doing, I think it turns the atmosphere in the group is…. I would say it is, it's not good. I think it's quite a bit dysfunctional and I know all groups get a bit dysfunctional. (YR3)

That's enough about the humans for now, it's time to meet the real stars of the show.

---

[23] Most post-graduate students were not paid for field work, even when the work was not related to their studies, which engendered some disharmony.

The AisaFENIX: a new imaging tool to revolutionise agriculture

Now we come to the central character in this narrative - the device that everyone was talking about in my first meeting with the P2P research team. It turns out that Fenix was the short name for the AisaFENIX hyperspectral sensor. Massey University purchased an AisaFENIX sensor in December 2014. At the time, the University, and its staff were enamoured with the new acquisition as this media release demonstrates:

> Massey University has a new state-of-the-art aerial imaging tool in its precision agriculture arsenal that was first developed for military reconnaissance and space exploration. The $500,000 Fenix hyper spectral imaging system from Finnish company Specim was purchased, with Massey, as part of Pioneering to Precision—a $10.3 million Primary Growth Partnership (PGP) programme between Ravensdown and the Ministry for Primary Industries (MPI) to improve how fertiliser is applied to hill country. The seven-year programme, which began in June last year, is expected to generate $120 million a year in additional export earnings by 2030 and net economic benefits of $734 million between 2020 and 2050. [The Director] says the remote sensor will enable New Zealand to capture unprecedented levels of data about the nutrient content of large sections of land that may have been previously inaccessible.
>
> “This is a game changer,” [The Director] says. “It’s like turning the whole of New Zealand into a living lab, where you can observe exactly what is going on and describe it in greater detail than ever before.” With this tool we can overcome the sampling limitations by mapping whole landscapes, and provide data about what type and quantity of fertiliser is needed, assess pasture quality over the whole farm to help farmers determine stock carrying capacity and to locate the good quality pasture where they can fatten younger stock,” [The Director] says. “And there are opportunities for huge environmental benefits too.” “This is an extremely versatile and powerful technology. You could determine the exact number of kauri trees in a forest for example, and any diseased trees would stick out. There is also huge potential for orchard-based industries, like kiwifruit growers who could identify things like the PSA vine-killing disease, way before the human-eye could detect it,” [The Director] says (citation withheld).

Pioneering to Precision employed two nascent hyperspectral imaging tools; the proximal ASD and the airborne AisaFENIX. Unlike the ASD, which is a proven scientific tool both in the field and laboratory, the AisaFENIX had been mostly laboratory-bound before the PGP. Instead, researchers typically employed light, less expensive multispectral sensors for airborne research.[24] However, the spectral limitations of multispectral sensors meant research was mostly restricted to measuring monocultures. Thus, when the P2P researchers began their project, which was designed to capture highly variable landscapes, the full-spectrum sensor was considered ideal.

Looking remarkably like a simple metal box, our Finnish friend is compact and lightweight compared to previous multi-sensors, making it an attractive option for the P2P project. It has two cameras, both focused through a single fore optic to avoid the need for co-alignment of two separate optical systems. The sensor has a series of detectors which capture the spectral information for an entire row of pixels perpendicular to the direction of travel, with a spectral interval of 3.5 – 12.2 nM. This information is produced in near real-time at a very high spatial resolution, which could theoretically show localised distributions of nutrients at a range of spatial scales (Von Bueren & Yule, 2013; Yule, 2015).

The P2P researchers describe the technical specifications of the AisaFENIX in a scientific publication:

> A full-spectrum, pushbroom AisaFENIX hyperspectral imaging system was used in the study. The sensor measures upwelling radiance from 370 to 2500 nm as Digital Numbers (DN) with a spectral interval of 3.5–12.2 nm. The AisaFENIX sensor has a Field of View (FOV) of 32.2◦, as well as an Instantaneous Field of View (IFOV) of Remote Sens. 2018, 10, 1117 4 of 14 0.084◦ (citation withheld).

At first glance, this push-broom type sensor appears to be bounded in terms of performance and nature, reserved for laboratory environments. However, it was given wings in the P2P project when the researchers deployed the normally land-, and lab-bound AisaFENIX in a Cessna 206 aircraft to soar above the hill country farms of New Zealand. For the P2P data collection, the AisaFENIX was fixed within the Cessna aircraft fitted with a single antenna RT Oxford Survey+ Global Navigation Satellite System (GNSS), see Figure 16. An Inertial Measurement Unit (IMU) system was used to collect geospatial data for image registration and orthorectification.[25]

---

[24] "Hyperspectral imaging differs from multispectral imaging in the continuity, range and spectral resolution of bands" (Mulla, 2013, p. 365).

[25] Orthorectification is the process of converting images into a form suitable for maps by removing sensor, aircraft motion and terrain-related geometric distortions from raw imagery. This is one of the main processing steps for evaluating remote sensing data such as that from the AisaFENIX (Muller, 2020).

Figure 16 *Massey University's AisaFENIX installed in a survey aircraft.* Source: Aerial Surveys Ltd

Deploying the sensor on an airborne platform soon proved to be fertile ground for publicity. The AisaFENIX headlined the University's campaign at major events including the National Field Days, and the Central District Field Days in 2015:

> **Revolutionary agriculture imaging and more at Field Days**
>
> Images from a new state-of-the-art aerial tool with the potential to transform farming were on show at this year's Central District Field Days, alongside the latest in sheep fertility research and food innovation. Massey University Vice-Chancellor Steve

> Maharey says the tool enables farmers to manage nutrients, soils, and water for each blade of grass.
>
> The new imaging system brings precise scientific evidence to pastoral management and helps land-owners ensure production systems are sustainable. (citation withheld).

The researchers' new gadget didn't come cheap, and that was a problem for our university researchers. The original plan was to lease the AisaFENIX as the NZ$500,000 capital outlay for the device was beyond the normal means of a university pressed for funds. However, the sensor's early promise, and the potential to offset some of the expense from the PGP's planned lease costs convinced the university to purchase a unit outright. The agreement did not come without strings attached however, as one media release discloses:

> A requirement of the PGP contribution to the purchase is that the FENIX hyperspectral imaging system will also be made available to third parties, with priority given to Ravensdown, MPI and other associated parties during the life of the programme (citation withheld).

Importantly, the purchase meant the researchers needed to find commercial and research applications for the sensor beyond the P2P project, while their science team was occupied in the early phases of the group's largest ever research project. Some members of the research team felt the need to diversify to diversify to finance the purchase capital outlay for the sensor was a costly distraction as a Year 3 interview with P2P researcher "Virginia W" reveals:

> We've got expensive sensors and things like that and we've got a job to do for Ravensdown. But we also have to find other work and we also need to find work for the sensor, so it's got products and services to offer. My concern is that a lot of time has been spent on [the spin-out company] and that it's meant to be one day a week, but I would say it's disproportionate time being spent on [the spin-out company] and meanwhile we've got contracts to deliver for Ravensdown.[26]
>
> My view is that [the spin-out company] was commercialised too early without any real products and services. We need to actually go out and try using the equipment for people, see what we can see at a heavily discounted price, so that we can demonstrate an ability to do stuff. I think the idea that we're going to sell our cleverness and that people are going to trust us to go do something at an expensive rate, isn't going to be easy, especially in New Zealand where we notoriously have low research and development spending. (YR3)

---

[26] The spin-out company was formed by two of the P2P researchers in conjunction with the University in Year 2 of the project.

The need to find new applications for the University's AisaFENIX led to the purchase of some rather expensive additional 'accessories', in a case of the sensor enrolling others to create a world for it. While the university's main objectives was to install the sensor in a survey aircraft, the Director thought the sensor's versatility in terms of data collection platforms would broaden the potential opportunities so the university made further investment in a laboratory scanner mount (Figure 17) and tripod mount for landscape scanning.

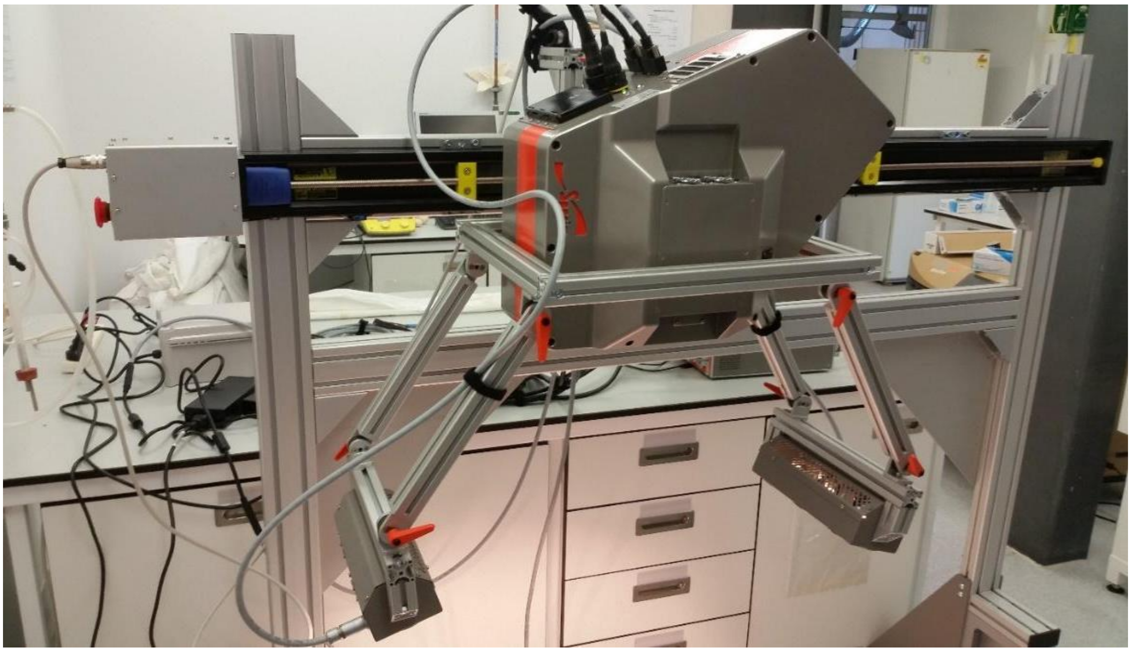


Figure 17 *Massey University's AisaFENIX sensor mounted in laboratory for proximal scanning*

While there had been previous studies using proximal and airborne hyperspectral sensors, their efforts appear to be largely confined to proximal studies. The first study to use the AisaFENIX from a fixed wing aircraft appears to be Rossini et al. (2013), who mapped water stress in an Italian maize crop in 2008. Usefully, the paper provides a detailed account of their method for capturing their airborne and ground-reference data. The methods for capturing ground-reference data were very similar to the approach employed later by the P2P researchers, including using the same model of proximal hyperspectral sensor and the use of reference panels. The authors also provided comprehensive information on the atmospheric corrections, data processes and statistical treatments applied to the data.

So, while the P2P researchers were not the first to use an AisaFENIX to sense crop biochemical characteristics, it appears they were first to employ an HSI sensor on a broad scale to measure pasture nutrients. Observations of the P2P researchers in action revealed they had little established foundational knowledge. Consequently, the researchers implemented a trial and error approach in the early stages of the P2P project. They also felt there were multiple pathways to a commercially viable solution as explained by P2P Researcher “Lance P” commented during an interview with Associate Professor Dr Brennon Wood and I in the early months of the research project:

> *Dr Wood: And some of those tools you're having to basically make yourself?*
>
> Lance P: Yeah, some of the tools. Because of the procedure, the chain of what I am following, no-one follows. That means, each person uses methodology, different ideas.
>
> *Dr Wood: Because there are still many different ways to solve the same problem?*
>
> Lance P: Exactly, there are so many ways.
>
> *Dr Wood: Yep, that's cool. The advantage of the PGP project is that you actually have quite a lot of ground data… though isn't it so you can trial some of those (ideas)?*
>
> Lance P: Yeah, since my PhD I deal only with the ASD. So, I have very good experience with the ASD data. So, I don't feel trouble analysing the ASD. It is very quick and now I can do just within a few hours, you know. But the Fenix, it takes still a lot of time because everything is a beginning, new to me.
>
> *Megan: So, it is a new challenge?*
>
> Lance P: “Yeah, yeah it is a new challenge. When we first analysed the first image last year it took us four months. Now it takes only takes less than a week. (YR1)

This extraordinary progress in terms of image processing time meant the project stood a chance of achieving commercial viability. This spurred the researchers on to further focus on making the data processing practices more efficient, with the main objective of automating as many processes as possible within a processing black box, called the HyperBox. This is discussed in further detail in Chapter 6.

Leaving the laboratory

Precision agriculture research thrives on the control of variables. After all, that is why researchers have those enigmatic spaces called laboratories, where researchers can control variables, or at least where they can be identified, measured, and recorded. In contrast, field trials present a myriad of challenges for researchers. Those brave souls embarking on field trials carefully design their experiments to identify and hold constant as many variables as possible to allow the accurate observation and measurement of the independent variable. Indeed, this approach has been the trajectory of agricultural science over the course of the twentieth century. Agronomic field trials are often conducted in small plots, where variables such as slope, aspect, soil conditions and plant populations are highly managed. For many experiments, a critical variable is plant species. It is no coincidence that agronomic field trials are dominated by monocultures, where the plant species is constant and homogeneous. *Pioneering to Precision's* aspiration to account for the nutrient content of soils across the notoriously diverse hill country environment is unparalleled.

Taking flight

In terms of remote sensing, the project represented an important shift from the laboratory-like conditions of proximal sensing to airborne data collection. Previously, accurate geolocation of data collected was only possible if permanent markers were in place, which made temporal data collection costly and often inaccurate. Consequently, the measurement of environmental dynamics has been difficult, expensive, and cumbersome to obtain. A key objective of the P2P project was to accurately map soil properties at high resolution. Until now, this has involved costly manual data collection and scientific methods designed to tease information from scarce observations.

While the AisaFENIX showed promise as a proxy for measuring soil nutrient levels of bare soil (citation withheld), the presence of plant material makes this impossible with the AisaFENIX as the sensor is unable to penetrate the plant canopy. At project outset, the funders decided the main objective was to determine nutrient levels in soil rather than use tissue analysis to drive decision-making. It appears the funder's bias towards soil testing had consequences for the project methodology and ultimately the outcomes of the project. To overcome this limitation, the P2P researchers made the brave decision to use measurements of plant reflectance as a proxy for soil nutrient levels, meaning that the researchers would also need to account for the relationship between plants and soil nutrient levels, i.e. the AisaFENIX data would be a *proxy for a proxy*. While the collection of the small, ground reference data was extensive in the P2P project, the end goal was to largely *replace* the small data collection with algorithms that model the relationships between the aerially acquired AisaFENIX data and the phosphate levels in the soil, see Figure 18.

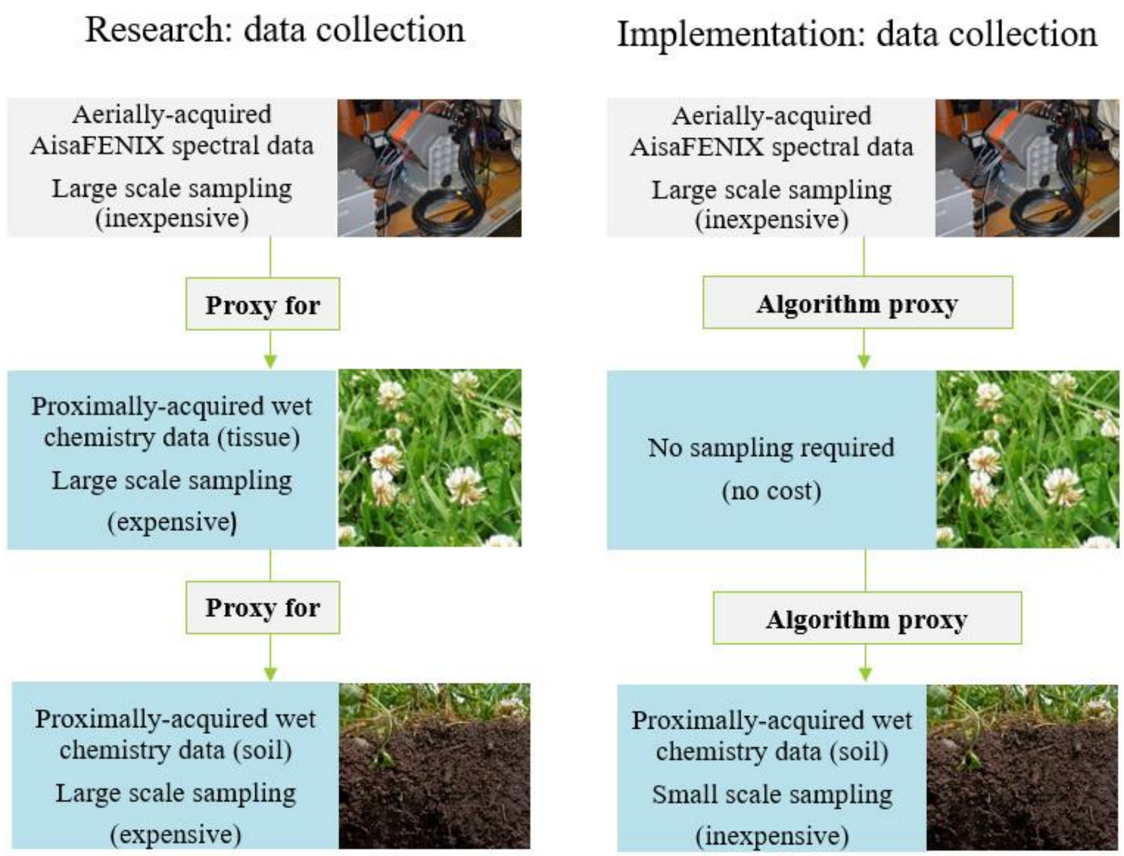


Figure 18 *The small, expensive data required for the research stage of the P2P project was extensive, whereas algorithms would largely replace small data collection in the implementation phase.*

Thus, for the project to be successful would require two key outcomes; (a) to determine the relationships between plant reflectance and soil nutrient levels, and (b) to model the relationships between plant reflectance (AisaFENIX measurements) and the proxy (plant biochemical characteristics) for soil nutrient levels (which may or may not relate to plant nutrient levels).

Two distinct datasets were created to facilitate this two-way data analysis and ultimately support algorithm generation. The first was based on the aerially acquired AisaFENIX data and supporting metadata. The second dataset represented the ground-reference data and contained a hybrid of ground-reference data measured in the field and laboratory. The P2P researchers treated the ground-reference data as their *control* for the AisaFENIX data, which was considered the independent variable. This represented an interesting shift in what was acceptable error for agricultural science; previously, field-acquired measurement has often been viewed as carrying inherent variability, which required calibration against laboratory samples. In this case the researchers aimed to eventually use the proxy to *replace* laboratory testing. These decisions had profound effects for the project. Employing a proxy for a proxy dramatically increased the workload required to deliver a scientific outcome, and it brought risks concerning validation. The impacts of these decisions are discussed in detail in the next chapter.

Data dependence in a community of machines

The researchers' scientific publications relating to the P2P project almost exclusively focused on the AisaFENIX dataset because in terms of scholarly novelty, the hyperspectral data are the 'primary' data source. However, observations of the P2P researchers in action suggested that the situation was more complex than this. Indeed, closer analysis revealed an assemblage of machines delivering ground-reference data played two key roles crucial to the project's success. First, large volumes of geolocated ground-reference data were needed to explain relationships between the AisaFENIX spectral readings and nutrient concentrations in plant tissue, with the defined goal to guide and inform fertiliser decision-making (citation withheld, Ministry for Primary Industries, 2014). Second, ground-reference data were needed to explain relationships between plant nutrient content and soil nutrient content in the analysis phase of the research. Other than the purchase of the AisaFENIX, the collection of the small ground reference data was one of the major costs involved with the project. The scale of the collection of carefully geolocated data representing numerous factors was unprecedented in hill country research in New Zealand. The funder's representative "Steve N" felt the ground-reference dataset was of value not only in relation to the AisaFENIX data, but additional value could be independently refined from the ground-reference dataset itself:

> Steve N: It is massively valuable to us. So we will be discovering things in the dataset ten years from now because some things we don't know now but we will know in five years and because we've got that data and the integrity of that data we kind of know it might be important in the future but, so the integrity of the data is sound so that allows us to keep going back to it with that new knowledge. (YR4)

For example, one post-graduate researcher in the P2P was repurposing tissue wet chemistry tissue results and proximal ASD readings to test the accuracy of a new remote sensing technique. Furthermore, the soil samples were retained from the research, meaning that samples could be re-tested in future if researchers wanted to examine other parameters.

Interestingly, with no background in pasture science, some P2P researchers believed they simply needed to measure the soil and where it was lower than desired ranges, fertiliser would be applied and thus the productivity of the pasture plants would be increased. Consequently, the researchers' focus was squarely on the soil nutrient levels rather than the plant and its many other needs. The science team made little or no effort to understand the plants involved, or their importance in the production system. Instead, plant measurement was considered simply as a conduit to understanding soil nutrient levels and indeed plant and animal productivity. In this sense, the P2P researchers framed plants as neutral intermediaries, rather than as mediators. The apparent lack of attention to the plant sensing component of the work, coupled with poor understanding of their target species would later cause issues in deriving knowledge from the data.

Data collection

The data collection strategy for the *Pioneering to Precision* project was different to a typical traditional PA research approach. The concurrent collection of airborne hyperspectral imaging data and ground-reference data was designed to fulfil two functions. The first was to facilitate the crafting and calibration of algorithms to represent relationships between soil nutrients and AisaFENIX readings. Second, and importantly, it was also designed to calibrate the airborne tool itself, so eventually few or no ground reference data would be needed when collecting data and the sensor readings would stand alone as an equivalent of measurement.

Data were collected from eight test farms over different seasons and data collection was guided by the project protocols. More than seven thousand tissue and soil samples were sent to a laboratory for calibration. The farms were selected to represent a mix of climate, soil, and vegetation types, with three farms in the South Island and five farms in the North Island. The experiment was designed so each farm was sampled at least once in the autumn and once in spring, although one farm was sampled five times. Every sampling event included the collection of aerially acquired AisaFENIX data, atmospheric readings to inform the atmospheric correction processes and the following ground reference data; 400 soil samples, 400 tissue samples, 400 ASD spectral readings, 400 plant species and plant cover readings. All measurements were tagged with the geolocation determined by the RTK-GPS. The AisaFENIX sensor was flown in a fixed-wing aircraft at 600 metre sampling altitude, resulting in a pixel size of approximately one metre (citation withheld).

Table 11 *Sampling event dates for the PGP farms up to March 2017*

Farm	**Location**	**Survey Dates**	**Total sampling events (=*n*)**
Farm 1	North Island	2013, Dec / 2014, Jan, May, July, Oct.	5
Farm 2	North Island	2014, Dec / 2015, May	2
Farm 3	North Island	2014, April, November / 2015, May, Oct	4
Farm 4	North Island	2015, Oct / 2015, April	2
Farm 5	North Island	2015, Nov / 2016, April	2
Farm 6	South Island	2015, Nov / 2016, April	2
Farm 7	South Island	2015, Nov / 2016, May	2
Farm 8	South Island	2015, Nov / 2017, March	2

Initial farm visit

Each of the eight test farms received an initial visit from members of the P2P team. I attended the farm visit to Farm 3 in Year 1. An administrator from the P2P team made initial contact with the farmers. The purpose of this visit was two-fold. First, several P2P personnel, usually led by a P2P Field Technician, were taken around the farm and to log safe tracks, look at the lie of the land, identify areas which will be sampled. Second, according to the project protocols, the members of the P2P field work team would meet with each farm manager/owner to gather information on the management of the farm (such as fertiliser information and history, weather data etc.), identify hazards and explain the approach to data collection. The full details can be found in Appendix E.

Pre-flight site identification and selection

After the initial site visit, the P2P researchers undertake planning for where each of the 80 test sites might be placed on each farm. The desk-top assessment was based on farm physical information, e.g. slope, aspect, soil type. Stratified random sampling was based on physical information including aspect (North facing = 300-60 degrees, East- or West- facing = 60-120 and 240-300 degrees respectively, South facing = 120-240 degrees), slope criteria (grouped into 0-8 degrees, 8-16 degrees, 16-25 degrees and > 25 degrees), and soil classification.

When selecting sites, the researchers also made their decisions based on the pasture species present (their preference was for ‘homogeneous areas’). They avoided selecting areas that are free from thistle (where possible) and avoided stock camps (due to the artificially elevated nutrient levels in these areas) and stock tracks. Figure 19 is taken from a map of suggested locations for the ground reference data collection for the marking out exercise. Note the difficulty in making sense of the scale of the image without textual clues; this emphasises the importance of accurate geolocation information prior to the marking out operation.

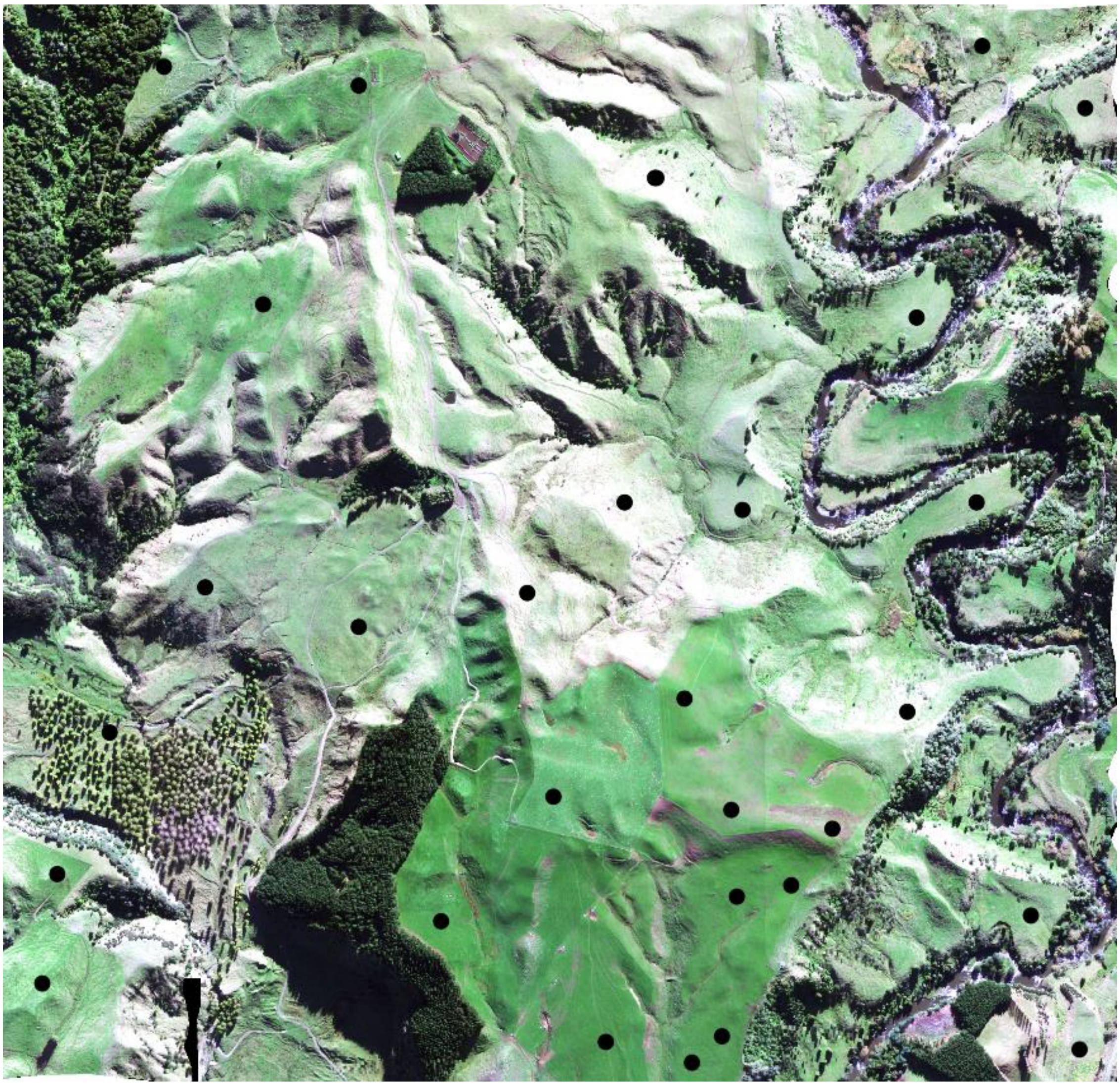


Figure 19 *Suggested ground sample locations from an area on a research farm.*

Marking out events

In the laboratory the researchers mostly worked autonomously on their own piece of work. While some of the researchers from the A team socialised together, moments of collaboration were rare for many of the P2P researchers. Fieldwork on the other hand was a totally different matter. In the field, researchers worked closely together for up to two weeks at a time, travelling, working, eating, and socialising together. Only by following the participants into the field could I really understand the role of data collection in this project and expose the dynamics of the group. Hence, in April 2015, I joined a group of researchers for a marking out event at Farm 5 in the Wairarapa region.

Observations of the marking event at Farm 5 revealed these affairs were somewhat chaotic. Most of those involved had never stepped foot on a hill country farm so tackling the terrain and operating the machinery required to navigate it proved challenging.

The purpose of the marking out event was to install geolocated ground-reference pegs prior to the airborne and ground data collection events. Before the researchers and I arrived at site, the P2P research technicians had selected 80 ground-reference sites in a desktop exercise using aerial photographs and topographic map information to guide their decisions. At each start of the event, two of the researchers, a student, and a field technician, positioned and assembled an RTK-GPS base station at a preselected high point on the site, see Figure 20. The P2P research team used a Leica Real-time kinematic (RTK) positioning unit to accurately determine each site position. RTK-GPS is a satellite navigation technique used to enhance the precision of position data derived from satellite-based positioning systems, providing up to centimetre-level accuracy (Wanninger, 2004). The RTK-GPS unit employed in the P2P project used measurements of the phase of the signal's carrier wave in addition to the information content of the signal and relied on a single reference station which was positioned at a high point on the farm by the P2P researchers. The RTK-GPS unit employed in the P2P project had three-centimetre accuracy.

Figure 20 *Setting up the Leica RTK-GPS base station*

Once the RTK-GPS was working, the researchers split into two teams of four people who used the system to locate the 80 pre-selected sites. Each team had at least one paid staff member, either a field technician or a researcher, and the remainder of the teams were usually populated with post-graduate students or casual field workers. Once the teams arrived at each correct reference site, they drove in five permanent marker pegs at each location, see Figure 21.

Figure 21 *P2P researchers geolocate permanent marker locations using Leica RTK-GPS*

During the setting out, it emerged that unlike the early trial at another location, which had a highly accurate Digital Elevation Model (DEM), the DEM employed for Farm 5 was often inaccurate. So, when the researchers arrived on site, some sample sites did not meet the design criteria. For example, the slope and/or aspect were very different to that marked out on the plan. On other occasions, the sites identified in the desk-top exercise were covered in thistle or other weeds, were not practical or were deemed unsafe. These issues forced the researchers to relocate several sites. This revealed a process dilemma for the group; while a researcher on site provided a recommendation on how to deal with the problem, the technicians all had different ideas. They believed the Director, who was absent, was the sole decision maker. Team members hesitated and debated as to where sites should be located, with the main point of contention being the ability of the new sites to be a fair representation of field conditions.

No mobile phone coverage meant they could not contact the Director for guidance. It appeared that the default position for one team was to relocate the site to the nearest flat area, whereas the other group took a different approach. They decided that if the pre-selected site were to be on a farm track, the new site should be located above the track to mitigate any run-off effects. On at least one occasion, the new site location was placed in pasture on a slope that matched that of the original target site but was on a different region of the farm. This inconsistency in approach was addressed for later farm mark-outs with a rule that the site should indeed be located above the track and the site should be in the same category. After each site was geolocated using the RTK-GPS, a marker peg was driven in and sequentially numbered. This numbering system was then employed to identify all data generated from that peg, such as soil samples, botanical composition, and tissue analysis.

The slope (using an Abney level) and aspect readings (using a compass) were recorded for each site. A GPS-located camera (Canon Powershot SX230 HS) was used to take scene photos for each site and a 'plot' photograph next to each ground-reference peg using a 0.25 m^2 quadrat, see Figure 22[27].

Figure 22 *A quadrat set out next to a ground-reference peg. The date and site reference number provide photographic evidence of species composition and plant cover for future analysis*

Data collection (sampling) events

Sampling events involved collecting both airborne and ground-reference data within a very short timeframe (usually on the same day). The data collection, processing, and transformation phases of the project generated an enormous volume of data, not only from the AisaFENIX, but also from supplemental (dependent) data and metadata, which add depth and complexity to the dataset, which by Year 4 had exceeded 30TB in size. Figure 23 represents data sources and data journeys produced during an interview with a P2P researcher which highlights the three main data collection platforms for the P2P project; field data (proximal), aerially acquired data and laboratory.

---

[27] An Abney level is a hand-held device used to read the angle of an object.

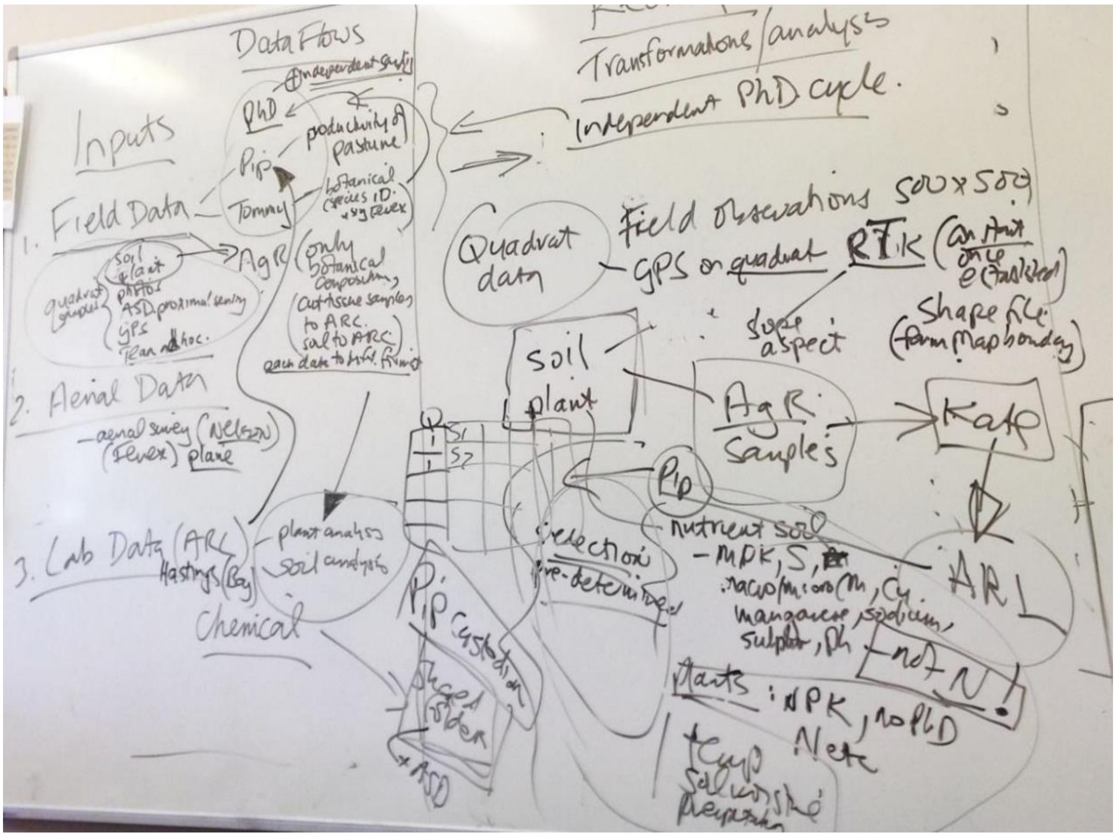


Figure 23 *Representation of data sources and the data's journey*

In practical terms, the collection of the ground-reference data was an enormous, labour-intensive undertaking and involved a community of machines and humans gathering geolocated data. With the exception of the spectral data collected using the ASD Field Spec, the proximal data sources were direct measurements.[28] The vast scale of the small data collection for a big data project is ironic, given that the idea of remote sensing is to reduce or eliminate the need for direct measurement. However, as I discuss in the next chapter, this signals the evolution of data collection to an all-data dataset. These measurements were collected by the P2P researchers to calibrate the remote sensing tools and to account for variables that may influence the reflectance results. The ultimate outcome of the project would be for the AisaFENIX to eventually supplant many of these proximal tools. The data are complex and arrive in various ways, types, shapes, form, and sizes and at different stages in the science process. Researchers must work with data that arrive in new (and old) forms (Kitchin & Lauriault, 2015), see Table 12.

---

[28] The wet chemistry soil analyses and tissue analyses may be considered to be indirect by some commentators.

Table 12 *Main data sources for the P2P project*

Data source	Device	Data collection platform	Data 'size'
Hyperspectral reflectance data	AisaFENIX	Airborne (aircraft)	Big with atmospheric & geometric corrections needed
Hyperspectral reflectance data	ASD Field Spec	Proximal (pedestrian)	Big with no atmospheric or geometric corrections needed
Wet chemistry soil nutrient data	Soil corer	Proximal (pedestrian)	Small
Soil type data	Soil corer	Proximal (pedestrian)	Small
Wet chemistry plant tissue nutrient data	Cutting shears	Proximal (pedestrian)	Small
Species composition data	Quadrat	Proximal (pedestrian)	Small
	Visually acquired via human eye		
Location data	RTK-GPS	Proximal (pedestrian)	Small
Location data	Inertial Measurement Unit	Airborne (aircraft)	Small
Location data	Global Navigation Satellite System	Airborne (aircraft)	Small
Location data	RT Oxford Survey+ Ltd	Airborne (aircraft)	Small
Site reference data	Flipchart	Proximal (pedestrian)	Small
Atmospheric parameters	Not disclosed	Proximal (pedestrian)	Small
Topographic data (slope)	Abney level	Proximal (pedestrian)	Small
Topographic data (slope)	Existing topographic maps	Existing	Small
Aspect data	Compass	Proximal (pedestrian)	Small

Ground-reference data for each of the eight trial flights were collected from the eighty pre-selected sites, with five measurements being taken near each site. The ground-reference dataset includes proximally acquired hyperspectral data gathered using a GPS-located ASD, and 'small data' gathered from pre-determined field sites using traditional methods. These small data included visual estimates of species composition acquired in the field; wet chemistry tissue and soil samples, slope and aspect readings, and RTK-GPS geolocation points. On the morning of each survey, the researchers collected ground data to assist with atmospheric corrections. This data was later employed in atmospheric corrections before the data were mosaicked into a single geo-rectified image.

The researchers were familiar with the ASD as the device had been employed in earlier trials to measure pasture characteristics. The researchers used the ASD because it offered control over some variables - a little lab in the field. The ASD's shroud eliminates light, removing the need for data processing relating to atmospheric corrections. Also, the position of all ASD measurements were geo-located using RTK-GPS, providing reliable, highly accurate spatial accuracy. For the case study, the ASD was used to sense the surface of the quadrat at each measurement peg shortly before the botanical composition was assessed and plant and soil samples were removed for wet chemistry analyses, see Figure 24.

Figure 24 *P2P team members use the ASD to collect proximal hyperspectral data at Farm 5*

Tissue samples were removed from a 0.25m^2 quadrat placed on the north east corner of each peg (the same area as the ASD readings were taken). All plant tissue within the quadrat was cut using hand shears near ground level; the tissue from each quadrat was bagged and labelled before being placed in an insulated storage bin for transport to the testing laboratory. Next, the researchers used a stainless-steel soil corer (approximately 20 mm diameter) to remove nine evenly spaced soil samples from each quadrat area, see Figure 25.

Figure 25 *Ground-reference quadrat where plant material has been cut and soil samples removed*

The researchers then placed the soil samples from each quadrat area into a labelled plastic bag for transport to the University's cool stores. The samples were weighed, dried and before being analysed using wet chemistry laboratory techniques. Retaining these samples in the laboratory meant that further testing could be conducted on the samples if necessary.

Figure 26 *Combining soil samples to be sent for wet chemistry laboratory analyses*

An AgResearch agronomist was employed to visit each peg site and collect visually acquired information on volumetric percentages of dead matter, legumes, weeds, grass stem, grass leaf and bare ground. The agronomist identified and recorded the first, second and third most abundant species present for each peg. Importantly, the same agronomist performed most of the surveys to promote consistency between sampling events. The agronomist represents one of the variables in the field work because while the assessments were guided by visual aids, the assessment was somewhat subjective. Data were then collated on a spreadsheet that included the geo-location peg number for future analyses.

Conclusion

In this chapter we met the actants in the P2P study and departed the laboratory to follow the researchers and their tools into the field to observe what and how data are collected. While I originally thought the focus of effort and attention would be on the AisaFENIX, I found the P2P project is populated with an unexpectedly complex assembly of humans, data, and machines. In this chapter I describe the data collection for the P2P project and describe some of the dynamics of the P2P science team. I found the division of scientific labour evident in the field was different to the laboratory where researchers worked largely individually. In the field, the researchers had to work with each other, and it didn't always go well.

The researchers continued to honour the site-specificity of ground-reference data, which was gathered in vast quantities. This chapter has detailed the construction of site-specificity as data collection by the P2P's community of humans and machines. This group of researchers were collecting enormous quantities of ground reference data. In an approach reminiscent of their predecessors in agricultural science, the researchers hoped that this complex community of ground reference data would let them account for as many variables as possible, which would in turn lead to the more accurate algorithms they would create in the laboratory. In the next chapter we return to the University laboratory with this vast collection of data to explore the science-making involved in aligning these proxies.

Chapter 6: The all-data revolution - new data aspirations for precision agriculture

Introduction

Hyperspectral sensors measure light. They measure the electromagnetic radiation originating from the sun and reflected back from objects in hundreds or even thousands of narrow and contiguous bands (in wavelength) across the electromagnetic spectrum (Chang, 2013; Yule et al., 2015). From that light measurement, researchers try to explain another portion of sampled data that cannot be easily collected, such as soil nutrient levels.

In Chapter 3 I found that until recently, the scientific method in PA was strongly influenced by the concepts of parsimony and efficiency as data were still sparse and arrived late in the science process (Gauch, 2012). Those employed in theoretical agricultural science required similar skills to empirical research, save for the increased influence of statisticians who were needed to generate and interpret models. I discovered that before the arrival of Global Navigation Satellite Systems (GNSS) in the 1990s, many agricultural fields were researched and managed as a uniform resource (Lark, 2001; Whelan & McBratney, 2000). Scientific progress has largely centred around small plot trials conducted by agricultural scientists using experimental designs whose findings could be extrapolated across a field with apparently similar characteristics (Lark, 2001). Executing temporal experiments in the field was fraught with difficulty.

Agricultural research became more complex in the 1990s when the arrival of GNSS facilitated the geolocation of agricultural data beyond the use of stereo-plotters(Larsen et al., 1994; Young & Carter, 2013), making it possible to accurately determine and continuous record the position of an appropriate receiver (Neményi et al., 2003). The impact of the arrival of GNSS on PA research was two-fold. Firstly, researchers could accurately geolocate in-field measurements year on year. Second, these new positioning technologies allowed researchers and farmers to reliably geolocate the application of inputs, with decisions informed by the geolocated data collection. The combination of geolocated data collection and application technologies spurred researchers to consider the adoption of an alternative hypothesis that spatial variation of soil and other characteristics should be accounted for in management (Lark, 2001). In Chapter 2, the bibliometric analysis revealed that while hyperspectral sensors were available to researchers in the 1990s, PA researchers did not start exploiting these sensors until recently. Chapter 2 also supports the view of Sonka and IFAMR (2014) that the influence of data-intensive technologies in PA is increasing. However, for big data to have a truly transformational effect on agriculture, researchers need to develop new products and services that rely on the data themselves rather than just being an input into existing business models (Fraser, 2018; Huberty, 2015; Pesce et al., 2019).

In Chapter 5 I introduced the actants in the P2P study and departed the laboratory to follow the researchers to observe what and how data is collected in the field. This chapter opens with an examination of the characteristics of the AisaFENIX data that was collected in the field and provides an insight into the implications of an emerging data ecosystem for scientific enquiry in the field. I follow the PA researchers back to their laboratory and chronicle their efforts to prepare the AisaFENIX and its supporting data for analysis. I explore the roles of the AisaFENIX data and its supporting data and find that the arrival of the AisaFENIX does not mean traditional data sources are ignored; indeed, I discover that it is important that the AisaFENIX data and their supporting data can talk to each other. Observations of the P2P study show that the processing stage was more complex and time-consuming than planned for. Finally, the findings suggest that data processing is taking more prominence in the scientific process, and as a result a new scientific division of labour in PA is also emerging.

Ontological characteristics of the *Pioneering to Precision* dataset

While remote sensing is relatively commonplace in PA, P2P represents a new era, the-all data revolution (Fraser, 2018). During my 30 months following the P2P project, the researchers often described their research as a ‘big data project’. However, closer analysis reveals that data arrive in many new and old forms. Unlike data typically gathered from previous technologies, the data are generated across multiple sources and platforms. I invested much effort in mapping the journeys of the P2P data, guided by the work of Bates, Lin, and Goodale (2016). The presentation of the ontology of data in this thesis follows ontologies from information and data in computer science as described by Kitchin and McCardle (2014), which are an extension of the philosophical use of ontologies. In philosophy, ontology is the basic description of things in the world, whereas in information science, an ontology refers to an engineering artefact, constituted by a specific vocabulary used to describe a certain reality (Fonseca, 2007). The use of ontologies to describe data characteristics here is deliberate. Instead of using taxonomies as I did in Chapter 2, I use ontology to not only study ‘what there is’, but also to encompass the most general features and relations of the entities which *do* exist as described by Vázquez (2018). This work was performed with the intention of illuminating the evolving ways data are produced, processed, visualised, and socialised.

Central data characteristics of the *Pioneering to Precision* all data dataset

A question that this thesis seeks to respond to is: What are the characteristics of new PA technologies? To help understand these technologies, I sought to identify ontological characteristics of data produced by emerging assemblages and explored how they differ from traditional PA data. An examination of the P2P dataset reveals a dataset with different ontological characteristics to its traditional small-data counterparts. While descriptions of the P2P project by the co-funder and the P2P researchers typically focused on the use of the AisaFENIX (Ravensdown Fertiliser Co-operative Ltd, 2018), closer analysis demonstrates that the project's data ecosystem was far more complex. Indeed, the dataset is a good example of the 'all data revolution' described by (Fraser, 2018). Viewing the P2P dataset as a single entity allows some central data characteristics to be described, see Table 13.

Table 13 *Central characteristics of the dataset*

Data characteristic	**Characteristics in terms of P2P dataset**
Spatial intensity	Some, but not all data are spatially intensive. The AisaFENIX data are fine-grained in spatial resolution, aiming to be as detailed as possible, and uniquely indexical in identification
Spectral intensity	Data from the AisaFENIX and ASD Field Spec are spectrally intensive compared to multispectral sensors
Speed of data collection	Some data created in near real-time (high velocity), e.g. AisaFENIX and ASD Field Spec
Variability of data	While data are structured, they are highly variable, including in spatial intensity
Data are layered	Some data are layered from single or multiple data source(s)
Veracity of data	Veracity of some data is imperfect; some data may have biases, noise, and abnormalities
Cost of data collection	Aerially acquired hyperspectral data are cheap, the ground-reference data is expensive
Versatility of data	Data are versatile (flexible): metadata and geo-tagging means that data, including the ground-reference data can be repurposed because they hold traits of extensionality - can add new fields easily if ground-reference data are collected
Temporality of data	Accurately geo-located data can be compared temporally
Data visualisation compatibility	Digital numbers can be worked into an image
Propensity for model error	Errors in models can be exacerbated by extrapolation
Scope of data collection	The scope of data collection from the AisaFENIX is vast; it strives to capture entire plant populations (*n* = everything)
Ability for data repurposing	Exhaust data may be of value for other applications
Dependence on processing	The aerially acquired AisaFENIX data require atmospheric, temporal, and geometric corrections before data can be analysed
Mutability	The AisaFENIX data are mutable mobile. Other data in the dataset are mutable immobile.
Scalability of data	Data are scalable (can expand in size rapidly);

The resulting data ecosystem is radically different to that to which agricultural scientists are accustomed. As listed in Table 9 of the previous chapter, this ecosystem includes both small data and data that meet the definition of big data as detailed in Kitchin (2013, p. 262). The dramatic change in volume, structure, and velocity of the aerially acquired AisaFENIX data, fulfils many of the characteristics associated with big data.

In this new data ecosystem data arrive from multiple sources and platforms and may have already undergone some processing before being received by the researchers, e.g. wet chemistry soil tests performed at an external laboratory. Observations of the P2P researchers established that each data source had different benefits and challenges that needed to be accounted for. This was especially the case for remote sensing where the data are expected to be relational. This suggests that small data collection will continue to be a vital component of the research landscape in the foreseeable future, but their position and role within it are changing.

I also discovered that small, traditionally acquired data played a crucial role in the calibration of airborne tools and in the crafting of algorithms and as such as playing a key support role in generating academic knowledge from big data technologies. Moreover, this study identifies that some small data can behave and be analysed as big data, where quality supporting metadata is present and appropriate data structures are in place. This was the case where over 7,000 accurately geo-located proximally acquired soil nutrient tests were used for both temporal and aggregated studies in the P2P project.

While data aggregation, and regional and even global modelling have been commonplace for decades in some fields of enquiry, such as meteorology (P. N. Edwards, 2010), the spatial variation of PA data was such that models generated from the data were applied at the field or farm level, rarely being of influence beyond the farm boundary. Most PA research has focussed on the mapping of soil and yield, with a view to adjusting on-farm management to maximise production (M. A. Oliver et al., 2013). Importantly, data collection and analysis has typically centred on calibrating a model. What this study of the P2P project shows, is that for some of these nascent technologies, such as the AisaFENIX, the data collected are not only needed to calibrate the model, but to calibrate the tool as well.

Aerially acquired hyperspectral data

Unlike multispectral sensors, Hyperspectral sensors sample a large frequency range. Importantly, not only are there substantially more geo-referenced sample points per hectare using HSI sensors, the data are also multi-layered; so each data location yields hundreds of layers over the VIS, NIR and SWIR, enabling biochemical analysis for any target (ASD Inc, 2016)[29]. Consequently, the data sets generated by the HSI sensors are enormous, complex and arrive early in the science

---

[29] Refers to the VIS (visible), NIR (near-infra-red) and SWIR (short-wave infrared) portion of the electromagnetic spectrum.

process, creating a completely different data ecosystem to that in which PA researchers are accustomed.

The P2P researchers deployed the AisaFENIX on an airborne platform to measure light over the broad landscape. With each pass of the aircraft, the sensor generates ‘strips’ of data that map about 4-500 metres wide at ground level. The AisaFENIX used by the P2P team has a spectral sampling interval of 3.4–7.3 nm from VIS and SWIR region (citation withheld), and the data is presented in 448 layers (a layer for each spectral band). Together, these form a three-dimensional data cube, which can be interrogated to generate information for a variety of purposes (Chang, 2013; Mulla, 2013; Yule, 2015), see Figure 27.

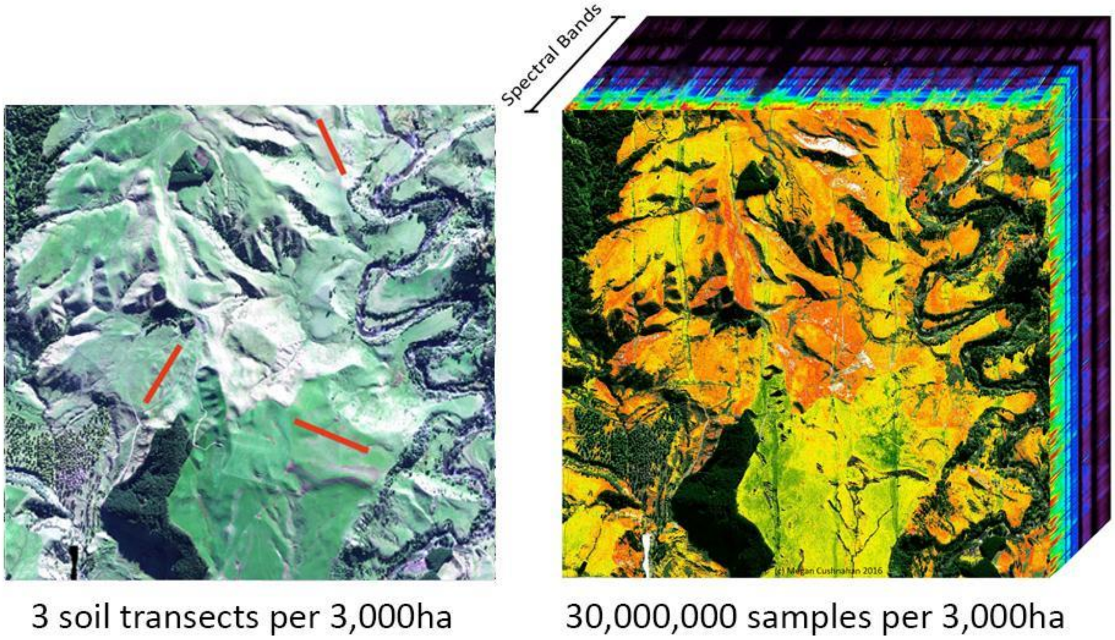


Figure 27 *AisaFENIX collects an enormous amount of high-resolution pasture nutrient data compared to existing approaches*

The data produced by the aerially deployed AisaFENIX differ greatly from proximal technologies. When deployed aerially, the researchers found that the AisaFENIX produces highly versatile, geolocated multi-layered data in extraordinary volume, in near real-time (citation withheld). At a ground speed of 110 knots, the system could collect data from an unprecedented 1,500 hectares per hour, capturing around fifteen million data points per hour (SPECIM Spectral Imaging Ltd, 2016). However, AisaFENIX data alone are meaningless, so extensive ground-reference data are collected to help decipher meaning from the data by calculating relationships between the AisaFENIX data and ground measurements[30].

---

[30] Ground reference data were typically collected within 24 hours of the AisaFENIX flight.

Volume

The AisaFENIX generates enormous volumes of raw data, with each P2P sampling event consisting of terabytes of data. This volume of data got increasingly large with the data processing as each image may have many iterations before the researchers completed the processing task. As a rule of thumb, a processed image of a 3,000-hectare farm was around 60GB in size. The volume of data coupled with the speed at which they were generated (near real-time) created data storage and transmission issues for the researchers. Eventually, the group installed their own server to store and access data. Having their own server also siloed the data from the university's data storage system, allowing the group to place wider restrictions on data access than normal in the university context. *Pioneering to Precision* Researcher "Lance P" described how "Jenny R" (P2P post-graduate student), organised and managed the data on the server in a Year 1 interview:

> *Does this data here go back to Jenny R?*
>
> Lance P: "Yes."
>
> *So, she gets ARL ... does "Joe A" get ARL data? (Lance P shakes head). No?*
>
> Lance P: "One of them, they get the ARL data because everyone is assigned some jobs, what they do. There is one folder that all the chemistry results are going into that folder."
>
> *Is that in a shared drive, where's that folder?*
>
> Lance P: "I think Jenny manages all the ARL chemistry data."
>
> *So, this goes into a shared folder. Every measurement, every time?*
>
> Lance P: "Yeah. every time. She also organises all the sheets in their proper order, proper order means time date."
>
> *Jenny R does?*
>
> Lance P: Yes
>
> *So, Jenny R is in charge of maintaining, organising, and maintaining the shared folder?*
>
> Lance P: Exactly.
>
> *She's the kind custodian. Okay that's interesting. Is there anything else in the shared folder?*
>
> Lance P: ASD data
>
> *So ASD, we haven't got to this....*
>
> Lance P: All the data in the common shared data, all the PGP data, the GPS because everyone do something. For example, ASD data "Joe A", he organises the ASD data in

> the folder and then GPS data “Frank L” (Field Technician) he analyses the GPS in the folders.
>
> *So, Jenny R is the custodian of this stuff, that does feed back to her? The chemical analysis, the chemistry and then she sequences them, orders that, but there's a whole bunch of custodians of various bits of the data. it would be interesting to know what the traffic is around the shared folder, was that as another kind of shared folder?*
>
> *Because “John S”* (Researcher) *had talked quite a bit about how he's protecting his data with regards to the shared folders. So he keeps pure data in another folder and multiples locations that's the original data then he transforms data himself on his laptop but there's also kind of a... the files in the shared folders are… available to be picked and prodded by the others but that original data is also protected elsewhere?*
>
> Lance P: I don't know about the PGP data management, but I think they are pretty secure. I think. Pretty secure. (YR1)

While the P2P researchers held a copy of the project data on their server, the co-funder’s representative “Steve N” explained that Ravensdown also held the data on their own servers:

> *Where are the data held?*
>
> Steve N: At the moment, it is held with the researcher (P2P researchers), but it backed up on a Ravensdown site every three months. (YR4)

Scale and speed of data collection

An important characteristic of the airborne AisaFENIX data is the enormous scale of the data collection. This presented researchers with a range of unique opportunities and challenges. The instrument generates huge volumes of data in the form of DNs in near real-time (high velocity). Observations of the P2P field work revealed that at best, researchers could hope to collect about one reading per minute on a wide scale data collection event (60 data points per hour). While this may seem like a small sample, farmers would typically only take up to three soil transects on a 3,000-hectare farm to inform fertiliser decisions. In contrast, the scale and speed of data collection for the P2P research promised to be on a completely different scale; in practice, this meant that on a clear, calm day two farms of around 3,000 hectares in a single region could be flown on one day.

Scope of data collection

The fine-grained data nature of the AisaFENIX data is radical advancement on their predecessors, as are the emerging techniques for analysing these data at the fine resolution. The P2P project also demonstrated that hyperspectral sensors are highly variable in the scope of data they collect.

Unlike the laboratory-like scope of its proximal counterpart, the ASD, the scope of the AisaFENIX's data collection is vast.

The AisaFENIX strives to measure and observe the whole field of application capturing entire plant populations where *n*=all. With the ASD set-up employed by the P2P team, the Field of View (FOV) was only ten centimetres. This tiny footprint was deliberately designed to capture spectrally intensive data in highly controlled conditions. This made the ASD suitable for capturing 'pure' spectra where the target plant is larger than the pixel, which in most cases eliminated the need for spectral unmixing as explained by "Joe A" (P2P post-graduate student):

> *Why is it important that the ASD collects pure spectra?*
>
> Joe A: I have used the ASD to collect pure spectra from the research plots of turf because they will give me pure samples. I mean that they are all one cultivar for example, so you only have one species in the whole picture. Having these pure spectra makes analysis of the data, and ability to refine information from the data, less complicated and with fewer assumptions applied.
>
> We saw the same thing when we were trying to discriminate between kanuka and manuka using the FENIX. Where the whole pixel was made up of the same species it was quite easy to build the relationships. But when I tried to do the same thing with a mixed pixel, it was exponentially more difficult, I think maybe impossible… to build meaningful relationships between the ground reference data and the FENIX data. (YR3)

Veracity of data

According to (Kitchin, 2014), the veracity of data is a common issue with big data. The veracity of the AisaFENIX data certainly presented the P2P researchers with considerable challenges as they moved from slower ASD to the airborne HSI data collection. By design, ASD data are highly controlled and regularly calibrated. In contrast, the P2P researchers soon discovered that AisaFENIX's aerially acquired data have many imperfections and the 'pure spectra' from the AisaFENIX needed to undergo significant intervention and transformation before they could be analysed (see Data Processing section).

Consequently, the P2P researchers viewed the two hyperspectral tools very differently, and different methodologies were applied to data processing and analysis. The ASD was viewed as 'an extension of the laboratory'. In experimental terms, the ASD's setup allowed two important variables to be controlled - light between the sensor and target, and geolocation of measurement. The researchers successfully installed a black plastic shroud on the ASD to eliminate light contamination. Importantly, this approach removed the need to include time-consuming atmospheric corrections in the data processing path. Issues with the veracity of the AisaFENIX data created by the move from the ‘lab’ to the air created considerable data processing challenges for the researchers, as “Lance P” (P2P Researcher) mentions:

> Lance P: I haven't seen any in-the-field Fenix instruments. Most of them are using (it) in the lab. The lab is a more comfortable environment.
>
> *Let’s put the farm in the lab, that would be a lot easier?*
>
> Lance P: [In the lab] You can calibrate the instrument easily using some standard tile or something, you know? You can put it easily. It's very comfortable, but in the field, it is a completely different environment.
>
> *So, is that a challenge for you?*
>
> Lance P: It is a moving environment platform you know. How can you calibrate? That means the Fenix is running (*across a plain*?). Yeah, yeah. (YR1)

The researchers used an RTK-GPS unit to geolocate the ASD measurements, ensuring the positioning of the data in the data layer was highly accurate. However, the accuracy of the geolocation of the AisaFENIX was far less accurate than anticipated by the P2P because they thought the aircraft’s own system was more accurate than it turned out to be. This assumption was understandable, given the plane was used for mapping landscapes in other applications. However, the P2P researchers later realised that the variation in the aircraft’s GPS system could be as much as 11 metres.

In hindsight, the installation of a relatively inexpensive differential GPS would have overcome these issues. This was a key learning from the P2P study and also serves as a useful illustration of the role that modifying technological assemblies plays in scientific practice for this research. The reasons for the misalignment of data layers, and subsequent consequences were not immediately recognised by all the science team and this caused much angst amongst the P2P researchers. The inexperience of the science team in working with airborne imagery meant they did not immediately identify the absence of differential GPS in the aircraft as the main issue. Instead, the misalignment was put down to a processing issue. In Year 2 of the project, the researchers thought that the data from the sensor were accurate and the problem could be addressed through processing:

> Fred K (P2P researcher): We also had some issues with the GPS misalignments and things like that, which was probably happened to be because of some processing issues. So, I've been processing quite a lot of these guys so I'm doing quite a lot of surveys at the same time and probably the IDEAL, because it's all heavily based on IDEAL because CaliGeo-PRO (Specim, Oulu, Finland) is written in IDEAL and ENVI is IDEAL based. Everything is using IDEAL. So, I was processing I think two or three surveys at the same

> time and the IDEAL, probably the IDEAL was corrupted, and it caused a major issue here in the geo-rectification.
>
> *Whenever you have IDEAL, it is a problem?*
>
> Fred K: Yeah, yeah, yeah. For that batch of processing. But because I was processing quite a lot of surveys then it affected the majority of the surveys so it was quite hard to figure out what might have caused the issue but at this stage, I do random check-up here.... because those affected these GLT files and I do a manual check-up here.
>
> *So, how do you do that check?*
>
> Fred K: I usually open all the files, well not all, but just sometimes I just open five or six and I do a quick mosaicking and I check how the alignment is looking and if they are bad then I try to ... process them again or try to figure out what might cause that.
>
> Since that, I'm checking these things and didn't appear again. There was one issue with the boresite calibration. The boresite calibration wasn't good enough and that also caused similar effects, so the strips were like... [draws on board].
>
> *It didn't match up?*
>
> Fred K: Yeah, it jumps a few metres or sometimes are few tens of metres, it depends on how bad the bore site is or what the georectification is.
>
> *Farm tracks are good for helping you figure out where they are aren't they?*
>
> Fred K: Yeah, yeah. (YR2)

The true cause of the misalignment issue was only discovered with one survey flight left in the project.

Data processing

A key challenge with data-rich data technologies such as the AisaFENIX is that the value model does not emerge fully formed from data themselves. In the research setting, the AisaFENIX data generated are simply inputs to a production process that depends on human insight (Huberty, 2015; McAfee et al., 2012). It is not until the algorithms have been created and selected, that the data from the device can provide meaningful information independently.

Data emerged from the AisaFENIX as a vast mass of unstructured digital numbers. Refining value from these data required the code to be unlocked and translated into a universal language that can be analysed. It could be argued that the value of the big data generated by the sensor lies not inherently in the data, but with the processing. While the funders appeared to fixate on the quantity of data produced, this study reveals the data is irrelevant without robust analytical processing.

While much effort was put into the data collection detailed in Chapter 5, “Mark L” (P2P Researcher) stressed the importance of the data processing and analysis to the success of the project:

> *It’s automated only in the sense of, in terms of data collection. You’ve still got to process it?*
>
> Mark L: Yeah. What I think we can do here as I think we can process it. So, we can gather it and process it and get it into a form where people can readily use it. So, that’s what we’ve concentrated on with the PGP and that’s what we will concentrate on in terms of, like we will give you a map of what we think is meaningful information that you can then pick up on. (YR1)

Figure 28 traces the journey of the AisaFENIX data from the aeroplane to the reflectance cube presented to the analyst. This work was almost exclusively performed by Fred K and was extremely time consuming. The diagram shows that data output from the AisaFENIX was in the form of Digital Numbers (DNs). Fred K converted the DNs into radiance (W m^{-2} sr^{-1}) using factory-provided radiometric calibration coefficients in CaliGeo-PRO software. Next, Fred K obtained surface reflectance values from the radiance data using ATCOR4 (ReSe Ltd., Wil, Switzerland), using protocols based on the work of Richter and Schläpfer (2002), which use geographic, temporal, and atmospheric parameters.

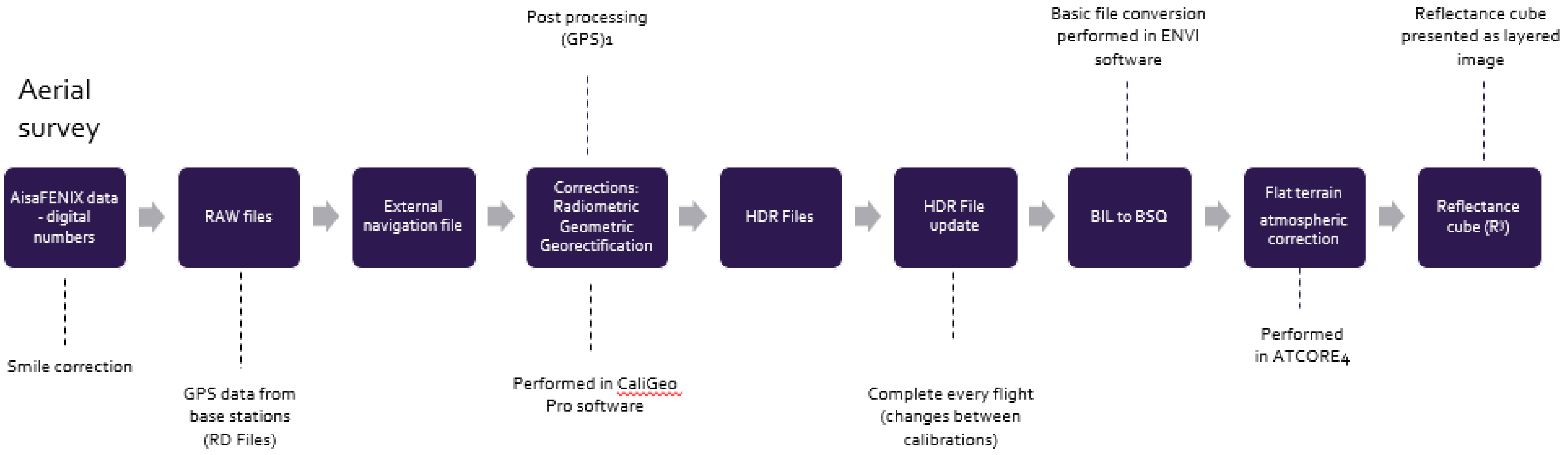


Figure 28 *Data journey through the pre-processing stage, as explained by Fred K during an interview in Year 2. Note. Post processing uses a navigational file to increase the accuracy of (data) strip location*

There were two key observations from this process. First, Fred K, performed many iterations of the process during his research. His goal was to find what he described as a ‘stable approach’ that could be applied to all the datasets emerging from the AisaFENIX, which could then be bound within a black box that the researchers called the HyperBox. Second, at the end of Fred K’s processing, the data emerging (a reflectance cube) was made available to the wider P2P research team so they could begin their own pre-processing and analyses. The P2P researchers performing the iterations described later in this chapter all emerged from the same reflectance cubes produced by Fred K. Interestingly, most of the project’s outputs emerged from the data he processed from the final flight, as these were the most accurately geo-located dataset.

The P2P researchers found the mercurial nature of aerially acquired hyperspectral imaging data made it difficult to refine value from each data collection as new atmospheric corrections are required for every flight. According to P2P Researcher “Lance P”, without these corrections, it was impossible to standardise results to ensure that data from different flights were relational and reliable algorithms could be crafted to create meaning from the data:

> Lance P: In order to give some meaning, full information to the data, we do some corrections.
>
> *How do you correct; correct in terms of what?*
>
> Lance P: First, there is no standard techniques because the problem is very few instruments are available around the world. So far only just the commercial instruments started very, very recently, only just, four or five years.
>
> *But the corrections are not to produce meaning right, they are just to make the measurements?*
>
> Lance P: Some of the corrections produce the meaning?
>
> *I'm thinking that by meaning by the way that something like nitrogen is going to turn up, but that's what you mean right, a whole set of them will apply to Nitrogen? The corrections that Fred K is doing to the Fenix data, are those corrections for the atmospheric conditions?*
>
> Lance P: Basically, the atmospheric conditions. What is happening is that the light emitting from the ground is reaching to the sensor objects. There is a lot of atmospheric factors interfering with the light (sic).
>
> *Such as the angle of the sun?*

Lance P: Angle of the sun and the aerosol particles, industrial particles, everything interfering with each other.

And he's taking some ground readings when the Fenix actually is flying to help that correction, is that right?

Lance P: Yeah, yeah. Once he has corrected those images from those atmospheric interferences then after that he did the georeference, that means the sites are in the correct positions. Then after he convert the raw data numbers to the reflectance. The reflectance is a standard value, that means zero to one.

Is there a zero to one for each band? I mean how many zero to one...?

Lance P: Exactly. Zero to one for each band.

So, each number is coming out on a scale and you scale them all to a single zero to one scale.... Is there a cost to doing that? Is that a risky part of terms of the enterprise?

Lance P: Yeah

You might end up with cabbages instead of nitrogen?

Lance P: There is some loss of information, but the thing is there is no standard approach. Still they are only developing, you know?

Why do you need the standard numbers? Why do you need to standardise?

Lance P: Because normally the number, the actual number should be 0-1 but it produces a number from -20 to 2000, something irrelevant you know. So, what we do, we convert those values into 0 to 1 in a standard way.

So, you log them or something… why do you need to standardise?

Lance P: Standard numbers. If I get an image today, tomorrow I will collect another image. If you don't standardise those values, you can't compare today's image to the next day.

So, temporal variation? Wow that's an amazing piece of tech.

Lance P: So, today is collecting a different numbers (sic), tomorrow is collecting a different numbers (sic).”

Is that a normal just in terms of instrumentation in general, is that, like if I look down a microscope right, I'm seeing the same scale all the time, if I look down tomorrow, I will see the same thing – right?

> Lance P: With the microscope you are dealing with, everything is controlled. The light is a standard, the microscope is a standard. Everything is standard. So, you get standard values and a standard scale, each and every day.
>
> *So, you've got natural light and the variation (and the wind...), dust.*
>
> Lance P: Exactly. And instrument performance and the atmospheric effects, light intensity. And photons, that means that the light reaching to the sensor.... Light is nothing but photons. Photons reaching to the sensor, so everything is different.
>
> *So, you adjust that so it is homogenous through time so your data can be compared? Are there any other instrumentations that are like that? Is that a common thing, or is that specific to light measurement, this lack of standardisation? It is such a moving target all the time. You think of an instrument that the... the reason why you have an instrument is because the instrument has done the standardisation for you?*
>
> Lance P: Yeah
>
> *It's already sort of pre-loaded into the instrument essentially and that is why the instrument is so handy because you don't get -200 one day so it's always standardised. It just strikes me as an odd challenge around instrumentation. Do you know other fields where other types of instrumentation have a similar kind of challenge around standardisation?*
>
> Lance P: No, no, it is only a problem when you are working in the field. The problem is that in these conditions we can't calibrate the instrument (*Uncontrolled environment*?). Because all the laboratory instruments, all the NAR machines, MRI, everything you can calibrate, you can put your white tile, you can standardise against a standard reference.
>
> *So, its field versus laboratory conditions?*
>
> Lance P: You can't use any standard reflectance there. (YR1)

As the purpose of the P2P project was to model these environments, the presence of shifting, mutable data did not make processing and analysis easier.

The P2P researchers and funders placed great emphasis on the creation of models from the AisaFENIX data. However, the research team discovered the data must undergo significant pre-processing to even reach the stage where they can be processed for analysis. It required pacifying before it could be made to work. The complexity of the pre-processing was far beyond what was expected and required significant time and resource to resolve. The project plan expected the large part of the scientific effort to focus on the analysis of the data rather than preparing the data for analysis, so timelines were unnecessarily tight at this phase of the project.

A consequence of this was that key P2P researchers spent more time on the processing instead of performing other work within the group, such as doctoral supervision and working on other projects.

The project stakeholders' focus on protecting intellectual property prohibited the researchers from collaborating with researchers outside the project. All the P2P students had to sign non-disclosure agreements and their theses would be embargoed upon completion. Science-making was tightly controlled to the immediate P2P researchers, who had signed up to the confidentiality agreement. One researcher who was involved with the wider research group, an experienced remote sensing expert, was not willing to accept the restrictions on their academic freedom and refused to participate in the project. Thus, his expertise was not available to the students in the group. The hard line taken by the university to protect its commercial interests and those of the project funders engendered an environment of isolation. The P2P students were under the impression they could talk to no-one outside the project and all publications were vetted by the funder before publication.

While in other situations this may not significantly constrain science-making, in this case some of the research team felt it unnecessarily hindered academic and commercial progress. This was primarily due to the science team's inexperience in the field. While some researchers had experience in proximal hyperspectral sensors, none had significant experience translating airborne data. Furthermore, the publicly available information in the field was very limited. As intimated in a Year 1 interview with "Lance P" (P2P Researcher), the researchers felt they were working beyond their experience level, in a field where there was little publicly available information to guide methods and provide reassurance their path was the correct one:

> Lance P: The work is very interesting, but the problem is you don't know how to process the data. Because it is still in the beginning. There is no standard softwares (sic). There is no softwares to download the data. Only recently started the availability of the data.
>
> *So, the data has only been around for four or five years?*
>
> Lance P: Before that, every instrument was owned by the government. It started in the late 1990s. That is the only instrument available around the world. So, it has been used since the 1990s, still running. So that is the only instrument. So, they done (sic) everything with that instrument. But now all the commercial companies are developing similar instrument (sic) to the Everiss. Everiss is a similar instrument to the Fenix but it is massive. It is a very good electronics (sic) that means very good signal to noise ratio.
>
> *Can you put it in the air, can you put it a B-52 or something?*
>
> Lance P: Yeah. It's a massive instrument. About 150 kilos.

> *So… is this... are you saying that there's no, ... everybody's trying to figure out processing? How to process the data and a lot of it in commercial contexts, I'm taking that means you don't go to a conference every year and everybody sits down and describes all the ways they've developed and processed data... You must have a sense of the people around the world that are working on similar problems to you?*
>
> Lance P: Everiss is owned by NASA. NASA, they have a group of people, they know how to analyse the data.
>
> *But that's not public information?*
>
> Lance P: [Shakes head]
>
> *No?*
>
> Lance P: 'Cause yeah. That is a story.
>
> *Are there peers that you can work for (with) or are you on your own?*
>
> Lance P: Other researchers, for example… I am doing some vegetation mapping in a particular area in the US(A). I borrowed to Everiss, I told to Everiss, the NASA people, to scan this area. They do all the corrections then after they will provide the reflectance cube, the final cube to me. (YR2)

While the intrepid researchers found little information to guide their technology-related challenges, they also encountered difficulties caused by their choice of proxies - light and plants. Both proxies can be highly variable. Light can be controlled and kept constant in the laboratory, however the move to the field means many corrections must be made to try and account for field variability. Cloud, which has proven so problematic to New Zealand remote sensing researchers working with satellite data, also provided a challenge for the researchers working with the aerially acquired AisaFENIX data. The P2P researchers were concerned that even flying below cloud layers could influence the variability of light on sensor readings. To mitigate the impact of this source of variability, the sensor was only flown on bright, cloudless days to reduce the influence of light levels on the spectral reflectance data. Even so, the light levels influenced radiometric corrections applied to the data according to a Year 2 interview with “Fred K” (P2P Researcher):

> Fred K: I have three options with three binnings, but we usually use the four times two binning. That's basically the number of spectral bands which we combine into one. That's for noise reduction. If for example, today, we should bin them quite a lot to enhance the strength of the reflectance of the spectra what (sic) we're taking.
>
> *Is that because it is a dull day so if the light was higher, you'd do less binning?*

> Fred K: Yes. And then up around eight ... it obviously affects the number of bands what we can get so if you used eight times two then we can get like 360 bands... I think 63 or 65 or something around that. But, because we use the four times two, we get the 448 bands most of the time, most of the survey.
>
> *Are you taking the light levels to help you assess the light levels to figure out how much binning?*
>
> Fred K: Yeah, yeah, yeah, we use the actual light settings. But because we aim to fly at... only on nice days then we should stick to four times two. (YR2)

While only flying the AisaFENIX on cloudless days helped the researchers to calibrate their instrument, in terms of commercial applications the project was barely an enhancement on the doomed Pastures from Space project mentioned in Chapter 3. The high number of cloudy days in New Zealand meant that without evidence that the researchers could successfully perform data collection on cloudy days and analyse the data, the technology would not be commercially viable.

The P2P science team aimed to perform commercial flights during spring and autumn seasons when moisture deficiency is less likely to be a limiting factor on pasture growth. Further restricting the limited window of opportunity by only flying on cloudless days would impact on the financial proposition of the project. During our interviews, Fred K suggested the researchers should focus on this problem, and also should address poor-quality slope data by collecting LiDAR data in conjunction with the AisaFENIX flights. However, neither suggestion appears to have been implemented during the data collection phase.

Geometric alignment of images

In addition to atmospheric corrections, the researchers also found the data generated by the airborne hyperspectral sensors require geometric corrections to render them useful. This was because the aerially acquired AisaFENIX images did not accurately align with the ground-reference data that had been geolocated to three-centimetre accuracy using the RTK-GPS. One such issue was related to the poor geometric alignment of some images, which was exacerbated by the absence of differential GPS in the aircraft. According to "Joe A" (P2P post-graduate student), this made it difficult to align the aerially acquired data with the ground-reference data[31]:

---

[31] Differential GPS enhances the quality of location data gathered using GPS receivers by providing positional corrections to GPS signals. DGPS uses a fixed, known position to adjust real time GPS signals to eliminate pseudo-range errors. "GPS signals coming from satellites down to the ground have to travel through layers of the earth's atmosphere, so they are subjected to delays. This affects the time taken for the signal to travel from any given satellite to a GPS receiver, which introduces slight error into the GPS engine, causing an error in the measured position" (Racelogic, 2018, p. 1).

> Joe A: In this case the images were calibrated using ground data (wet chemistry analysis) from geolocated sites on the properties analysed. While the ground data was accurately geolocated, differential GPS was not installed in the plane so readings could be up to eleven metres out, therefore any ground data collected could only be located within that range or manually calibrated.
>
> Larger features that are visible in the aerial image can be still be used, such as trees and buildings. The spatial error would not be relevant if the target had a larger homogenous area than that displacement (for whatever you are measuring), such as a field of maize or soybeans etc. for species identification, but because hill country pastures are highly heterogeneous in most respects the co-registration of the GPS ground targets and the image is critical to the accuracy of image calibration is critical. (YR3)

To add further complexity to the task at hand for the P2P researchers, the geometric corrections needed were not always consistent, making the automation of the corrections complex. In an effort to overcome these challenges, researchers worked the data into strips of images that were stitched together before they could be analysed. The researchers found the strips were not always well aligned, such as in Figure 29, and had to resort to manual stitching, see Figure 30. This time-consuming process was not easily automated and had flow-on effects for alignment with ground-reference data.

Figure 29 *Image with poorly aligned image strips*

Figure 30 *Corrected image after manual stitching*

This illustrates the importance of assembling the correct machines when working with these nascent remote sensing technologies. While the need for the differential GPS may not have been readily determined from the outset of the project, it was a useful finding from the P2P study for future research. Unfortunately for the P2P project however, the unwitting omission of a single dependent technology, differential GPS in the aircraft, had scuppered the researchers' ambitions of aligning the aerially acquired and ground-reference data for all but one flight. To solve this problem researcher Fred K suggested that ground-reference tarpaulins be placed next to the georeferenced site pegs for the final flight, see Figure 31.

Figure 31 *Researcher positioning a reference tarpaulin employed for georectification purposes*

This approach allowed the researchers to manually align the ground data collection sites to the correct location within the airborne images later. The process was described by "Joe A" (P2P Post-graduate student):

> Joe A: Each tarpaulin is placed five metres South of a survey site, with both the site and tarpaulin geo-located using a real-time kinematic (RTK) DGPS system. This confirms the actual site locations to allow any positional error in the mosaicked hyperspectral image to be calculated so a correction can be applied. (YR4)

The installation of the ground-reference tarpaulins was a simple, yet genius solution. For some P2P students, this was the only flight that produced valid data for their studies as their analyses relied on accurate alignment of the aerially acquired and ground-reference datasets. The multimillion-dollar project had been saved by $5 tarpaulins from a local hardware store, and some timely ingenuity.

The data processing was not optimised until the ground-reference tarpaulin approach was employed in the final flight. This is a useful example of the value of how a small data set can have significant implications for the success of a big data project.

Managing missingness

Another issue faced by the P2P researchers was addressing the missing values identified by Fred K in the data generated by the AisaFENIX sensor. Missing values are not uncommon in remote sensing data, especially aerially acquired data. Missing values may be due to a variety of reasons including a malfunction in the sensor and also occur when the sensor data show clouds instead of the Earth's surface. Clouds show as a white pixel in the image (or 1 numerically, which represents 100% reflectance). In the P2P project, the aircraft was flown on cloudless days to avoid missing values due to clouds. However, they still encountered missing values in their data, which displayed as a zero in the dataset, or as a black pixel in the imagery.

One of the first things the researchers needed to consider with missing values is whether they were missing systematically or missing at random. The second important aspect is the rate of missingness. Clearly, when there is an abundance of data such as with hyperspectral imagery, the number of missing values is viewed differently. Fokoué (2015) suggests three main approaches to address missingness when using big data: (a) Deletion, which consists of deleting all the rows that contain any missingness; (b) Central imputation, which entails filling the missing cells of the data matrix with central tendencies like mode, median or mean; (c) model-based imputation using various adaptation of the ubiquitous Expectation-Maximisation (EM) algorithm.

In the case of the P2P project, the researchers used the deletion approach to address missingness in the sensor data, a risky approach given the researchers were not entirely sure whether the values were missing at random or not. However, the researchers were relatively confident that the issue was with the sensor (so assumed to be consistent) and to take another approach would have been very time-consuming and possibly led to further delays on the project. The implications for the project are that when mutable values are updated, their state applies across all references to that variable. So, changing a value in one place, changes it for all references to that object (assuming a systematic missingness). Unlike immutable data types, there is no way of changing the internal state of the mutable data, so the reference always gets reassigned to a new object. The biggest implication of this is that for immutable data, equality is more reliable since the value's state won't change. Another interesting observation was that the researchers applied the second approach, central imputation, to address possible outliers during the analysis phase. This is considered risky when dealing with inherently variable data, such as hill country farm soil nutrient data. P2P Researcher "Fred K" described some of the challenges the researchers were encountering in Year 2:

> Fred K: If there is some noise then it's like, it's more like peaks... a spike occurs and then this is usually... belongs to one band or one detector, which didn't work properly.
>
> *Interviewer: So that's in the sensor itself hasn't worked?*
>
> Fred K: Yeah. But that's all natural, I mean every sensor has these hiccups and sorts of things.
>
> *Int: Are they usually the same ones?*

> Fred K: No. They change in random. That makes them hard to detect and remove. We simply detect that based on the deviation between these lines and that. Because they are usually around 65,000 value, digital number, which is the maximum. So, it's alright to detect them. And we replace this value with the neighbouring value. (YR2)

The HyperBox: a black-box solution for pre-processing.

The researchers applied different approaches to atmospheric corrections to the raw AisaFENIX data in an attempt to improve the accuracy of the processed data. A key observation of this study is that the pre-processing of data before the analysis stage was far more time-consuming and complex than had been planned for. Importantly, data from the different approaches were incommensurate due to the different corrections applied to the raw data.

The P2P researchers were initially spending weeks making atmospheric and orthorectification corrections in preparation for analysis. By year two of the project, the researchers had the pre-processing time down to approximately two days. This processing time and its associated costs were likely to scupper any opportunities to commercialise the surveys, so the researchers had to find a solution. Once the researchers decided on a set approach that was ‘good enough’, which would reduce the processing time, and ensure that the data between subsequent flights were commensurable. The P2P researchers engaged the services of a young computer programmer to help them automate as many of the pre-processing steps as possible. What they developed was essentially a pre-processing black box called the HyperBox, which is short for hyperspectral processing black box. In practice, the HyperBox was a set of automated processes that ran correction processes with minimal human intervention. Importantly, the researchers believed that the steps involved in their HyperBox method were valuable IP, so information about their method would not be made public.[32] The development of the HyperBox was a pivotal step towards creating a commercially viable spectral survey service. It also reduced the labour required to pre-process the research data as described by “Fred K” (P2P Researcher):

> Fred K: But the biggest deal is the software what we're developing so I don't need to click all of these strips through but I just usually get all the files I need for the processing, like, for example here this external navigation file, I also have the boresite calibration, the sensor definition and many other supporting files what I need for the processing and I also gather all the data, all the raw files into one folder and then the software what we develop can do these automatically from here to the mosaicking.
>
> *Wow.*

---

[32] Whilst the pre-processing steps were recorded in detail at the time, commercial constraints mean these are not published as part of this thesis.

Fred K: That made a big difference because then originally it would take me to process a strip about a day. But now I can do it in a few hours so that's a big improvement.

And when you say a strip, that's just one strip width or..?

Fred K: Oh, did I say strip or survey. Oh, one survey. In the old workflow I did probably one, or I worked one or two day per survey. Yeah. But now it can be done in two hours let’s say, three hours. There are many things to do for example we have troubles with the mosaicking part but that's the easiest to fix. And also, the CaliGeo-PRO is not completely operated from the software because what we usually do, we have many softwares CaliGeo-PRO, NVI, ATCORE4 and ENVI again and we do like an overarching software here. (writes on board). We call it HyperBox at this stage. And we practically control these softwares and there's also a little algorithm here. So, we control everything from there.

And we specify everything, all the settings necessary for the processing at the beginning so it has a single user interface and we hit RUN or you can also execute these steps separately. But at the end it will be better to execute everything in one go.

So how far along are you with the HyperBox? Is that, that’s not fully developed yet, you're still working on that?

Fred K: Yeah, we're still working on bits. The mosaicking is still not working, and the CaliGeo-PRO is not integrated into that one, so we create everything for the CaliGeo-PRO, but we stopped using HyperBox at the moment and we open CaliGeo-PRO and process the data there. It has a batch processing function so it's easy because you just open the CaliGeo-PRO, add the already defined option files and then hit run and it will do everything.

But HyperBox is doing until that point at the moment?

Fred K: Yeah, So, it does everything from here... and from there. It's not bad. The (inaudible) bits for example, is the external navigation file, I create that per survey, and I create these. Those cannot be really automated or there is no point because you do it one... And... we are working with this, so we want to integrate CaliGeo-PRO into that workflow, into the HyperBox later. So, I already asked Specim to do some updates for us.

> *So, is the CaliGeo-PRO one of their software products?*
>
> Fred K: Yes. It's a free, free of charge, it's a freeware software and they are quite enthusiastic about developing that further so hopefully they can give us some better user interface. Because it’s very hard to automate these through the user interface so we have to get rid of the user interface and control everything from the common prompt and with the orders and then we can incorporate that fully into the HyperBox.
>
> *Well it looks like you're making lots of progress. Are you feeling like you're making progress?*
>
> Fred K: Yeah, yeah, yeah. I'm quite happy with the HyperBox and the concept and everything so that's a good improvement. (YR2)

The problem of incommensurability

The accurate geolocation of the ASD and the controlled light conditions mean little intervention is needed to render ASD data from different measurement events commensurable. This makes the tool ideal for longitudinal studies where robust models are desired for year-round application, such as the P2P project. It also means that the data are more easily aggregated – an important feature where the data are expected to have life beyond the initial project. The P2P researchers adopted a quick and easy calibration process using a white reference tile that was carried and used during measurement events. This method ensured the ASD was calibrated between measurements to minimise variability, mimicking laboratory conditions, so the data from different ASD measurement events are relational.

Calibrating the AisaFENIX proved much more troublesome for the science team. The instrument was sent to the manufacturers in Finland for calibration, including during the data collection period (between data collection flights). The calibration was required to ensure the AisaFENIX was accurately sensing reflection. This meant the tool was not available for weeks at a time. Even more importantly, the researchers found that data collected before and after calibration were not easily commensurable. The researchers encountered other problems relating to calibration of the AisaFENIX. Unexplained lines appeared in the data for some of the trial flights that were interrupting the creation of algorithms. At the time, the P2P researchers thought they had found a way to ‘correct’ the data; however, they could only do so for the data generated since the last calibration. This meant that lines in data from measurement events before the most recent calibration could not be corrected, as P2P Researcher “Fred K” explains:

> *Interviewer: When you found a way of fixing that did you go back to those other images and run the new process over them, all the raw data?*

> Fred K: Ah, no I started with the last survey. So, everything after, I think September 2015, they have been corrected but I didn't correct for the older ones. And also, those are with different calibration and set-up and we cannot do backwards, so it has to be done, these things at specific calibrations.
>
> *Because you don't necessarily have the calibration set-up?*
>
> Fred K: Yeah, yeah, yeah. So that was one issue. And the other day we had not vertical lines, but horizontal lines.
>
> *How did that happen?*
>
> Fred K: That happened after the survey. So I got nothing to do with that one because how I got the data it was already in the raw data so I opened the data and I saw horizontal lines everywhere and we figured it out, I asked the Aerial Surveys guys to upload the files again to our server and they upload it again and the lines disappeared so that was something that has to do with the copying or something must have happened with the files during copying. Yeah but we sort it out, it was a one-off thing.
>
> *Visually I'm just trying to get my head around what you're actually seeing with visual lines and horizontal lines?*
>
> Fred K: Imagine the Fenix image and like every 15th or 16th row is like, is in coloured in one colour, coloured red, coloured blue, or something. Some random colour. And, that's it. So, those lines ruin the whole image and also, they ruin the spectra as well, so it wasn't only one band, but throughout the visible and throughout everywhere.
>
> *So, did you have to rectify that, or did they make the adjustment?*
>
> Fred K: No. no. no. We deleted those files and they re-copied that, then re-downloaded from the Fenix. (YR2)

This problem was compounded by calibration and the researchers' experimentation with pre-processing techniques being applied to the datasets. During the interviews, one researcher revealed that during the early flights they experimented with various methods of processing for different flights, meaning the lineage of each data collection was different. Subsequently, researchers from other organisations working with the AisaFENIX sensor have confirmed that the problem with stripes in the data was indeed a calibration issue. However, at the time, the reality for the P2P researchers was that each time the AisaFENIX was calibrated a whole new batch of calibrations were needed and the data between calibrations were not commensurable because the stripes could not be accounted for. This meant that data from different flights (and their output models) could not be compared or aligned with each other, limiting the opportunities for data and model validation.

Conclusion

In this chapter I extended my ethnographic investigation to identify and explore the ontological data characteristics of airborne hyperspectral imager AisaFENIX and the complex community of machines that it relies upon in the research context. I returned to the laboratory to chronicle the efforts of the P2P researchers to prepare the AisaFENIX data and their supporting data for analysis. I found that the AisaFENIX provided the P2P researchers with enormous volumes of data that are of high spatial and spectral resolution. These data arrived early in the science process at an unprecedented velocity and scale. However, I found that these data are not always accurate and require much processing to render them suitable for analysis.

The complex data assemblages generated by the P2P project contained data that fulfil the criteria of big data, however deeper analysis revealed that in the research context, small data are still vitally important data sources where they are accurately geo-located. Until recently, most PA researchers came from backgrounds in agronomy, soil science, agricultural engineering, and geo-statistics. Experts in their own field, it is unlikely many of these academics would also have the specialist expertise required to process or analyse remote sensing data. Thus, this may signal a possible shift in the scientific division of labour within research teams to include more personnel with remote sensing and data analytics skills. Observations of the researchers at work revealed that data from these technological assemblages required significantly more processing to prepare them for analysis than anticipated, and as a result a new scientific division of labour in PA is also emerging.

Chapter 7: Beyond scarcity – the investigation and discovery process in the all data ecosystem

Introduction

Big data have the potential to be truly transformational for agriculture (Alrøe & Noe, 2014). However, applying more data to existing business models is unlikely to drastically improve the livelihood of farmers, nor address the wicked problems facing the agricultural sector, such as freshwater contamination or climate change. In the case of *Pioneering to Precision*, the funders and researchers were attempting to ‘revolutionise hill country farming’ (Massey University, 2016). To obtain a more comprehensive view of the soil conditions on hill country farms, the P2P researchers had to look beyond the information that could be immediately extracted from the AisaFENIX imagery and use proxies, or indirect measurements based on this spectral information.

While solving the P2P’s original research question alone would have been an impressive feat, arguably the true potential of the data generated in the project lies beyond the original project and depends on the ability of others to repurpose some or all of the data for other uses. Indeed, the P2P project provided the researchers with an opportunity to develop new products and services that rely on the data themselves rather than just being an input into existing business models (Fraser, 2018; Huberty, 2015; Wolfert et al., 2017). Once the geo-located dataset was in place, there were opportunities to create other proxies based on the information. Experience from other industries suggests however, that refining value from big data technologies would be a challenge (Huberty, 2015; Sonka & Cheng, 2015a; Sonka & IFAMR, 2014; Sunding & Zilberman, 2001).

Transforming information into knowledge requires researchers to first understand and assess the information. Second, the knowledge that is generated must be distributed to potential users, such as farmers, in a way that they can implement in their own situation. This requires farmers, for instance, to be motivated to use the technology and the knowledge must be in a form that is understood. In the P2P project, a key challenge was to bridge the divide between technical solutions and implementation in the field. To achieve commercial success would require the P2P researchers to translate data intensive scientific practices from a laboratory-like situation and its specialist journals and their audiences to knowledge that could be understood and accessed on the farm.

In Chapter 6 I described the characteristics of P2P data and how the processing required to render the data suitable for analysis. In this chapter I examine the data investigation process, verification and scientific reporting performed by the P2P research team to refine value from the P2P dataset.

Framed by new data relationships, a new scientific method in PA is emerging, and I find it is a far cry from the linear classical science traditionally promoted for agricultural science. As how science is done changes, so to do other relationships. In this vein the chapter also investigates the scientific division of labour within the P2P research team and examines how this is adapted to meet the needs of a new data ecosystem. I discuss the intractability of linearity in scientific reporting and find a disjunction between how science was done in the field and in the laboratory, and what is reported in academic literature. A gap that has implications for how the PA field may progress knowledge in future.

Value discovery in the *Pioneering to Precision* project

In the early months of my study, it appeared that the research effort in the P2P project would mainly focus on the data generated by the hyperspectral sensor, the AisaFENIX. This proved to be an oversimplification of the reality of the project. It became apparent that the study was really a large-scale calibration exercise, which was highly dependent on data processing from a complex community of ground reference data. It was these ground-reference data that would give meaning to the data produced by the AisaFENIX.

From the outset, the P2P researchers knew the task of refining value from the variable hill country pastures would be far more difficult than the simplicity of monocultures where per pixel classification methods could be applied, as P2P Researcher “Lance P” explains in an interview excerpt from early in the project:

> Lance P: In case of pasture, we need to keep one thing in our mind. It’s a lot of, a lot of work done before with the remote sensing (sic). Before, there was only a multi-spectral sensor. That means only discrete portions of the spectra... used to take the signatures of the pasture. But they failed because pasture is highly heterogeneous. When there is a multiple species (sic), you need a high spatial and spectral resolution. That is very essential. That is why not much remote sensing work is done in the pasture area.
>
> *Because it's hard?*
>
> Lance P: Because when you come to the homogeneous crops like rice or wheat, very easy. (YR1)

A key area of interest in the ethnographic analysis was to unearth the organisation and practice of scientific enquiry in the P2P study. While remote sensing is common in PA, the P2P project was unique in New Zealand because of the scale and characteristics of the data collected and analysed. The tools used in the project promised to replace expensive, laborious and time-consuming data-collection methods traditionally used in agriculture with the non-invasive, rapid collection of high volume, versatile data (Pullanagari et al., 2011).

Notwithstanding the project's ambitious intent, Chapter 5 revealed that the project was structured in a typically linear fashion, with tight timelines and ambitious milestones. Indeed, the project was framed around the funders' notion that innovation is stable and predictable rather than a model deliberately crafted to refine value from a scientific process. In a Year 4 interview, the co-funder's representative, "Steve N", described the process from his viewpoint:

> Steve N: The process has been fairly linear in that we had to do some steps that have accomplished what we have wanted and our milestones kind of fitted that linear process but still gave us the flexibility to do the discovery. What I mean by that, is that we knew had at certain points we had them... we had to collect a whole lot of data and whatever that may come to that data is and the next point is that you have to build a relationship and whatever that relationship is and whatever its sensing, whatever its picking up we … and then you have to test it.. (YR4)

Within the first six months of the P2P project, the researchers had accumulated the largest dataset of its type in the country. Yet, refining meaningful information from the dataset was proving very challenging. Unlike established PA research projects, data from a wide variety of sources were produced up-front in the process. Observations of the P2P researchers in action reveals they responded to the deluge of highly variable and versatile data generated by the project by adapting their scientific method.

While the researchers initially tried to follow the classical, linear structure of investigation laid out in the project plan, they soon departed from this method. As I observed them and talked to them about their work it became apparent that unbeknown to most of the researchers, they were not following the script of a classical linear approach to knowledge development that had been laid out by the funders. Perhaps due to the disciplinary origins of the researchers, the P2P research team adopted an approach akin to that used in the field of data science to translate superabundant data into valuable knowledge (Pfister & Blitzstein, 2013), see Figure 32.

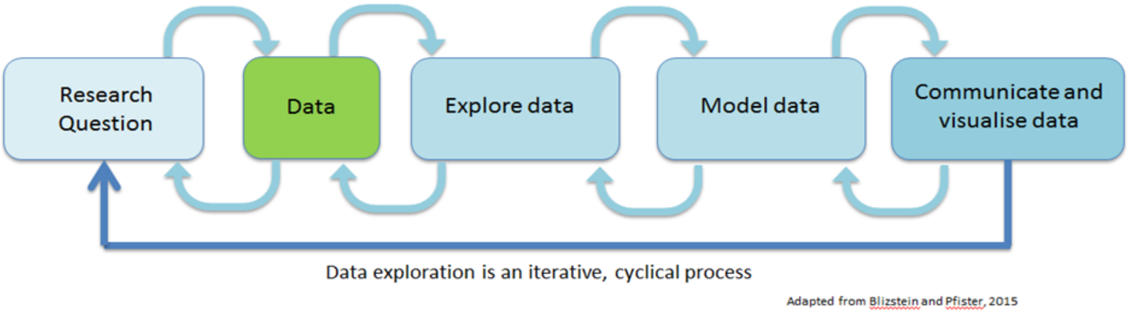


Figure 32 *The simplified science process promoted by data scientists working with big data*

The data science approach is helpful where researchers are looking to identify a correlation, i.e. a relationship and association to another variable, such as identifying a relationship between spectral readings and direct measurements.

According to (N. Rao, 2017) the main difference between the data science approach and established data analysis is that:

> The latter is hypothesis driven, while big data analytics rely on machine learning to arrive at best fit models. The central feature of machine learning is its essentially 'theory free' or hypothesis free' approach with focus on learning from data (p. 4).

However, discovering a correlation does not imply causation, where that one variable causes another to change, which means one variable is dependent on the other. Therefore, this approach can result in challenges of interpretation, spurious correlations, and model fits (N. Rao, 2017). Importantly, where only a correlation needs to be identified, the domain expertise of the data scientist is more powerful than other domain expertise. As the P2P researchers adapted their practices and processes to accommodate the characteristics of the data, a new science process in PA emerged.

Life in Proxyland, where to measure everything is to know everything

Creating algorithms that refine data into valuable information requires an understanding of what variability means. In a push to keep to their highly constrained development schedule, the P2P researchers decided to rely on correlation coefficients with ground data gathered from their initial data collection events to validate their models. Some researchers abandoned their early efforts to engage with farmers and domain experts, instead they turned to statistics alone. To make their emerging method work, the researchers pushed the validation work to very late in the science process. When describing validation, the P2P researchers were generally referring to a set of validation flights, which were flown after the test flights as part of the P2P project. Like all remote sensing products, the P2P algorithms needed to be validated, because uncertainty was produced in every data processing procedure. In the P2P study, validation involved not only quantifying uncertainty, but also providing thresholds to the funder to determine whether the product was reliable.[33]

In Year 1, the focus of the P2P project had been to gather data from which the research team could later build proxies, or indirect measurements. The researchers' goal was to create an algorithm that would represent the relationship between AisaFENIX reflectance data and soil phosphate levels. However, the arrival of the post-graduate students on the P2P project meant that it was

---

[33] Leaving validation to very late in the process increased the chance the findings or assumptions were invalid.

necessary to return to the processed data to derive new insights from the dataset. This was because all, bar one, of the students had projects that relied on repurposing the data rather than using data collected to answer the P2P project's main research problem.[34] As the project progressed, the researchers also returned to the *unprocessed* data. While returning to the unprocessed data was not initially planned, it was driven by the researchers' need to calibrate the AisaFENIX against the wet chemistry data, as well as trying to model relationships between the aerially acquired (AisaFENIX) spectral data and the ground-reference spectral data, see Figure 33. Hence, by Year 2 there were two different datasets the researchers were interested in re-purposing; the aerially acquired spectral (AisaFENIX) data and ground-reference spectral (ASD) data.

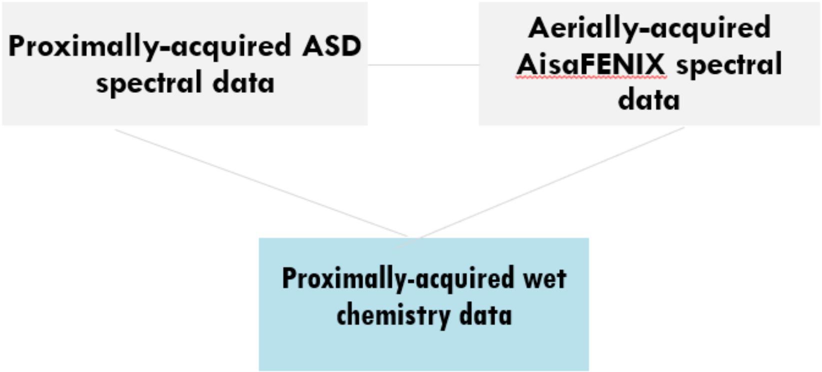


Figure 33 *The P2P project aimed to calibrate the ASD and AisaFENIX data with each other, and wet chemistry data*

By Year 2, the P2P team had amassed a large volume of data, yet the researchers were still posing new hypotheses for the post-graduate students to test. The point at which the experiments were designed and when research problems and hypotheses were posed was no longer set at the front end of the process; instead, they had become midpoint issues. The redistribution of experimental methods and hypotheses from the start of the science process to midstream raises some interesting epistemological and ontological questions. For example, researchers have previously developed and documented their experimental methods and procedures *before* data are collected to ensure that research is genuinely independent and efficient. To adapt to this new ecosystem of data they were creating, the P2P researchers adapted their processes and looked to create validation opportunities later in the science process. By Year 3 the researchers concern turned their attention to refining value from the data, as P2P researcher "Virginia W" pointed out:

---

[34] Jenny R's thesis directly linked into the PGP project outcomes, i.e. she was tasked with finding a relationship between the AisaFENIX data and the nutrient levels in soil. The other students' projects were supplemental to the project. For example, Joe A's goal to map manuka populations based on the AisaFENIX data and his own ground reference data. Another student investigated soil moisture content using ground reference data collected via a wireless sensor network and the AisaFENIX data.

> We do need to extend the clever, to the value and actually have some dollar figures where people benefit because most people want the cost benefit analysis. They don't want to buy clever, because clever is for sale everywhere so they need to have useful clever that is going to make them money.
>
> We’re saying we can describe the environment in a lot of detail but in terms of how we extract benefit from that there is I think a bit of a lag here as to understanding how we can actually use this and once we get to be able to actually use it. So, people like me saying this is a fantastic opportunity, but somehow... here we are we are sort of having a mind-fuck basically about how the hell do we make sense of all of this. I think that’s kind of the issue at the moment. (YR3)

Virginia W’s concerns were soon quelled as the opportunities for value creation from the processed data came thick and fast in Year 3 and 4 of the project. One student, Joe A, repurposed the dataset to develop a novel method of identifying manuka and kanuka plants from the dataset. Manuka honey production is a lucrative new industry, which can provide a higher return than sheep and beef production in the right circumstances. The student suggested that this was a possible avenue for helping farmers select areas to place beehives based on existing manuka populations. Furthermore, it may be possible to use the data to identify areas of low production potential that also would provide suitable growing conditions for manuka. This could inform decisions on which pasture areas could be retired to manuka and provide a viable honey production zone. Also, in Year 3, a second opportunity for value creation was borne out of Joe A’s conversation with a participant farmer from a test farm.[35] The farmer had seen some of the student’s pasture area maps and believed there was an opportunity to add value to his operation by extending the maps across his whole land area. It emerged that the farmer was previously a land valuer and he wanted Joe A to measure the effective pasture area on his entire farm, not just for pasture management purposes, but to provide a more accurate land valuation than the available existing practices. Again, Joe A repurposed processed data from the original AisaFENIX surveys to build more accurate estimates of effective pasture area. The work was eventually published in a national publication for rural valuation specialists. A third example of this data repurposing was initially borne out of a formal farmer feedback meeting held in Year 1 of the project. The farmer feedback session held in August 2015 was held with the participant farmers whose land was being surveyed as part of the project. The session gave the farmers the opportunity to review the proposed maps, including one map which visualised copper levels in the plant tissue across the farm. The presentation of this map sparked an unexpectedly enthusiastic response from some of

---

[35] Participant farmers included landowners and farm employees whose properties were surveyed as part of the P2P project. The participant farms are listed Table 11.

the farmers present. This interest was because the farmers were spending significant time and money on supplementing their animals with copper.

Another researcher, Virginia W, found the concept of the copper maps warranted further investigation, but had little domain knowledge of animal nutrition so looked to partner with an expert, external to the P2P project, in the field to better understand the opportunity to create value from the concept. However, IP constraints limited the ability of all the P2P researchers to network beyond the borders of the project. Fortunately, other University staff were permitted to collaborate with the P2P team with permission. Hence, Virginia W was permitted to interact with a Massey University colleague, a veterinarian who was an expert in the field of animal health and nutrition. The domain expert reviewed the feedback from the farmers and provided valuable insight, confirming that the potential to map copper concentrations in plant tissue was indeed of significant value. Interestingly, the veterinarian also proposed a third research question; this time to investigate if an antagonistic relationship between copper and molybdenum could be identified in the data. The researchers returned to the data to identify a relationship between the copper and molybdenum levels, which was validated using the calibration samples from the original study. In the end, the original data were used to answer three important and valuable research questions, see Figure 34.

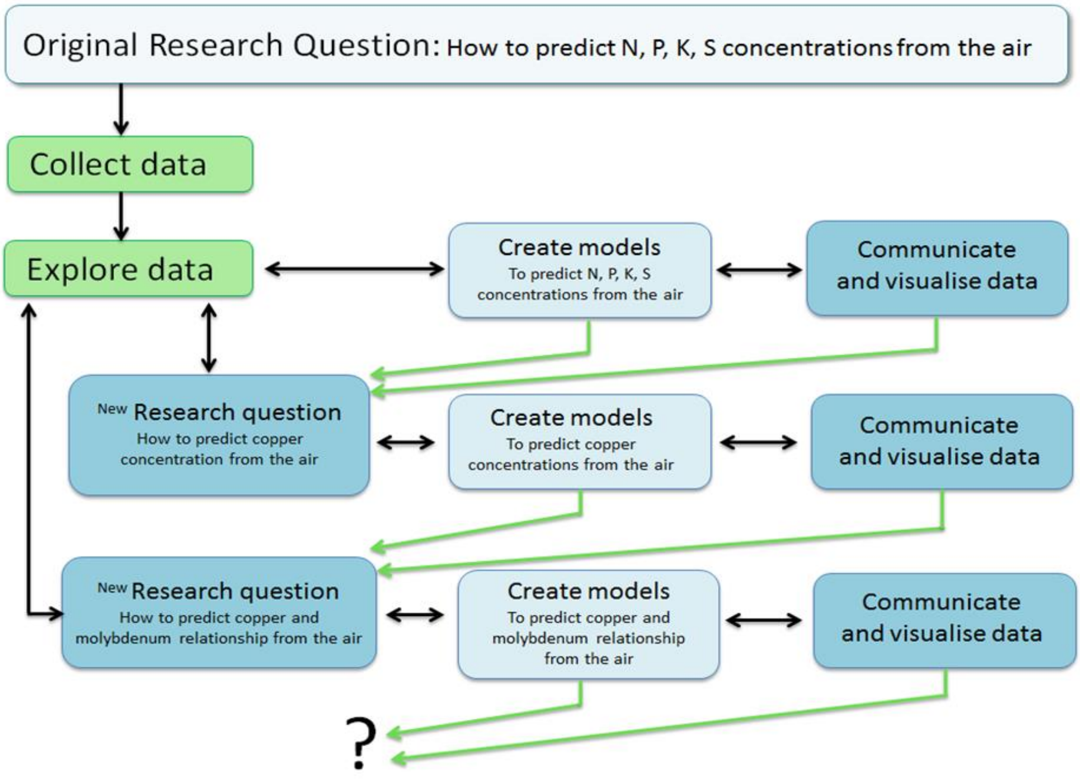


Figure 34 *The iterative, cyclical science process used to create value from the P2P project beyond the original research question (source withheld).*

There are two key points to note here; first, the communication and visualisation of the data played a pivotal role in creating this value from the data. The feedback from the farmers demonstrated that maps could be ‘disruptive objects’ rather than linear, one-dimensional information transmission belts. Second, to realise the value required input of a domain expert beyond the siloes of the original research team. In an interesting footnote to this story, Fred K, the P2P researcher involved with the data processing revealed that the copper map was not modelled on actual data. Virginia W, a member of P2P B-Team was largely working in a silo with some of the post-graduate students.

Consequently, and unbeknown to Virginia, the P2P A-Team researchers had created the predictive model for copper using the nitrogen models they already had. The researchers simply applied a predictive model to estimate copper levels, instead of creating a new model based on plant tissue analysis results for copper. Low nitrogen levels were automatically assumed to mean low copper levels. The P2P researchers ignored the many other factors that could influence copper levels and assumed that estimates of nitrogen could directly translate to copper levels. In a classic example of science being unaware of its limits nor the possible perverse consequences, the researchers were prepared to comment on supplementation requirements based on these unproven relationships. Importantly however, the researchers had retained the physical soil samples for future laboratory analysis, so the model could be tested by correlating the model with the ground-reference data if there was an appetite to test the model in future[36].

In Year 3, I was still unconvinced that I had recorded an accurate picture of the P2P researchers’ practices and processes, so at this point I completed a review of the science process employed by the researchers. In Year 1, the researchers had noticed that their analyses were taking longer than expected as they needed to dip into the data again and again to test out iterations of models until acceptable correlations were achieved. While their science process in Year 1 had appeared to resemble that of the iterative, cyclical model promoted by data scientists, by Year 3 the process had evolved to include multiple entry points. Interestingly, I found that the researchers were returning to both the unprocessed and processed datasets, depending on the process they were trying to adopt, see Figure 35. The role of validating the models had also evolved by Year 3, with the P2P researchers placing more emphasis on validating models earlier in their research process. While this was achieved by setting aside some ground reference points to check the accuracy of the correlations, in some cases more information was needed from domain experts to decipher what the ground reference data measurements actually meant.

---

[36] In 2018 several of the P2P researchers published a conference article mapping copper concentration in pasture. P2P researcher Virginia W was not acknowledged in the publication.

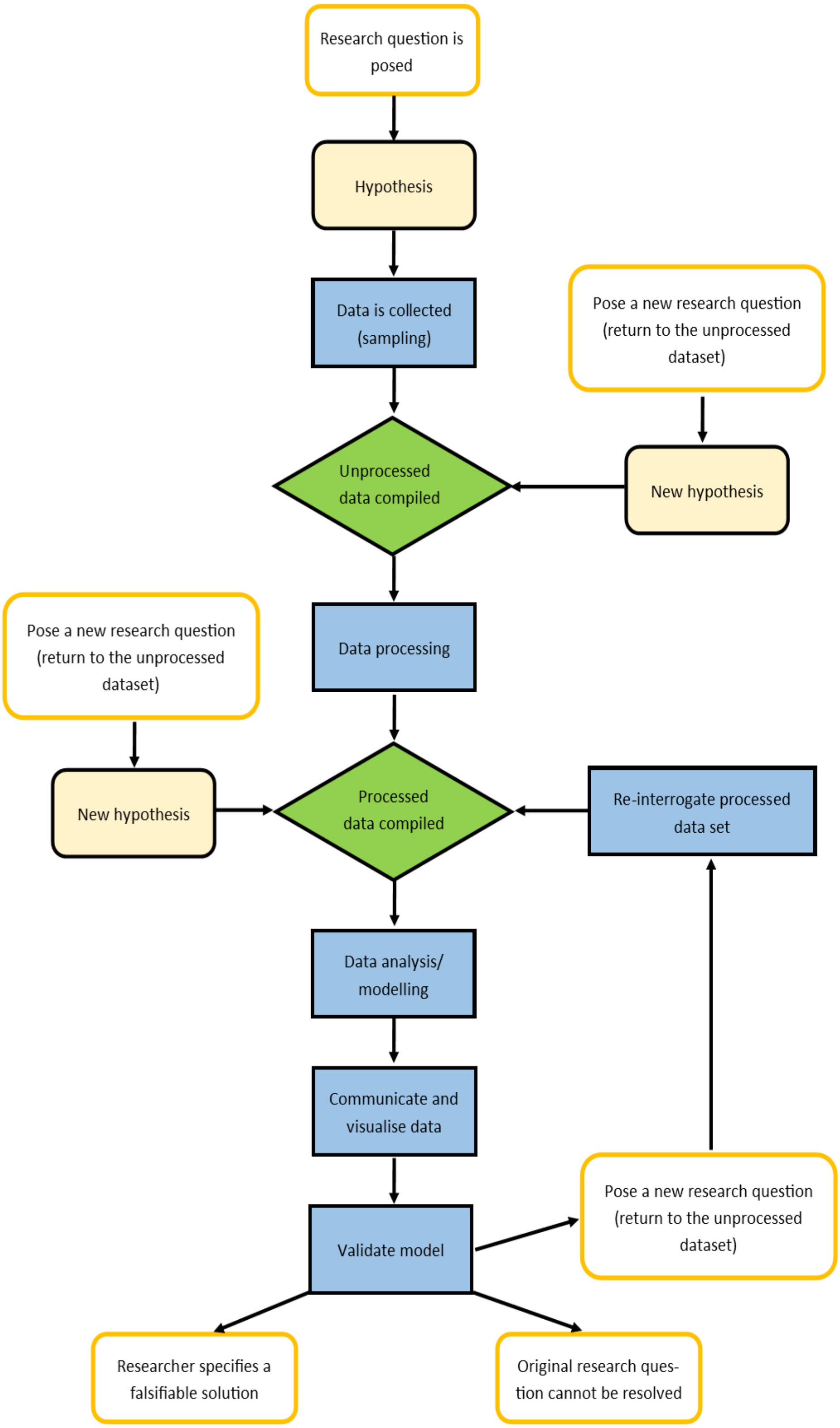


Figure 35 *The process used to create value from a hyperspectral imaging project*

The significance of this shift in process was three-fold. First, the complex science process employed by the P2P researchers represents a radical departure from the classical model of existing PA research. Second, both the processed and unprocessed datasets were considered to hold value in their own right in distinction to the existing assumption in PA that unprocessed data rarely held significant value. Third, the process revealed that while data-science expertise was critical to develop the correlations, domain expertise was needed to validate them.

Data analysis: deciphering the language of the AisaFENIX

With so much time and effort being put into the pre-processing of the data, I was surprised to find that when the data was released to the analysts further processing was needed. In Year 2 of the project, I had collected a detailed record of the journey of the P2P data from Fred K and Lance P. Fred K was primarily involved with AisaFENIX data. He outlined the section of the process he was involved in, i.e. data corrections of the AisaFENIX from the aircraft through to the creation of a data cube, which was recorded in Figure 28.

In terms of the science process outlined in Figure 35, the data produced by Fred K contributed to the dataset “unprocess compiled”. Once Fred K had reached this stage with the data the AisaFENIX was then made available to the other researchers to perform their own processes and analyses. It is this dataset that the P2P researchers mainly returned to when posing new hypotheses from the dataset. The most common process performed after this stage was preparing the relevant ground-reference data to compare with the reflectance data cube provided by Fred K. Lance P talked me through an iteration of this process, see Figure 36.

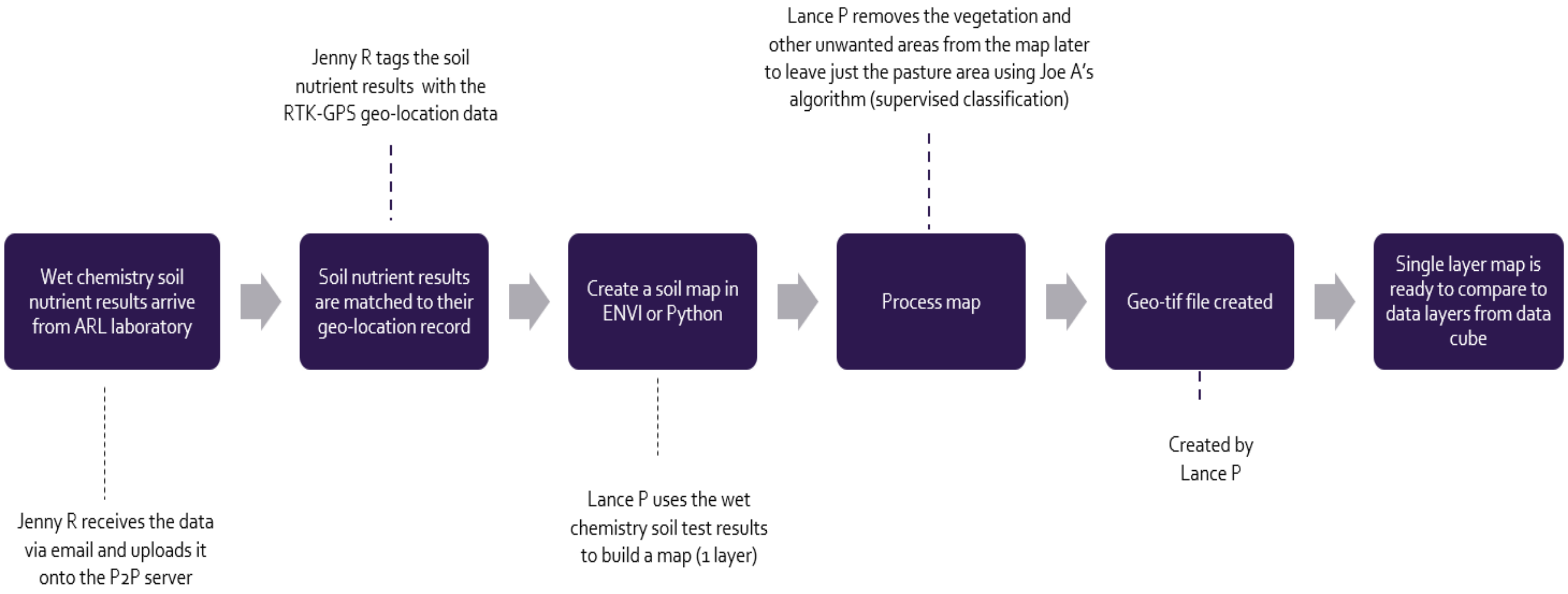
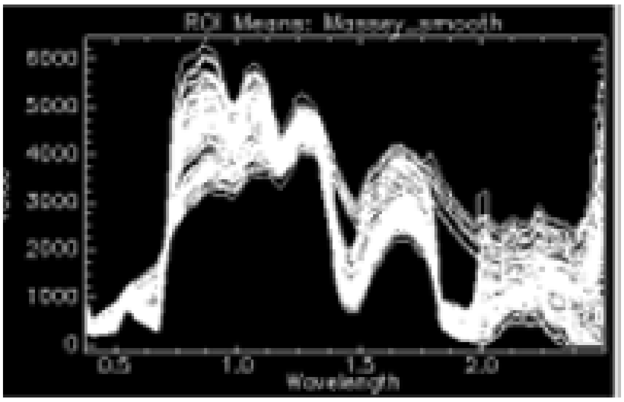


Figure 36 *Lance P's process to prepare wet chemistry data for comparison with the reflectance data cube*

By Year 3 the P2P researchers had developed several different techniques to compare the ground reference maps with the reflectance cube from the AisaFENIX. To explore this concept, I asked the P2P researcher Joe A to share his analysis workflow in a series of interviews, photographs, and videos.[37]

According to Joe A in a Year 3 interview, the AisaFENIX data that was made available to the research team had already been corrected and processed into reflectance data for analysis. With much of the processing already complete, the researchers could all build their knowledge from the same base of processed data:

> *What data did you use for your analysis?*
>
> Joe A: I used the species ground reference data and the AisaFENIX. We had identified areas dominated by manuka and others that had kanuka. We used the RTK-GPS to identify these locations and these were recorded in the dataset. The AisaFENIX data was a reading of reflectance when it came to me for analysis. When I got the data, it came in the form of an image with 448 layers, one for each wave band.
>
> *What does that look like?*
>
> Joe A: I can click on the image at any given pixel and I can extract the spectral profile data for that pixel, which looks like this [picture provided by Lance P].



> *How did you extract the information from the pixel?*
>
> Joe A: So, for each pixel in the image I had 448 data points that represented reflectance at each one of the wavebands. To view the image… it was only possible to physically see three of those layers at any given time. Because each of the chosen layers was converted into red, green, or blue, for interpretation by the human eye.

---

[37] The following sequence compiled from artefacts collected, including excerpts from my research diary.

> Unlike the human eye, which can only view three colours, the computer is capable of understanding the entire spectral profile. So, the computer was used to analyse not just three bands at a time, but all the bands. The reason for using the computer, as well, is that for any given target of interest, the straight reflectance from each band may only form part of the analysis. Other parts of that picture or analysis could be how two bands vary from target to target in relation to each other.
>
> *What software were you using?*
>
> Joe A: The reflectance values were extracted individually using the software, ENVI. Incoming light is considered 100% (total light incoming from the sun that gets through the atmosphere), the light that is reflected is therefore a percentage of that first number. (YR3)

Joe A then talked me through the full process that he used to analyse the data. I was surprised that this workflow included more pre-processing. He explained that the processing until this point was not specific to the analysis task, however processing specific to the type of analysis he was performing was still needed:

> Joe A: Various authors have found that pre-processing of the data to remove noise, or reduce bands, or convert the data to a different format can have improvements for some analysis tasks. As such, the data was pre-processed in a number of ways prior to the analysis procedure below to further refine the analysis tasks. (YR3)

After completing the pre-processing, Joe A moved onto analysing the data to identify algorithms that would accurately represent agreement between the reflectance data and the ground-reference data collected on the day of the AisaFENIX flight. Joe A explained the workflow he used to select algorithms that would help him map specific plants across the landscape. He pointed out that this was the method he used when the target plan was larger than the pixel as he had difficulty in selecting successful algorithms where spectral unmixing was required (see Research Diary Excerpt).

[RESEARCH DIARY EXCERPT] (YR3)

Joe A’s workflow for plant classification

Data arrives in the form of a raw image (see Figure 37) from Fred K, who had completed the pre-processing.

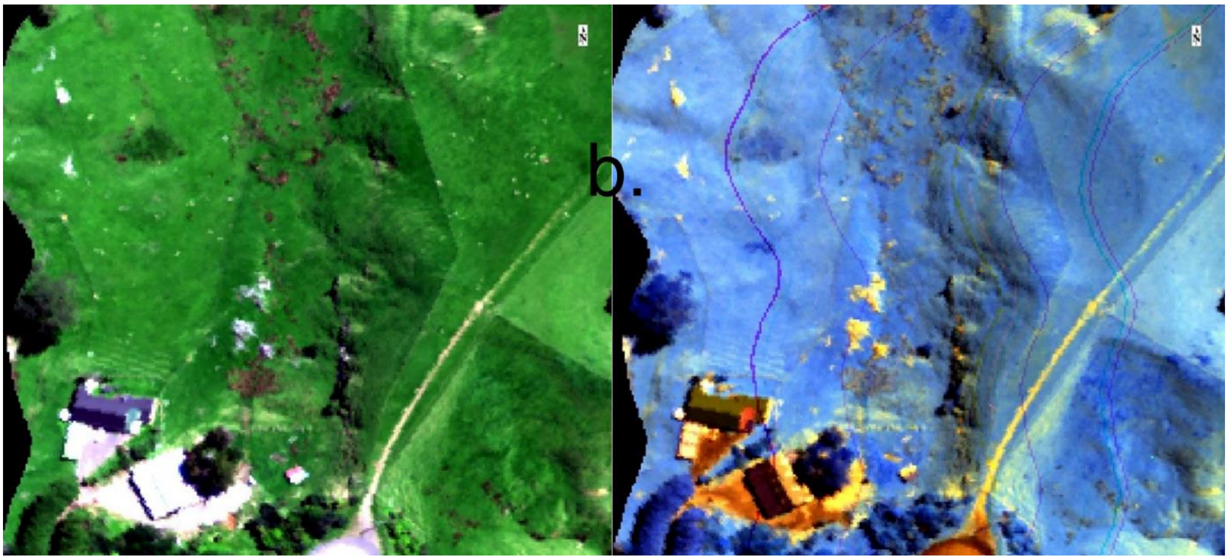


Figure 37 *Visualisation (raw image) of Joe A's data using real colour wavelengths. Supplied by Joe A*

[RESEARCH DIARY EXCERPT] YR3 continued

Joe A’s workflow for plant classification

Step 1. Take all the ground reference data and separate them into testing and training subsets.

Step 2. Target identification. Using the training subset, Joe located the target species, e.g. manuka or kanuka within the AisaFENIX image. I did this by matching the geolocated ground reference data (showing the species is present) to the AisaFENIX reflectance data from the same location (matching to the same georeferenced position).

Step 3. Extract the spectral profile data, see Figure 38, for those target locations and supply it to the computer software (called Envi) as ‘training data’. Training data are an example of the spectral profile for each of those targets of interest. The software then uses that information to identify other positions where the same species may be found in the image.

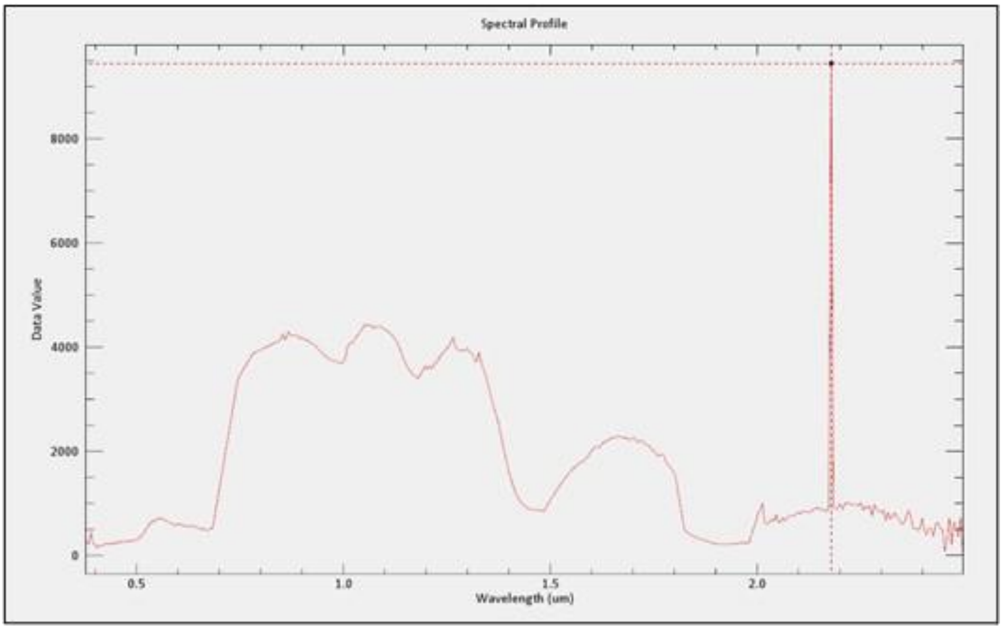
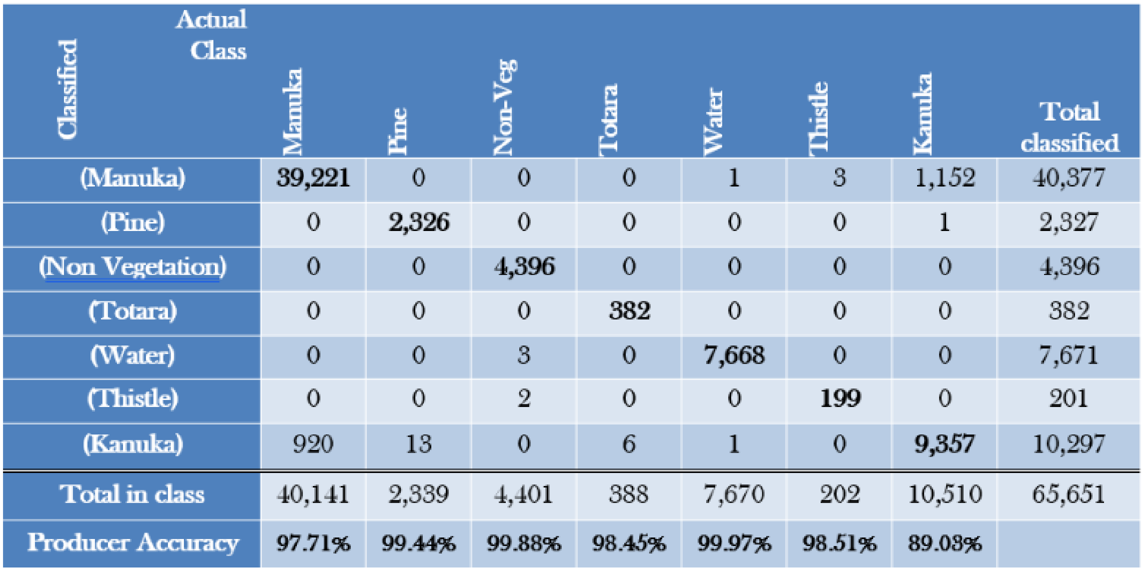


Figure 38 *An example of spectral profile data from Joe A's classification analysis. Supplied by Joe A*

Step 4. Take the output information and compare it to the areas that were held back for validation purposes (the test subset). This provided an indication of the accuracy at those particular known sites, which was displayed as a ‘confusion matrix’, see Table 14. The table shows the confusion matrix and producer accuracies for the classification of image components from Joe A’s hyperspectral imagery.

Table 14 *Confusion matrix and producer accuracies for Joe A's analysis. Supplied by Joe A*



Classified / Actual Class	Manuka	Pine	Non-Veg	Totara	Water	Thistle	Kanuka	Total classified
(Manuka)	**39,221**	0	0	0	1	3	1,152	40,377
(Pine)	0	**2,326**	0	0	0	0	1	2,327
(Non Vegetation)	0	0	**4,396**	0	0	0	0	4,396
(Totara)	0	0	0	**382**	0	0	0	382
(Water)	0	0	3	0	**7,668**	0	0	7,671
(Thistle)	0	0	2	0	0	**199**	0	201
(Kanuka)	920	13	0	6	1	0	**9,357**	10,297
Total in class	40,141	2,339	4,401	388	7,670	202	10,510	65,651
Producer Accuracy	97.71%	99.44%	99.88%	98.45%	99.97%	98.51%	89.03%	

[RESEARCH DIARY EXCERPT] (YR3) continued

Joe A’s workflow for plant classification

Step 5. Produce accuracy metrics. The confusion matrix data allows a variety of accuracy metrics to be produced, such as Producer Accuracy, Overall Accuracy, and Kappa Statistics, which are provided as a percentage of accuracy.

[As a result of this work, Joe used ENVI to produce a classification map, which is broken down into targets, such as manuka, kanuka, roads, water, and any other targets which you ask it to provide.]

Step 6. The classification image (see Figure 39), which still holds its georeference information, is taken to the field for validation of the classification in the field.[38] In ENVI, each classified component can be viewed individually or in unison with some or all included. The image, or a portion of an image (e.g. a single classified zone) can also be exported to ArcMap for data visualisation purposes and/or dissemination as ArcMap is a more commonly used software than ENVI.

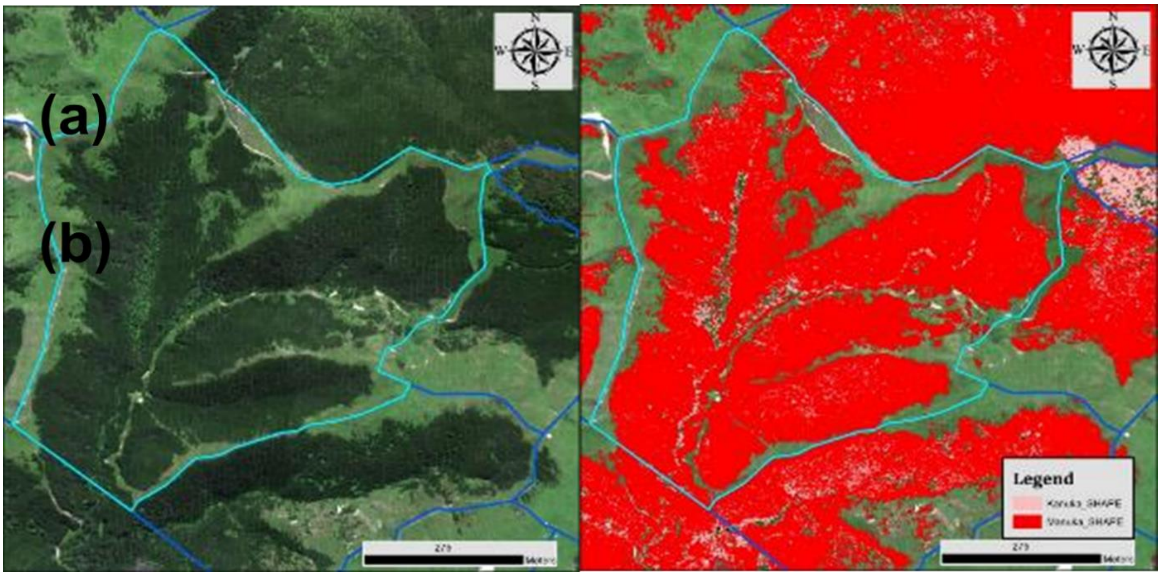


Figure 39 *Joe A's classified image showing manuka and kanuka populations (right).*

[RESEARCH DIARY EXCERPT] (YR3) continued

Joe A’s workflow for plant classification

Step 7. Repeat the exercise for each different potential algorithm (there are many included in ENVI) and data pre-processing approach.

Step 8. Select the algorithm that is the best fit for the analysis of the particular target.

---

[38] The classification image can also be used for a variety of management tasks, such as identification of optimal sites for beehive placement.

Back to the field: verifying scientific outputs in the *Pioneering to Precision* project

Earlier in Joe A's account of his data analysis workflow he mentioned the portioning of test data from his dataset so it could be used to verify selected algorithms either in the field, or if the image and target permitted, via a desk-top exercise. A basic question concerning researchers in the contemporary PA is whether we should believe what we hear about algorithms, and what the algorithms tell us. Understanding how researchers, project managers, and funders can ensure that scientific ideals are maintained when the outcomes of a science process are not predetermined will be important to maintaining public confidence in scientific developments. Spiegelhalter (2020) recently introduced the possible contribution of statistical science to both evaluation and intelligent transparency. Although public confidence in science remains high (Funk & Kennedy, 2017), exaggerated claims about artificial intelligence threaten the uptake by end users (Hall Jamieson, McNutt, Kiermer, & Sever, 2019).

It is important for projects such as the P2P that the algorithms presented to end users are trustworthy and benefits and disadvantages of them are accurately communicated.[39] While Speigelhalter's work focuses on science-making in the pharmaceutical field, there are useful parallels with precision agriculture. That is, researchers and farmers should consider the trustworthiness of claims made *about* the algorithm, and those made *by* an algorithm.

Indeed, all scientific research stands on the ability to perform an independent validation of the presented results (Sagan, 2011). So how could the PGP co-funders and farmer end-users be sure the algorithms being presented were trustworthy when applied to new sites on a new day? One accepted approach is for researchers to document pre-processing steps and ensure data are readily available for experts outside the project to allow reproduction on the same experiments to either falsify or validate previously created results (Christiansen et al., 2018). This was the approach taken by Joe A when he kept data aside for subsequent testing of his selected algorithm(s).

Importantly, the P2P project included a validation phase (Year 4) where the researchers returned to the field to fly farms and test the accuracy of the algorithm against proximally acquired wet chemistry soil tests. The project set-up set validation of the algorithms as its own 'stage' after all the fieldwork was complete. Until this time, the validation of results was based on existing wet chemistry, with sample sites set aside to test using the algorithm selected.

---

[39] This does assume that it is possible to achieve this with complex algorithms, who might exert a calculative agency beyond design. This reinforces the argument that algorithm designers should have sufficient domain expertise to understand the decisions made via algorithms they have created.

There was some disagreement among the P2P researchers as to the role and timing of this validation work. Some of the researchers argued that more significant validation work should be undertaken earlier in the science process to detect any methodological problems. They suggested there should be some independent validation work performed after the first two or three flights to check the accuracy of the findings. Another researcher felt they should sacrifice some of the initial test flights to install LiDAR in the plane to see if this impacted accuracy.

It was not until all test flights were complete that any significant validation work of the algorithms was performed. In hindsight, earlier validation flights may have detected the issues caused by the absence of differential GPS in the aircraft. Importantly, despite being a new area of research, the scientific publications made by the group did not highlight the importance of the differential GPS, so the error could easily be made by fellow researchers in future.

The normal process of falsifiable science would subject a hypothesis to stringent criticism and testing. When the hypothesis successfully withstands a wide range of rigorous tests and is eventually falsified, a new problem, hopefully far removed from the original solved problem, would emerge. This new problem calls for the invention of new hypotheses, new data collection, followed by renewed criticism and testing. And so, the process continues indefinitely. While the researchers were able to keep data aside for future testing, or indeed validation. However, this study demonstrates the difficulties that contemporary researchers face in ensuring their work is scientifically robust. For examples, observations of the researchers working on the P2P project reveal that concerns relating to protection of intellectual property prevented them from publishing their methods and practices. Thus, it becomes impossible for other researchers to redo the experiments and either validate or falsify presented findings. Moreover, the cost of the sensor means few science laboratories could test the results of the study, unless the raw data and processing methods were supplied.

Furthermore, the sociological negotiation around which algorithms ‘make the cut’ is also concealed. When PA researchers are constructing and negotiating algorithms they also build in the question, does it work? After all, researchers want knowledge that works. But what counts as working? That discussion is not had when we fail to have transparency around methods and iterations. It may also be that for all the problems this study identified with the P2P project, it actually worked for key actors in terms of career progression and learnings for future research. In this sense what worked like success and failure is deeply relational.

I suggest that researchers need to reconsider current conventions around the discourse, reflective practices and publications associated with science methods to support data-driven science. For instance, collaboration will be crucial if researchers in the field are to tackle society’s wicked problems. While this progress will centre around appropriate data infrastructures, transparency of failed and successful methods has the potential to accelerate research programmes.

Finding and protecting IP in the *Pioneering to Precision* project

The constraint on the P2P researchers' publishing meant that each time a student or researcher produced a paper, they required the co-funders' permission to publish. When interviewed, the co-funder acknowledged that the company wanted to protect the intellectual property in both the methodology and the dataset itself:

> Steve N: There is an exclusivity period. That is how the PGPs operate. That is the payoff for the government for the investment that they have made. The IP in the methodology is important.
>
> Because, that's [the dataset] the core, you know. There is an ability for people to replicate what you do whether you look at any kind of remote sensing, there is always an ability to replicate. But you just can't acquire a dataset overnight. And an understanding of what is in that dataset.
>
> The IP is in the dataset. (YR4)

Importantly, during the P2P project, some of these requests to publish were rejected or constrained, including, for example, one doctoral student's submission to a conference. Whilst the researchers did not like these constraints, they complied with the requirements. The constraints did concern the P2P Director "Mark L", as some of the young P2P researchers wanted to progress in their academic careers and publishing lay at the centre of that progress:

> Mark L: I think John wants a career in academia, as a researcher-come-teacher.
>
> *So those ROs (Research Officers), are they essentially post-docs? They're recently completed their PhD?*
>
> Mark L: Yeah, like 2012, 2014, 2011.
>
> *So, are they publishing and sort of like gearing into an academic life?*
>
> Mark L: The problem is with the PGP. So, we've already come across that, is the nervousness in Ravensdown about publishing. (YR2)

Scientific reporting in the *Pioneering to Precision* project

These limitations on publishing were a concern for the P2P Director because scientific papers are the currency of the modern academy, and researchers typically benefit from research from their production. The growing volume of PA publications reported in Chapter 2 support the notion that

PA research is no different. However, we also learned in Chapter 2 that academic outputs from commercial researchers are very low compared to outputs from the University sector.

For ‘commercial science’ projects such as *Pioneering to Precision*, the need to protect intellectual property rights limited the opportunities to publish. This is a particularly challenging situation when commercial research is performed in the University setting, such as in a P2P project. F. Provost and Fawcett (2013) raise the potential conflict of researchers working on data projects wanting to publish their work and the need to protect intellectual property. However, they suggest that organisations consider such requests carefully (F. Provost & Fawcett, 2013). While there may be advantages for commercial organisations allowing staff to publish, such as increased publicity, exposure, and external validation of ideas, in the University setting it is more likely that students will want to publish to build up a portfolio of work. It appears that while the P2P project director communicated concern that the publishing restrictions placed on the emerging P2P researchers may impact their career progression, the project and University leaders did not create strategies to work around this issue.

The progression of scientific knowledge is neither linear nor predictable, but it can be facilitated by explicit efforts to promote the accumulation of research-based knowledge. Indeed, a key objective of this work was to accurately document and record the scientific method employed by the researchers working on the *Pioneering to Precision* programme. This ethnographic study reveals that it is difficult to see what is really happening when the scientific outputs don’t accurately reflect the real process – so we rely on observations, interviews, and other artefacts to get to the truth of what is really going on.

In Chapters 2 and 3 we learned that there is little reflection on, and publishing of, detailed scientific methods and practices, including failed methodological iterations. Chapter 2 supports the notions that PA publications are typically concerned with the materiality of the science experiment - instrumentation, machines, and plant responses. The production of knowledge largely depends upon struggling with experimental setups, instruments, and machines. Where methods are published, they are often reduced to simplified, linear processes. Indeed, as we learned in Chapter 4, the publication of one of the P2P group’s first publications on hyperspectral imagery analysis proved to be a catalyst for me to attempt to record the P2P science-making practices and processes as the researchers had re-classicised their iterative methods into a linear process, supporting the notion that of the messy story of science is rarely disclosed in public, see Figure 40.

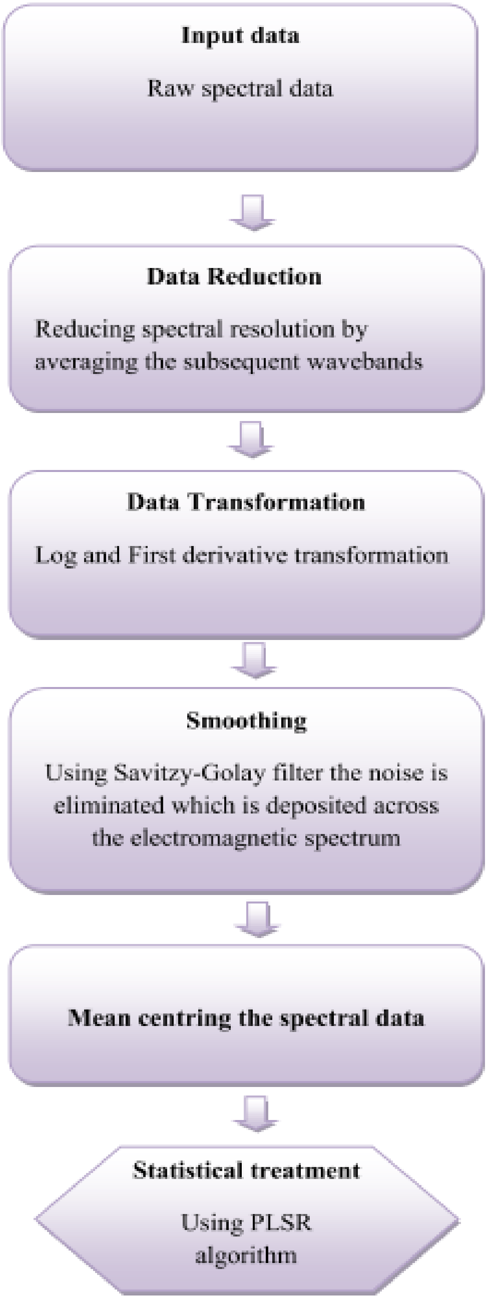


Figure 40 *Flow chart of hyperspectral data analysis depicted in a P2P publication. (Source withheld*

Precision agriculture was conceived as a project of prediction and control, but few things turn out as intended. Researchers are at the mercy of the experimental setup and are endlessly surprised. The less researchers reckon on surprise, the nastier the surprises turn out to be (Cayley, 2009). Yet, most scientific publications in PA only show one iteration of their research – the one that worked. Traditionally, there has been little reflective practice published in precision agriculture.

Failed experiments or models that produce a low correlation coefficient are rarely freed from the laboratory into the light of scientific publication. While this is possibly defensible in the small data economy, where datasets are unlikely to be repurposed, in the all-data economy the publication of iterations or failed experiments could dramatically improve the efficiency and effectiveness of science-making.

A wider implication of strict publication policies is the impact on the construction of science. In Chapter 3 I explained that science builds on previous ideas and is constantly growing. Scientific reporting, where information is deliberately shared with colleagues through the process of peer review is vital to the growth and integrity of knowledge-making (Carpi & Egger, 2011). However, the study of the P2P researchers reveals a crucial limitation of the prevailing research model in PA, which relates to the diffusion of new scientific knowledge and technologies, that is, invariably the practices proposed by researchers rely on *farmers* to implement them.

Precision agriculture has a well-documented farmer uptake problem (Schimmelpfennig & Ebel, 2016; Tey & Brindal, 2012), with the field proving to be a frustrating technological revolution manqué. The prophetic writings of British researcher Jules Pretty in the mid-1990s argued that researchers, extensionists and policy makers underestimate the challenges they faced in getting farmers to adopt sustainable approaches to agriculture (Pretty, 1994). In Chapter 3 I showed that since its inception, PA has been a technology-based field of enquiry. While little analysis of technology push in PA can be found in the literature, there are many examples of technologies tabled as revolutionary that have not flourished in the field. While Gemtos et al. (2011) blame farmer characteristics for the poor uptake of PA technologies, others, such as Jago et al. (2013) suggest that the conventional innovation approach means that some products are not fit for purpose, or do not fit into existing farming systems. This assertion is supported by the plethora of new technologies showcased at conferences that slip into obscurity (Oester et al., 2017). Ultimately however, the reasons that many technologies fail to succeed is still unknown, hence the need to sustain concern and reflection with failure.

The Science Advisor to the Precision to Precision Project "Sam U" was troubled about the unknown reasons that other similar projects had failed "A lot of people have tried it and not been successful. I just made the comment, 'You better look at those to see why they failed'." (YR2)

Publication practices vary between fields of study; in some fields the publication of methodologies and failed iterations is the norm, such as in medical research. However, PA research appears to have followed the norm of the wider agricultural science field, where in recent decades, researchers often provide scant detail of their scientific methods. This practice appears to be relatively recent as a brief content review of scientific papers in the 1980s and before revealed that the methodology sections of papers were highly detailed.[40] I suggest that contemporary researchers are not deliberately obfuscating methods, but rather I suggest that historically researchers only published one methodological iteration because that is all they actually performed. The tendency of PA researchers to only publish a single iteration may reflect these historic expectations.

Renowned novelist and environmental activist Wendell Berry proposes that modern researchers should "keep a more truthful balance sheet in which we show our losses as well as our gains" (Cayley, 2009, p. 162). If nasty surprises or the many failed or deficient iterations are hidden from view, we open a trapdoor for future researchers to fall. Furthermore, research may look easier than it really is – giving funders grounds for unrealistic expectations of timelines and chances of success. The P2P Science Advisor "Sam U" commented:

> Sam U: …so you do have to recognise that they're a different set of pressures and expectations for the outcome. And so, because of that pressure and the concerns, they want to see results earlier than later. And, you can provide results early, but they can also be very misleading. Because if you don't have a robust dataset then the chance of having success initially can lead you down to a false sense of security or a false sense of success.
>
> I think some of it is a lack of comprehension of the complexity of the problem. The complexity of the terrain that they're trying to map. (YR2)

There is a precedent to support the transparent and detailed reporting of both research successes, and failures in precision agriculture. The highly cited Dutch study, "Murphy loves potatoes - Experiences from a Pilot Sensor Network Deployment in Precision Agriculture", was published in 2006. Instead of abandoning the findings of the failed trial, authors Langendoen, Baggio, and Visser (2006) used the opportunity to share their nasty surprises, experiences and learnings with the research community to avoid their mistakes being repeated. Particularly relevant to this thesis, was the author's reflective public revision of their development process and suggestions for researchers embarking on future studies.

Rather than hiding behind data ownership and intellectual property disputes, the authors risked their reputations to disclose their failed methods and unsuccessful project. It appears that the publication of the paper did not seriously harm first author Professor Langendoen's career; two years after publishing the paper he became a full Professor of computer science at Delft University of Technology, where today he oversees an influential research group (IEEE, 2019). Indeed, instead of ending in academic failure, the paper stands as one of the most highly cited publications in the field. While there are thousands of papers outlining their successes, Langendoen et al.

---

[40] This work was completed as unpublished pre-work to the bibliometric study included in this thesis.

(2006) did what few have done before, or after – they were candid about their failure. In doing so, the paper is often cited as evidence that agricultural research is not always successful, an assumption one could be forgiven for making given the dearth of publicly recorded failed projects. It is also evidence that publishing both successful and unsuccessful, detailed methods promote the efficient progress of the field. Moreover, this transparent, detailed, and honest approach may provide a blueprint for academic publishing to progress the field.

A radical departure from classical science model

The P2P project is a good example of working in a new data ecosystem where the consequence of these risks are reduced compared to previous small data studies, because it is easy for the researchers to dip back into data to develop or select another algorithm. In the case of the P2P, the researchers presented multiple algorithms with differing r-values to the funders. By having a selection of unverified algorithms available, the P2P researchers could validate several algorithms in a validation trial; an approach that while unconventional, ultimately proved to be particularly efficient.

Despite their challenges, the progress made by the researchers in this case study does support the notion that advanced datasets, such as the P2P dataset, can have a life and value beyond the initial project with careful planning and curation. Their work reveals some of the value opportunities that exist at a wide range of spatial scales. Indeed, the data visualisations they were able to produce offer an unprecedented view of farming systems from the "bottom-up,".

Shifts in the scientific division of labour: The paradoxical demise of domain expertise

Chapter 3 reveals that agricultural researchers have until recently, deliberately narrowed problem definition and controlled scope to allow the science and the expertise required to know it to be restricted to a predictable, homogeneous group of specialist experts. So, soil research is done by soil scientists, plant research is done by plant scientists etc. This classical method for constructing data *and* for producing the expert researchers needed is a predictable and controlled process, although itself is a product of a modernist ontology that divides knowledge up into a series of discrete silos. However, domain specialists and farmers were deliberately left out of the P2P team. Instead, the focus was on remote sensing expertise and algorithm creation on the expectation that algorithms would supplant domain knowledge.

Observations of the researchers reveals that by early-Year 2 of the project, the researchers' work processes had become dominated by data analysis. Most of the P2P science team were accustomed to working with variable data, however, re-analysing data from a live project to propose and answer new research questions was unchartered territory for all the researchers. Importantly, the researchers were not necessarily always sure what they were looking for in the hyperspectral data, as the AisaFENIX sensor had rarely been used for this purpose before, so they relied on statistical correlations between the ground-reference data for the target to point towards successful algorithms. For example, in Joe A’s account of his plant classification analysis, he identified the location of pixels containing manuka and kanuka and then searched the various bands of data to find the bands that showed the best correlation between the two.

By this stage the researchers themselves were increasingly working in silos. This is particularly important as the team had very little pasture or soil expertise, as well as very limited experience in the *creation of algorithms*. With tight timelines and budgets, the P2P Researcher “Lance P” revealed the team had turned to the most modern of resources, Google:

> Lance P: If I do an agronomy PhD, no-one asks you, everyone knows agronomy, it is just the management, the agronomy is nothing. You can't compete with other people, so applied science is completely different; there is a lot of opportunities.
>
> *So, you saw a niche?*
>
> Lance P: I was struggling in the beginning but later on you realise the advantages of applied science.
>
> *But when you made that transition did you have to go off and do a whole lot of courses, basically?*
>
> Lance P: I didn't do it, but the thing is, interest. interest.
>
> *So, did you go on short courses though and things to pick up particular, you know, tools. Like tool-centred training right - which is what I do as well, if I need to know X, here is the place to go.*
>
> Lance P: But here most of them are very expensive.
>
> *Yeah, I know.*
>
> Lance P: You know three or four days is a couple of thousand dollars. So, I use Google. Google is a very, very good thing. You get everything. Well, there are threads of discussion and you soon discover there are a bunch of people.
>
> The problem is that you need to sort out... you get a lot of information.
>
> *And not all of it is good?*
>
> Lance P: Sorting out is the only issue.
>
> *So, you're basically self-taught?*
>
> Lance P: Yeah, analysis, yeah, yeah.

> *And problem driven? No, it seems as though you came out of pure science and you've moved into this data analysis - it's like what problems do I need to solve? So, what tools do I need to solve those problems?*
>
> Lance P: Yeah. (YR1)

A question that arises from this, and similar interchanges with other project participants is why there was such little investment in professional development in the P2P project? Despite the project budget exceeding NZ$10 million, the project was largely resourced with post-graduate students and recent doctoral graduates, none of whom had expertise in processing airborne data, or creating algorithms from all-data sets. Furthermore, there appeared to be little appetite for investing in upskilling the members of the science team, other than using their time, which was an inexpensive input compared to the wider budget.

It is important that data are accessible and usable for different users. What appears to be simple is in reality a highly complex process requiring extensive systems knowledge. Contemporary co-production innovation processes employ both farmer knowledge and domain expertise to ensure that this is achieved (Eastwood et al., 2012). In contrast, the P2P researchers and funders took the view that such expertise was not needed. There was a dismissive attitude towards specialists such as agronomists and soil scientists. When questioned, the funder “Steve N” felt that excluding domain expertise was justified as it may be a barrier to progress:

> Steve N: That's actually an interesting one where you've got what you call a very long standing and traditional expertise that is embracing the newer technology. In other words, like this morning saying well... traditional soil testing is dead, as we know it.
>
> Yeah, so if they can't handle it, then you can't involve them. Because they won't adapt or adopt.
>
> So, and I think if push comes to shove, they would be hesitant to change shall we say. (YR4)

With little input from plant science domain experts both in the planning and execution stages of the project, researchers took a naïve empiricism ‘to measure is to know’ approach to knowledge creation. This tactic came with risks, especially when seeking to understand complex biological systems (Hoffmann, 2011). For example, creating reflectance models of homogeneous plant systems is difficult by itself, and the reflectance from the complex hill country pasture systems being modelled in the project could be influenced by many, many more factors that could not be controlled in this study.

In remote sensing it is imperative to understand that you are actually measuring what you think you are. Indeed, the complexity of the biophysical characteristics of plants can make determining accurate relationships between nutrient levels in plant tissue and soil nutrient levels extremely challenging, even under laboratory conditions (Morgan & Connolly, 2013). So, including researchers with expertise in remote sensing plant physiological characteristics may have assisted in the experimental design of some of the project's key research to ensure the approach taken was appropriate.

For example, one student was tasked with modelling a reliable relationship between the plant reflectance readings (as a proxy for nutrient levels in tissue) and phosphate levels in the soil.[41] The key piece of information that seemed to be missing for this work was which part of the plant tissue should be measured to create this relationship. In the end, a mixed tissue sample (measure everything) approach was adopted after consultation with the co-funder. Mixed tissue samples are typically used where assessments of an animal's diet are made, such as when measuring metabolisable energy intake. A mixed tissue sample approach for determining soil nutrient levels was a significant departure from the accepted contemporary pasture science and the student struggled to find meaningful relationships from the data.

The complexity of the task at hand was likely due to the relationships being impacted by many, many factors, many of which cannot be controlled in the field. However, in the laboratory it is possible to control many key variables, possibly the most important of which in this context is soil moisture content. The complexity of the task was exacerbated by the group's limited understanding of pasture systems and farming practices. The project's researchers did not comprehend that New Zealand's hill country farmers generally apply fertiliser to promote growth of clover species, especially subterranean and white clovers. Contemporary pasture scientists have discovered that very specific tissue sampling methods are needed to derive accurate results when looking to determine fertiliser requirements for these species. For example, (Olykan et al., 2019) recently demonstrated that the laminae (leaflets) of subterranean clovers have very different plant tissue nutrient levels than petioles (leaf stems). The authors contend that laminae (of pure clover samples) be analysed to determine fertiliser recommendations (Olykan et al., 2019).

When considering the practicalities of *Pioneering to Precision* data collection, where subterranean clover was present, the reflectance readings were likely to be dominated by the canopy of clover laminae. Yet, the P2P researchers chose to adopt a sampling method historically used to reflect animal intake rather than techniques designed to simulate animal grazing. The method employed involves cutting entire plots of mixed tissue samples. The samples collected contain not only significant clover petiole material, but also a mix of grasses and any other plants present. This approach invariably leads to a huge amount of data noise in the collected plant material for nutrient analyses, making it unlikely that an accurate assessment of the clover's

---

[41] The student's thesis is embargoed so some details of this example are deliberately vague.

nutrient status or a robust relationship between the remotely sensed award and underlying soil nutrient status could be achieved.

Despite the funder’s earlier concerns that including domain experts in the science process may slow down the progress of the project, in Year 4 of the project “Steve N” acknowledged that the input of domain experts may have been useful in deciphering the data:

> Steve N: ... it is easy to get an output but if you don’t actually understand what’s driving that output, you may go to somewhere else, different farm, different geographical region and the output is erroneous because you don’t understand what the triggers are underneath it. And, so it could lead you down a pathway. So, the simplest way to say it is that if you just chuck all the data together and say what’s the relationship, oh yeah we’re going to use that without actually understanding what is actually driving the numbers, then you will come unstuck. Because you will apply it in an inappropriate situation. (YR4)

Miller (2010) suggests that progress in knowledge discovery in such ecosystems is being hindered by researchers who ignore the rich body of theory and models already available. He sees these as valuable sources of background knowledge that could guide the exploration of massive spatiotemporal databases such as those produced in the P2P project. Indeed, rather than being atheoretical or anti theoretical, Miller posits that the knowledge discovery process harmonises well with traditional avenues to knowledge construction in science (Miller, 2010). Miller suggests the knowledge discovery process *benefits* from domain expertise and theory to focus searching through vast information spaces and distinguish between real and spurious patterns discovered in these spaces. Furthermore, to exclude domain expertise from the innovation means that science cannot build upon itself; essentially the domain knowledge is captured in a static blackbox, where self-reflection and improvement is impossible.

Challenges of working with AisaFENIX data

Understanding and addressing the complex implications of the arrival of all-data datasets is a key challenge for researchers hoping to realise value from these technologies. Two downsides of working with airborne-acquired AisaFENIX data compared to proximally acquired data were discovered during this research. The first relates to the characteristics of the data generated. Laboratory-based proximal AisaFENIX data are relational in nature; the controlled laboratory environment means that temporally acquired data are commensurable. However, the airborne-collected data proved not to be relational in nature. While they contain common fields, atmospheric corrections and in some cases variability due to calibration of the instrument means that conjoining of different data sets was not possible. Thus, as discussed in Chapter 5, the P2P researchers needed to generate significant ground-reference data from every sampling event to provide reference points for analysis.

The second key challenge encountered by the science team was that significant data processing was needed to render the data ready for analysis. This requirement was not factored into the already tight project timelines, putting the timeliness of the project in jeopardy. Indeed, the P2P project was framed around the notion that innovation is stable and predictable, but things such as the data processing requirements, which the funders and researchers thought were stable were not. So, while *Pioneering to Precision* highlighted the promise of hyperspectral sensing and imaging for crop characteristics, "Joe A" (P2P Post-graduate student) revealed deep challenges generated by the data itself:

> A major problem is the quality of ground data. Errors in the calibration data are repeated through the extrapolated model, nullifying the effort that was expended to collect the data. – Accuracy is highly dependent on the quality of the calibration data. Errors in the methodology early on, if not identified and corrected, can persist through the programme.
>
> The AisaFENIX instrument produces raw data in the form of digital numbers for every pixel. The raw output from the sensor is converted to a measurement of reflectance. The image is then processed and calibrated, and algorithms are models developed from that. A model is something that can be applied in multiple places, so far in this programme they have been calibrating individual images one at a time. The ultimate goal is to have an algorithm that can be applied across multiple landscapes that does not require individual calibration of the image. (YR4)

Other data-related issues encountered by the researchers during the processing stages, but were not addressed, leading to problems in the analytical stages of the project. For example, the researchers found problems with ortho-rectification and -mosaicking the images, some images were missing large strips of data where the sensor stopped working, and there were unexplained lines in the data that were interrupting the creation of algorithms. Furthermore, no LiDAR data were available for the site, nor was the aircraft fitted with a LiDAR sensor. Hence, as no reliable terrain model was available, a flat terrain orthorectification model was applied to the dataset instead of one that fitted the image to the correct terrain model.

A consequence of this was that the data collected from a single pixel could be larger than the one metre of the pixel due to terrain geometry. This was a problem because the stretching of the image meant that it was almost impossible to accurately account for shadows in the image. P2P Researcher "Fred K" explained the issue:

> Fred K: So, I do this in ENVI.... So, these are the softwares (sic). And in ATCore4 we can only read this one, the BSQ file and then I do 'Flat Terrain' atmospheric correction.

> *So, what's the ‘Flat Terrain’?*
>
> Fred K: It assumes the topography is flat. We can do 'Rugged Terrain' as well when we have to use a DEM but at this stage we don't have a good DEM so we stick to this one and we started this one because we didn't have DEM and we will continue using that for this.
>
> *But in the future if you have LiDAR with that?*
>
> Fred K: Then we should use the 'Rugged Terrain' option, which is much, much slower but it will give you better results because then using the LiDAR for example we can calculate the shadows and we can apply different corrections for shady areas and sunlit areas and so on.
>
> *So, does that smooth out the data so that you are getting a more accurate reading of the bands?*
>
> Fred K: Sort of yes, because we can also correct for other issues like with directional reflectance so for example the plane is flying from north to south or from north from south and so on and then it can do different readings.
>
> *So, the LiDAR... why is it so much slower to use the 'Rugged' is it just because it's got more corrections going or more computations going on?*
>
> Fred K: Yeah More computation. Because it calculates the exact geometry for every pixel relative to the plane and relative to the sensor and it processes also the topography, so it calculates the slope and the aspect and also the shadows at that time of the day. So, it takes a little bit of time... Using this one it usually takes like 5 minutes, everything included. If you do... control it from a user interface. The 'Rugged Terrain' would take ten, fifteen minutes at least. Computationally more intense. (YR2)

Practices of science: Precision is not accuracy

While the P2P were spatially very precise, the results were not always accurate. Given the name of the PA includes the word ‘precision’, there is surprisingly little discourse in the field regarding the role of precision (and accuracy) in science-making. Precision and accuracy are two *different* ways that researchers think about error and the P2P Researcher “Lance P” recognised that while the instruments they employed were very precise, this did not automatically equate to accuracy or meaning:

> Lance P: The Fenix is basically an on-board instrument; the camera it collects the… just the numbers. So, it doesn't have any standard values so in order to make it standard values....

It doesn't have standard values for what?

Lance P: “These layers”

For anything basically?

Lance P: For anything, yeah.

It's just saying these are the light scores basically, these are the spectral scores?

Lance P: Yeah, yeah. It's only just it collects the light which is reflecting from the ground objects. It collects just the numbers, that's it.

It's very precise, but it's completely meaningless?

Lance P: Yeah. (YR2)

Indeed, accuracy refers to how close a measurement is to the true or accepted value, whereas precision refers to how close measurements of the same item are to each other (University of Hawai‘i, 2020). Precision is independent of accuracy. In remote sensing projects such as the P2P project, it is possible to be very precise (for example very highly spatially detailed), but not very accurate. It is also possible to be accurate without being precise. The best quality scientific observations are both accurate and precise.

There is enthusiasm about the increasingly high resolution of remote sensing, but precision does not equal accuracy. This reinforces the findings of Khun et al. (2016) and also illustrates that high resolution does not necessarily mean better quality information. In a trial that compared the accuracy of two remote sensing devices for measuring nitrogen levels in corn, the lower resolution GreenSeeker Variable Rate Application and Mapping System (Trimble, Sunnyvale, CA) and a converted commercial 12-megapixel camera (Canon Powershot S110) mounted on a fixed-wing UAV (eBee Ag unit, senseFly SA, Cheseaux-Lausanne, Switzerland). The researchers concluded that the lower resolution device produced more accurate information than its higher resolution counterpart (Khun et al., 2016), probably due to the Greenseeker being an active sensor so its radiometric information was less influenced by lighting conditions.

Observations of the P2P project support the notion that precision is not the same as accuracy. For example, the project’s researchers needed to align datasets as layers to model relationships between the airborne data and the geolocated ground-reference data. To do this the P2P researchers placed great emphasis on accurate geolocation of the ground-reference data, installing permanently geolocated markers at predetermined locations using (three-centimetre accuracy) Leica RTK-GPS technology.

This approach was successful insofar as the researchers were able to use the marker pegs to ensure that temporal readings were taken from the same location each time. Unfortunately, this method also relied on highly accurate GPS location of the second layer of data from the AisaFENIX. However, inaccuracies in the geolocation of the aerially acquired data were not detected immediately, as no checks were made. It was not until late in the data collection schedule that problems became evident. It emerged that the absence of a differential GPS system in the aircraft meant that the alignment of the AisaFENIX data layer was not as accurate as predicted. The inaccuracy for the second layer effectively meant the effort spent on the previous layer was wasted time, energy, and money.

The alignment issues between the aerially acquired and ground-reference data raised questions about the validity of applying the model across more than one farm, as this interchange during a project meeting in Year 1 of the project demonstrates:

> Lance P (P2P researcher): The thing is, I told you before... because the mapping, we need to change the strategy of what we are doing because so far only individual calibration models of mapping. So rather than using only just one model based on the ASD data and applying to all the images. There is a more reliable and more accurate and scientifically sound.
>
> Mark L (Director): Is that when you're validating the Fenix data?
>
> Joe A (Post-graduate Student): So, what do we do?
>
> Lance P: Not doing any Fenix validation because Fenix data is not really... Fenix data... the pixels exactly are not representing the sampling location. Even though we are using... in fact that is the wrong pixel we are selecting but trying to make a good correlation but that's not making real values on the map. So rather the ASD is pretty much accurate and all the data is good and robust and properly validated so only using that models and imposing on the Fenix data. (YR1)

In this interchange Lance P suggests relationships and correlations to target parameters should not be carried out using AisaFENIX data because the aerial and proximal data were not well aligned. He suggests only using the ASD for model creation, and then applying it to the AisaFENIX data without regard to ground conditions, ground-reference, or validation. Even at this late stage in the programme, with numerous flights undertaken (and the work that goes with them) and millions of dollars spent, there was no clear methodology in place for the project. While this may initially sound like a failing to a classically trained plant scientist like myself, it is an existential fact in the new data ecosystem that the data will often arrive before even the hypotheses are formed, let alone the method. For the importance of experimental design and consequences of methodological failure in the new world have shifted. Iterations are to be expected.

Without a clear methodology in place and faced with the misaligned datasets, the science team pivoted and deferred to the position of using data they were familiar with and applying it to the airborne data. The highly accurate spectral resolution of the ASD meant the researchers could model the relationship with the wet chemistry soil nutrient data over time. Interestingly, despite having highly precise measurements, the researchers aggregated the ground-reference data from all the flights to create a less detailed average figure that could be imposed on the AisaFENIX data. Moreover, it does not appear that AisaFENIX data were validated during the research stage and the relationship was purely based on correlation. While this may make sense at first glance, omitting validation is inherently unsafe (Passioura, 1973), and in hindsight validation work at that point may have benefited the project in the long term.

The delay in validating the data was not the choice of the researchers but was instead an example of the intractability of linear science processes. In the original project plan, the funders had scheduled validation trials once the test farms were complete. The funding associated with the validation was linked to the timing so despite the wishes of the researchers to conduct validation trials after the first few flights, they had to rely on data partitioning and testing of algorithms against the test datasets to validate their methods until after the algorithms had been developed based on the results from all eight trial farms.[42]

Conclusion

This iterative, complex data journey in the P2P project is a far cry from the linear journey data of classical science traditionally promoted for agricultural science. It has a quite different logic from the sampling and generalisation of the classical model and promises to solve the problem of not capturing variability in the application world. Observations of the researchers grappling with the role of data in the science-making process in the ethnographic account was a turning point in my research. It ultimately led me to consider what such a radical shift in the data ecosystem could mean for PA researchers.

I discovered that the P2P researchers were not only working with data from big data technologies, but instead, they were working with data from a complex community of technologies. These technological assemblages and their data may have complex and far-reaching implications for PA research and the science-making teams behind it. Understanding the data characteristics of new technologies and their assemblages will help future researchers to adjust their science-making process and resourcing to accommodate the emerging data assemblages and to optimise the value generated from the technologies. Observations of a science team working in the all-data economy

---

[42] Data partitioning refers to the process detailed earlier by Joe A, where the researchers kept some ground reference data aside to test the accuracy of the selected algorithms.

identified that a single data collection event not only provides the data used to predict levels of soil nutrients related to growth, but the data carry information that can be re-used to solve more than the predetermined problem.

Chapter 8: Embracing the beautiful mess: realising value in contemporary precision agriculture

Introduction

Precision agriculture research is shaped by both social and material forces and is therefore sociomaterial (Fenwick & Nimmo, 2015). When I began this PhD journey the emphasis of my initial investigations was on exploring how farmer uptake of technologies such as the hyperspectral sensing and imaging could be improved. However, in the first few months of the project a theme emerged from the early data analyses that changed the course of the study. Early interviews and observations of the P2P researchers suggested they were struggling to adhere to the linear science process prescribed in the project set-up as they wrangled the deluge of data produced by various technologies, including the AisaFENIX. It became apparent that the data generated by the project were unfamiliar and challenging to the group and the researchers were adapting their scientific method to refine value from the unexpected characteristics of the data, and this had consequences for data processing and analysis.

My preliminary work outlined in Chapters 2 and 3, combined with these early observations revealed that PA research is awash with data. However, there has been comparatively little consideration of the nature and capacities of data itself. My supervisors and I agreed that this emerging theme deserved deeper analyses, ultimately shifting the trajectory of my own research towards using insights from ANT. Hence, I set about employing an ethnographic analysis to inform my research around how the arrival of data-rich tools is impacting science-making in precision agriculture. The tools and sensitivities of an ANT approach (Law, 2009) helped me to make sense of what was going on in the P2P laboratory that was the object of my enquiry. In particular, ANT was helpful as it seeks to describe how connections that link humans and non-humans (in this case data and technological assemblages) come to be formed, what holds them together, and what they produce (Desai et al., 2017). Indeed, I found ANT offered a powerful resource to explore data, networks, and scientific processes emerging in the field of precision agriculture.

My reflections on the challenges the P2P researchers were encountering at the time also led me to reflect on *my own* research practices. Having come from a quantitative methods background, I realised that I too was struggling with a need to shift away from linear thinking to capitalise on the data I myself was gathering. In my previous career in the agronomic sciences, revisiting data was usually unfeasible due to the nature of the datasets gathered. With the encouragement of my supervisor Associate Professor Wood, I, like the P2P researchers, stepped out of my comfort zone and adapted my line of inquiry to ask deeper, more relevant questions informed by real-time observations, rather than strictly following a predetermined research path.

While relics of these early inquiries reside in this thesis, notably through my questioning of the field's progress in terms of wide-spread farmer uptake, the focus of the study moved to the researchers themselves and the data characteristics of the nascent technologies they employ. The research questions were revised to support the development of theory on how the arrival of these nascent technologies may transform PA research itself. Hence, the primary research question posed in this thesis is: "How do the characteristics of technological assemblages affect precision agriculture research and development?"

To address this research question, I began by exploring the following questions:

1) What are the characteristics of new precision agriculture technologies?
2) How do information-rich technologies promise something different for precision agriculture?
3) What challenges, risks and opportunities emerge from the assemblage of data rich technologies in precision agriculture?
4) How do big data influence science-making in precision agriculture?

Spending 30 months observing the P2P research team through the lens of an ethnographer and domain specialist yielded valuable insights into limitations in our understanding of who contemporary PA researchers are, how assemblages of PA technologies are changing, and how the arrival of data-rich technical assemblages may impact the practices and processes of science-making.

To support this work, I also assembled and analysed a longitudinal dataset of over 11,500 PA publications, which provided valuable insight into the trends in the field over time. This bibliometric study coupled with the ethnographic account provided rich data from which many learnings have emerged. In this chapter, I examine some of these learnings and discuss processing methods and practices that may assist researchers when repurposing PA data. I suggest how funders and policy makers can support value creation through flexible project management and the provision of authentic data worlds where collaborations between data, machines and humans thrive. In addition, I explore some of the challenges and opportunities that researchers and funders face in verifying and communicating data in the contemporary context. The thesis concludes with a reconsideration of how PA research funding and practices could be adapted to meet the needs of contemporary science-making, and, finally, some suggestions for further research.

Findings and discussion

This section brings together the findings and themes emerging from the bibliometric study and ethnographic account and draws conclusions on what data-driven science-making might consist of for PA research. While this thesis is not designed to provide a prescription of what can be done to overcome problems identified with science structures in terms of the arrival of data-rich technological assemblages, I do reflect on the implications of the research findings for researchers, policy makers and project funders more broadly.

The findings paint the P2P science-makers as improvisers who adapt their methods and practices to accommodate a rapidly changing data ecosystem. Employing an ANT approach meant I could investigate the arrival of complex technological assemblages and explore how they are compelling researchers to employ novel methods that involve and generate new strategies for refining value from data. The outcomes of this thesis contribute toward understanding of contemporary PA research in three ways. First, the work provides important empirical evidence that PA research is increasingly moving from single technology focus to assembling and adapting a complex community of technologies. This is supported by the findings of the ethnographic account and bibliometric analyses in Chapter 2. Moreover, the bibliometric analyses found that while big data technologies have indeed arrived in PA research, complex assemblages of ground reference data are also being collected at unprecedented rates to support the creation of algorithms that calibrate these nascent devices. An important early finding from the bibliometric analysis was a somewhat unexpected trend towards increased use of small data technologies in PA research. The ethnographic account extended these findings to demonstrate that indeed, not only were the P2P researchers working with big data, they were capturing, processing and analysing data from multiple sources, and that the operationalisation of big data relies on its integration with ‘small data’. Importantly, the findings of Chapters 2 and 3 showed that these data assemblages are rarely described. Instead the focus of reporting is on single technologies, even though they are part of a complex wider technological assemblage. Thus, this study also adds specifically to the PA literature by advancing knowledge of the ontological characteristics of the data assemblages gathered together in the P2P project. The central characteristics of the P2P dataset are identified and described both in terms of their characteristics and their sources.

Second, this study provides insight into what an evolution of data assemblages might mean for scientific enquiry in contemporary precision agriculture. Analysis of the ethnographic account revealed that the data characteristics from the P2P’s all-data dataset had a profound impact on science-making and the scientific division of labour within the P2P research project. A key implication of these data characteristics was a shift in the science process, where effort moves from the planning and performance of a data collection stage to a messy, complex process, where data collection is still important, but where substantial expertise and resources also need to be invested in the data processing and analysis stages compared to the existing paradigm.

This study recorded the P2P researchers as they adapted their science making methods and practices to accommodate the characteristics of the new data ecosystem. I discovered that while the project was set up to accommodate significant effort and expertise in the data analysis stage, the significant resources required in the data pre-processing phase of the project was unexpected, and consequently sufficient provision had not been made for it in the project planning. Interestingly, almost all of the effort expended in pre-processing of the P2P data was invested in the aerially acquired hyperspectral data. So, while the aerially acquired AisaFENIX data are inexpensive to collect, the expertise and expense involved in preparing the data for analysis was not factored in and need to be considered in the time schedules and budgets for division of labour on science-making projects.

Instead of all the data being expensive and arriving late in the P2P project, a complex hybrid of data collection was observed. While the data from the AisaFENIX were cheap, massive, varied and arrived early, the collection of the small ground reference data that supported the analysis remained laborious and time-consuming as described in Chapter 5. The characteristics of the all-data dataset also had implications for the data analysis stage of the P2P project. While the P2P researchers initially thought their research efforts would focus on deciphering the language of the AisaFENIX, the task was in fact much more complex as they also needed the AisaFENIX to talk to the small data from the ground reference measurements. This was a problem for the researchers because the project was set up and funded on the basis of the classical science method, where much of the effort is up front and value is refined through a linear process.

Third, the study builds an understanding of how reflective practice can identify and inform changes in science-making practice. Early in the thesis I found that there is a scarcity of social science research in PA, with farmers and researchers being relatively ignored compared to the technological focus of research. There are two sides to this; farmers are positioned as distant and largely ignored end-users, while there is a lack of understanding about how PA research is actually done by the researchers that do it. In Chapters 6 and 7, I found that assemblages of big data technologies and their supporting data are changing how researchers are doing their work. However, the field needs to reflect on and revisit our science-making to ensure our practices and processes meet the demands of contemporary research. It is through reflective practice and social understanding that we will understand the significance of the data ecosystem in which PA now operates in.

A new data ecosystem for precision agriculture

While New Zealand’s vibrant agricultural research scene already generates large volumes of data (Medyckyj-Scott et al., 2016), PA researchers and funders are yet to establish a data ecosystem that sustainably supports science-making in this new era. The advent of technologies such as the AisaFENIX provides an opportunity for researchers to explore multi-layered, geolocated and fine-grained data at an unprecedented scale. However, while the availability of vast quantities of data offers researchers new opportunities, these new data assemblages also pose infrastructural challenges. Sharing data publicly can accelerate scientific discoveries and save research funds and effort by avoiding unnecessary duplication of data collection. However, global investigations into the repurposing of research data show that data quality and data infrastructures are often insufficient to facilitate the reproduction of results, or the repurposing and reanalysis of data (Roche, Kruuk, Lanfear, & Binning, 2015).

I posit that the behaviour of all actors, including researchers, funders, farmers, and policy makers working in these new data ecosystems will strongly influence whether value can be created from data beyond the initial projects that data were gathered for. Earlier I reported that some of the data characteristics in the new ecosystem strongly impact science-making. For instance, in terms of velocity, much of the data not only arrived at speed, but arrived earlier, upstream, disrupting the science process. Observations of the P2P project reveal that not only can this characteristic significantly impact the way researchers conduct their science-making; it can influence the researcher’s perception of experimental design. Relative to traditional data collection approaches, the timing of the arrival of data, coupled with the data being inexpensive meant the researchers were relatively unconcerned about the data being incorrect or having problems. The perception was that if there was a problem with the data, they could be re-captured *en masse*, relatively quickly and cheaply. Thus, little emphasis was placed on experimental design. Instead, effort was moved toward the back end of the science process to data processing, transformation, and analyses.

Another important finding from this study is that much of the data undergo significant processing and transformation before arriving for analysis. Observations and interviews with the researchers unearthed two central implications of this processing requirement, namely the time it takes to perform the processing, and the influence that the process can have on accuracy and repeatability. While the need for these transformations have been described in previous chapters, it is worth noting that if data are to be made publicly available, both the raw and processed data should be supplied. Furthermore, the ability to reproduce results using the dataset for verification purposes will be greatly compromised if detailed methods are not provided.

The ethnographic account of the P2P researchers suggests that the field may progress by repurposing carefully crafted datasets instead of collecting new data each time a new hypothesis is posed. In this study I observed that researchers encountered data from two main groups: digital exhaust from a community of machines and sensors, and ground-reference data from research experiments. Importantly, data from these sources were structured very differently and as Huberty (2015) suggests, data from each source comes with its own biases and assumptions. An important learning from the P2P project is that some data characteristics influence the ability of the hyperspectral sensing and imaging dataset to be repurposed; a) accurate geolocation of data collected, including support data, b) time, c) easily accessible data d) adequately described methods, especially relating to atmospheric corrections and data processing.

(Geo)Location, (Geo)Location, (Geo)Location

Understanding where data come from, spectrally, spatially, and socially are all relevant to understanding how the data connect beyond the farm boundaries to a wider world. This study reveals that data from our main actor, the AisaFENIX, offer little benefit until they are supported by accompanying ground-reference hyperspectral data that was accurately geolocated. As an assemblage, the AisaFENIX's friend, the ASD, literally grounds the data. It is not surprising then that many of the difficulties encountered by the researchers in the P2P project centred around inaccuracies in spatial data. However, size does sometimes count (at least in remote sensing), and somewhat surprisingly, research conducted by the PGP's post-graduate students suggests that remote sensing data without accurate, detailed geolocation information can still be usable for some modelling purposes, especially when the pixel is consistently smaller than the target. The key message here is that while data may be useless to one researcher, they can be of great value to others looking through a different lens.

Time

A key characteristic of the contemporary data ecosystems is that unlike their predecessors, the new ecosystems are living and when curated with care, will grow into dynamic communities. Traditional data ecosystems have little life or value beyond the project from which they were borne. This was largely due to the challenges faced by agricultural researchers in reliably returning to the same position in a field more than once. Today, the advent of accurate geo-location technologies mean that it is possible to return to an exact location again and again. So, while the *date* of data collection was traditionally viewed as being important, observations of the PGP project highlight the importance of exact time metadata being attached to all data sources in the new data ecosystem, including proximal ground-reference data. For example, the study revealed that a study by one student involved in the PGP discovered that reflectance readings taken from grass plants could be correlated to the time of the day, thus revealing that recording time of day can be useful in identifying biochemical trends in plant growth.

Easily accessible data

Observations of the P2P project suggest that exhaust data from this new ecosystem can deliver additional, often unexpected, value for researchers in the PA context. While exhaust data can be of value (Kitchin, 2014), this study demonstrates the benefits of consciously creating a data ecosystem that supports surplus science. But what makes a good data ecosystem for iterative PA research and how do researchers approach science data management, curation, and access to facilitate surplus science? While comprehensive guidelines on data completeness and reusability have been produced for the biological sciences by (Roche et al., 2015), it is the seminal work of Medyckyj-Scott et al. (2016) that lays the foundation for the creation of successful data ecosystems in the New Zealand context. Medyckyj-Scott et al. (2016) assert that the fitness-for-purpose of data used by tools is critical and that "the requirements and interpretation of output for some tools will be specific to the New Zealand environment" (Medyckyj-Scott et al., 2016, p. 10).

The PGP project also demonstrates the need to understand what supporting data or enabling technologies are needed to create a versatile dataset for the context in which it is to be used. For example, In Chapter 7 we discovered that the absence of LiDAR data for the aerially acquired hyperspectral imaging data made it impossible for the researchers to apply a rough terrain model to the dataset for atmospheric correction and ortho-mosaicking purposes. This meant that only a flat terrain model could be applied, reducing the geospatial accuracy due to stretching of the pixels over highly variable terrain. Also, in Chapter 6 the P2P researchers found that the inclusion of differential GPS in the aircraft would have drastically reduced the need for data processing to correct the geometric alignment of the images. However, the researchers found a low-tech alternative solution, which involved placing tarpaulins to allow the researchers to manually adjust the imagery.

Scientific division of labour

To date, knowledge discovery and data mining have been driven by computer and data scientists. Importantly, the dramatic shift in the timing of the data and the data characteristics identified in this research signposts a significant change in the scientific division of labour in the agricultural sciences. So, the question is how will this impact on the scientific division of labour within precision agriculture research teams? As we learned earlier, the traditional epistemology sees the final outcome and end of scientific activity determined at the early stages of the science process. A key observation was that the scientific division of labour was adapted to meet the challenge of working in the new data economy by employing more data programming, computer science and statistical analytical skills in the team than initially planned.

Based on the observations of the study, I suggest that continuing to work in remote sensing or precision agriculture teams without input from other fields, such as the veterinary, soil and agronomic sciences may further widen the gap between information and knowledge and there is a risk of missing out on value generated from surplus science. Importantly, it is unlikely that all new precision agriculture technologies will have the same data characteristics; we need to understand how representative hyperspectral sensing and imaging is of emerging precision agriculture technologies and to taxonomise nascent technologies in terms of the new data economy.

Reflective practice in PA research

The findings contribute to our understanding of the forms that reflective practice may take when following science teams such as the P2P researchers; personal reflection, dialogic reflection, reflection within laboratories, reflections within communities of practice (seminars, conferences) and reflections within the academy (scientific and review articles) all contributed to the findings of this thesis. I encountered little evidence of reflective practice in funding, science-making, and communication practices both in the bibliometric study and the ethnographic account. In the case of the P2P study, while the researchers were struggling to adapt their practices, they were largely unaware of these changes until discussed with the social scientist involved with the project. This underscores the importance of the role reflective practice can play to support transformative change. I suggest that to capitalise on the promise of emerging big data technologies, the research community supports publishing practices that promote knowledge building, such as the reporting of detailed methods of successful and failed science approaches. This should include the academic work that is performed outside of publishing, for example allowing researchers to participate in workshops and conferences where IP issues might otherwise prevent them from participating and learning. Importantly, to consider how PA science is ‘done’ and to what end, policymakers and funders should consider including social scientists in the planning and execution of large-scale PA research projects.

Implications for policy makers and project funders

While some research has been performed with the intention to inform agricultural policies (Jouanjean, 2019; Kritikos, 2017; Pesce et al., 2019), there has been little discourse on the interface between policy makers and science-making practices within the laboratory. In Chapters 2 and 3 I found there is sparse social science scholarship in the field, save for a few important publications in the 1990s such as that by Wolf and Wood (1997) and Nowak (1997, 1998). This thesis set out to understand how emerging technological assemblages in contemporary PA are very different to the production-enhancing technologies that preceded it in the evolution of our agriculture system. Yet, the ethnography revealed that policy makers, funders, and researchers themselves continue to apply the same linear, largely unsuccessful thinking to these new data ecosystems without considering the needs of the end-users or how the science-making is actually practiced.

Analysis of the ethnographic account revealed that researchers are adapting their practices and scientific division of labour to capitalise on emerging data ecosystems. Specifically, by employing an ANT approach, I explored both human and non-human actants, which allowed me to document a significant repositioning of scientific expertise and its division of labour in response to the characteristics of the data generated by the project. This is important, because the policy makers and funders involved with the project appeared to be largely unaware of the profound changes the scientists were making to accommodate the new data characteristics identified in Chapter 6. Indeed, it appears that in the broader context, funders and policymakers continue to structure, fund, and manage data-rich research projects as though nothing has changed. This critical disjunct between the structure and management of research projects and the reality of research on the ground is a key finding of the ethnographic study. Observations of the P2P project revealed that by failing to acknowledge the changing characteristics and role of data in PA research, funders and policymakers may unwittingly set up researchers to fail by structuring data-rich research projects to align with the processes more suited to small data studies.

In the case of the P2P project, the researchers had to retrofit their methods to comply with misaligned project design, structure, and monitoring. More research extension is unlikely to be the key mechanism to bridge the gap between the expected world and the real world here. Instead, research effort needs to be directed into identifying the appropriate distribution of scientific division of labour at a policy and project scale.

Project set up for proliferation

History conditions policy makers, funders and researchers to frame analysis as a process where data are assessed to provide an answer or solution to an immediate problem (Alrøe & Noe, 2014). The classical science model has traditionally provided a reliable and efficient means of finding such solutions by reducing problems into ‘bite-sized chunks’. However, a downside of this approach is that it falls short when addressing complex and ‘wicked’ problems. The linear process fails to reference the dependence of the analysis on other data sources, and how they are integrated into the science-making process when addressing complex problems. What appears to be data exploration is rather a mission to discover relationships for predetermined outcomes. The analytical lens is often selected by funders, not the researchers. The data are used to support or dispute the initial evidence which gave birth to the hypothesis being tested. When the analytical lens is tightly controlled, it leaves little opportunity for proliferation or any departure from the present approach. Moreover, using the data to inform other decisions or to solve other problems is viewed negatively, framed as project creep.

Researchers operating in the post-scarcity data economy do not need to go back and restart a blind experiment although they may tweak or change processes to test an idea or hypothesis. So, how can PA researchers exploit the iterative thinking and enquiring spirit of the bricoleur in this new data economy?

Policy makers, funders, researchers, and farmers all have an important role in supporting value creation in the new data economy. Evidence from the P2P study demonstrates how new questions can keep being asked of a large and versatile dataset. An example of this was when Virginia W collaborated with the veterinarian to repurpose the P2P dataset if there was a relationship between copper and molybdenum from the data collected on pasture characteristics. The bricolage-like approach suggested by Yule and Wood (2014) demonstrates that the repurposing of data can be routinely treated as a source of value, not as a risk to be avoided or minimised.

However, progressing the field will require the science world to reconsider negative connotations associated with the iterative approach. Scope creep is viewed as an aberration to be corrected by keeping a project on track rather than being perceived and managed as opportunities for proliferation. While researchers may identify opportunities to refine value from datasets outside of the scope of a project, exploring such opportunities in the contemporary research environment is almost universally perceived negatively, with researchers needing to ‘get back on track’. To prevent researchers, funders and farmers missing out on these opportunities, researchers such as Harrin (2007) suggest that the scoping of projects can be tailored to accommodate new iterative approaches. However, there is little evidence of this approach being applied in the PA research arena. Hence, a question arising from this work is how open are funders to this flexibility, especially in applied / near industry research?

To promote surplus science, funders and researchers could consider internalising scope creep within a project to seize the opportunity and realise value. This study demonstrates it is possible to support iterative science in precision agriculture, so perhaps researchers and funders need to adopt a less negative notion for scope creep to realise the value of hyperspectral data and other big-data technologies. For example, the P2P Director was cognisant of the opportunity to build additional knowledge from the P2P dataset and the project funder supported this by funding doctoral projects that would dip into the data. Given the group’s inexperience in the domain, there were some interesting successes, which hold promise for future projects. The students created and tested new hypotheses with varying degrees of success, using the existing dataset to inform their simulation models; these included models for tree identification and health, models to discriminate between the small native trees manuka and kanuka, and models to identify pasture area. While none of these models progressed to commercial applications, they did contribute to the academy where permitted, and the students eventually achieved their academic goals.

While not the main focus of research, this study also reveals some socio-ethical risks relating to emerging PA technologies that should be considered at a project level. Three narrow but important questions identified in the thesis that warrant further investigation are:

1) What policy responses can be put in place to counteract potential negative effects of PA technology, such as the centralisation of knowledge, concentration of power and centralised agricultural supply chains?
2) What policy discourses around PA emerge, and how can they be used to identify and manage negative socio-ethical consequences of PA research?
3) How is the notion of financialisation of agricultural data beyond the farm boundaries impacting farmer livelihoods?

Implications for science-makers and future research directions

The research conducted in this thesis has led to some useful results and conclusions. These findings include the identification of gaps in our knowledge around science-making in PA research that would benefit from further research. The philosophical and methodological tenets needed to provide insight into what an evolution of the data economy might mean for second-order scientific enquiry in PA are still to be crafted. Observations suggest the data from single standalone technologies, such as hyperspectral sensors are unlikely to solve second-order problems faced by farmers and wider society. The research does however shed light on the opportunities that these emerging technologies provide for generating data that could contribute to the science-making on a regional or even global scale, specifically:

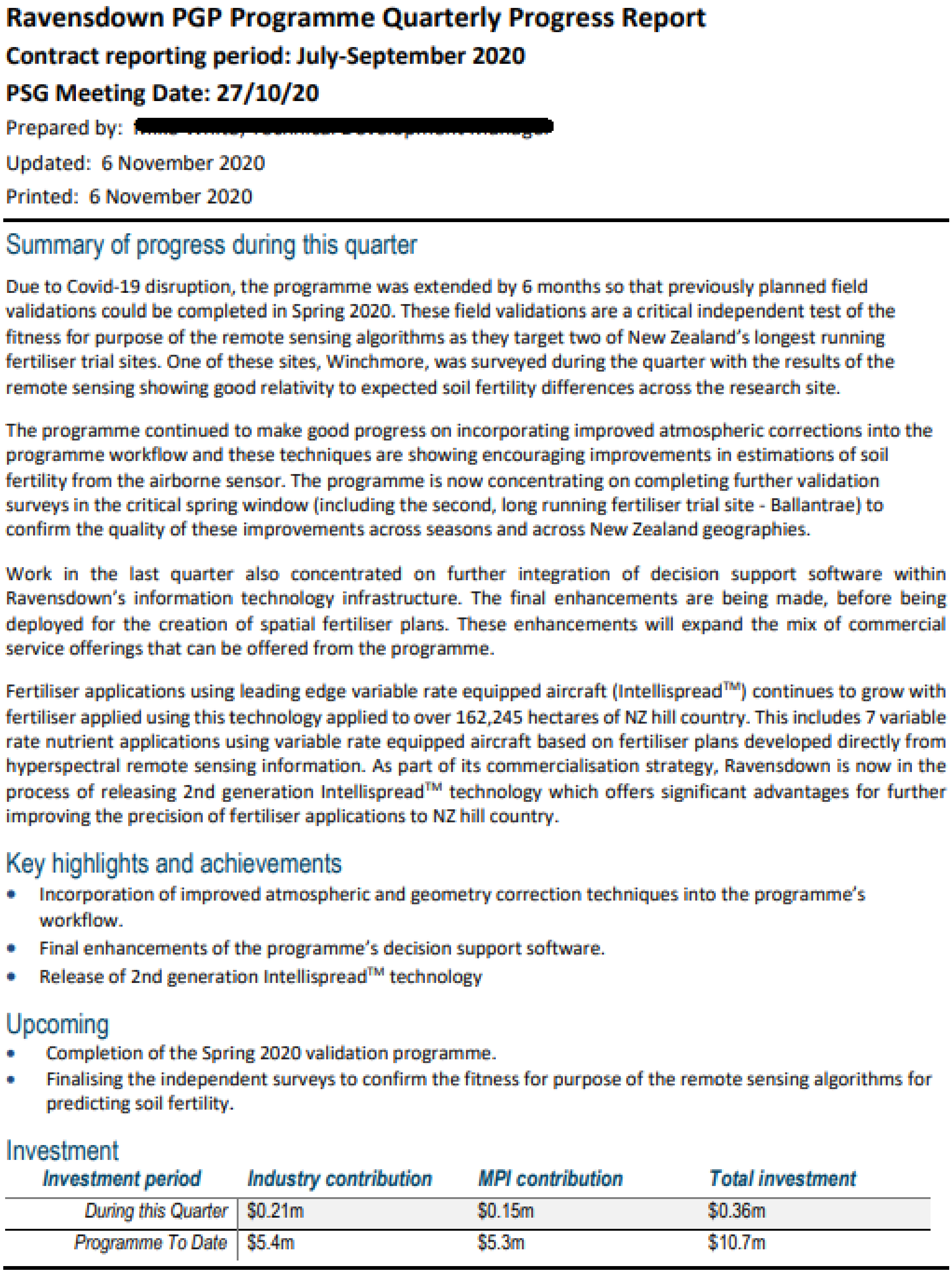
1) Examining how to identify data structures and research practices that best support the assemblage and aggregation of contemporary PA data, including small and big data.
2) Further investigation to identify data structures that support the assemblage and aggregation of data and to foster the proliferation of surplus science is needed.

Finally, the field will also benefit from increased critical reflection to build a greater understanding of the structure of scientific enquiry and further test the findings of this study. In particular, more methodological work is needed on how to promote positive community reflection and collegial discussion practices in the contemporary research laboratory context.

Appendices

Appendix A: Example Ravensdown PGP programme quarterly progress report

(Source withheld)



Ravensdown PGP Programme Quarterly Progress Report

Contract reporting period: July-September 2020

PSG Meeting Date: 27/10/20

Prepared by: [redacted]
Updated: 6 November 2020
Printed: 6 November 2020

Summary of progress during this quarter

Due to Covid-19 disruption, the programme was extended by 6 months so that previously planned field validations could be completed in Spring 2020. These field validations are a critical independent test of the fitness for purpose of the remote sensing algorithms as they target two of New Zealand's longest running fertiliser trial sites. One of these sites, Winchmore, was surveyed during the quarter with the results of the remote sensing showing good relativity to expected soil fertility differences across the research site.

The programme continued to make good progress on incorporating improved atmospheric corrections into the programme workflow and these techniques are showing encouraging improvements in estimations of soil fertility from the airborne sensor. The programme is now concentrating on completing further validation surveys in the critical spring window (including the second, long running fertiliser trial site - Ballantrae) to confirm the quality of these improvements across seasons and across New Zealand geographies.

Work in the last quarter also concentrated on further integration of decision support software within Ravensdown's information technology infrastructure. The final enhancements are being made, before being deployed for the creation of spatial fertiliser plans. These enhancements will expand the mix of commercial service offerings that can be offered from the programme.

Fertiliser applications using leading edge variable rate equipped aircraft (Intellispread™) continues to grow with fertiliser applied using this technology applied to over 162,245 hectares of NZ hill country. This includes 7 variable rate nutrient applications using variable rate equipped aircraft based on fertiliser plans developed directly from hyperspectral remote sensing information. As part of its commercialisation strategy, Ravensdown is now in the process of releasing 2nd generation Intellispread™ technology which offers significant advantages for further improving the precision of fertiliser applications to NZ hill country.

Key highlights and achievements

- Incorporation of improved atmospheric and geometry correction techniques into the programme's workflow.
- Final enhancements of the programme's decision support software.
- Release of 2nd generation Intellispread™ technology

Upcoming

- Completion of the Spring 2020 validation programme.
- Finalising the independent surveys to confirm the fitness for purpose of the remote sensing algorithms for predicting soil fertility.

Investment

Investment period	*Industry contribution*	*MPI contribution*	*Total investment*
During this Quarter	$0.21m	$0.15m	$0.36m
Programme To Date	$5.4m	$5.3m	$10.7m

Appendix B: Example contact questions –Phase 1 – understanding the tech

Introduction: Job title and a brief description

Qualifications/ expertise

Where from before P2P

How and when did you join the P2P?

Sup. What motivated you to join the P2P?

Understanding the technology that P2P has used, is currently using and may use in the future.

What do you do for the P2P?

What are the key tasks or projects you're working on at the moment?

What do you measure (tools instruments)?

What do you do with the measurements (analysis, objectives)?

Sup. Scale (resolution/ up and down).

Sup. Calibration = what is being calibrated/ validated?

Also, measurement modelling

Job in relation to others in the P2P team

Who do you measure with?

Who do you give measurements to, receive from, share data?

Why? How? What software etc.

What are some of the barriers you foresee to achieving P2P's research goals (technical limitations/ financial limitations/ organisational limitation – more staff needed/ upskilling the team etc.)

Precision

What is precision?

How is what you do precise?

Why is it good to be precise?

Compared to what?

What about error (how is it corrected/ minimised/ acknowledged etc.)

How are innovations developed/ changes to tech made?

Opportunities ahead

What could be done to improve innovation within the P2P team?

Anything that you would like to add that we haven't discussed?

Appendix C: Ethics approval

MASSEY UNIVERSITY
ALBANY

11 May 2015

Megan Cushnahan
[redacted]

Dear Megan

Re: Optimising scientific design for the on- farm uptake of precision agriculture

Thank you for your Low Risk Notification which was received on 8 May 2015.

Your project has been recorded on the Low Risk Database which is reported in the Annual Report of the Massey University Human Ethics Committees.

You are reminded that staff researchers and supervisors are fully responsible for ensuring that the information in the low risk notification has met the requirements and guidelines for submission of a low risk notification.

The low risk notification for this project is valid for a maximum of three years.

Please notify me if situations subsequently occur which cause you to reconsider your initial ethical analysis that it is safe to proceed without approval by one of the University's Human Ethics Committees.

Please note that travel undertaken by students must be approved by the supervisor and the relevant Pro Vice-Chancellor and be in accordance with the Policy and Procedures for Course-Related Student Travel Overseas. In addition, the supervisor must advise the University's Insurance Officer.

A reminder to include the following statement on all public documents:

> *"This project has been evaluated by peer review and judged to be low risk. Consequently, it has not been reviewed by one of the University's Human Ethics Committees. The researcher(s) named above are responsible for the ethical conduct of this research.*
>
> *If you have any concerns about the conduct of this research that you wish to raise with someone other than the researcher(s), please contact Dr Brian Finch, Director (Research Ethics), telephone 06 356 9099, extn 86015, e-mail humanethics@massey.ac.nz".*

Please note that if a sponsoring organisation, funding authority or a journal in which you wish to publish requires evidence of committee approval (with an approval number), you will have to provide a full application to one of the University's Human Ethics Committees. You should also note that such an approval can only be provided prior to the commencement of the research.

Yours sincerely

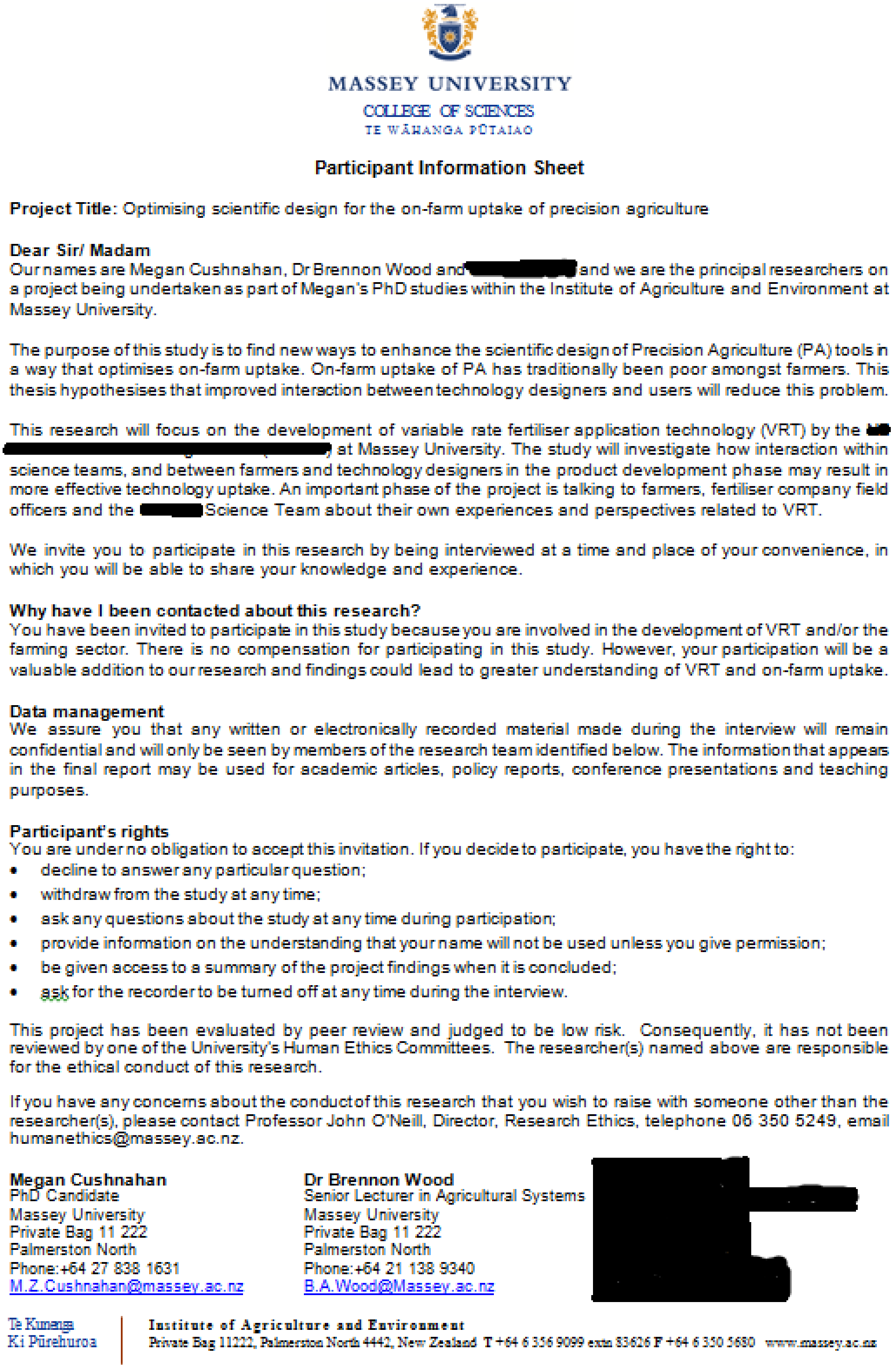
Brian T Finch (Dr)
Chair, Human Ethics Chairs' Committee and
Director (Research Ethics)

cc [redacted], Dr Brennon Wood
Institute of Agriculture & Environment
Palmerston North

Professor Peter Kemp
Head of Institute of Agriculture & Environment
Palmerston North

Massey University Human Ethics Committee
Accredited by the Health Research Council

Appendix D: Participant information sheet



MASSEY UNIVERSITY
COLLEGE OF SCIENCES
TE WĀHANGA PŪTAIAO

Participant Information Sheet

Project Title: Optimising scientific design for the on-farm uptake of precision agriculture

Dear Sir/ Madam
Our names are Megan Cushnahan, Dr Brennon Wood and [redacted] and we are the principal researchers on a project being undertaken as part of Megan's PhD studies within the Institute of Agriculture and Environment at Massey University.

The purpose of this study is to find new ways to enhance the scientific design of Precision Agriculture (PA) tools in a way that optimises on-farm uptake. On-farm uptake of PA has traditionally been poor amongst farmers. This thesis hypothesises that improved interaction between technology designers and users will reduce this problem.

This research will focus on the development of variable rate fertiliser application technology (VRT) by the [redacted] at Massey University. The study will investigate how interaction within science teams, and between farmers and technology designers in the product development phase may result in more effective technology uptake. An important phase of the project is talking to farmers, fertiliser company field officers and the [redacted] Science Team about their own experiences and perspectives related to VRT.

We invite you to participate in this research by being interviewed at a time and place of your convenience, in which you will be able to share your knowledge and experience.

Why have I been contacted about this research?
You have been invited to participate in this study because you are involved in the development of VRT and/or the farming sector. There is no compensation for participating in this study. However, your participation will be a valuable addition to our research and findings could lead to greater understanding of VRT and on-farm uptake.

Data management
We assure you that any written or electronically recorded material made during the interview will remain confidential and will only be seen by members of the research team identified below. The information that appears in the final report may be used for academic articles, policy reports, conference presentations and teaching purposes.

Participant's rights
You are under no obligation to accept this invitation. If you decide to participate, you have the right to:

- decline to answer any particular question;
- withdraw from the study at any time;
- ask any questions about the study at any time during participation;
- provide information on the understanding that your name will not be used unless you give permission;
- be given access to a summary of the project findings when it is concluded;
- ask for the recorder to be turned off at any time during the interview.

This project has been evaluated by peer review and judged to be low risk. Consequently, it has not been reviewed by one of the University's Human Ethics Committees. The researcher(s) named above are responsible for the ethical conduct of this research.

If you have any concerns about the conduct of this research that you wish to raise with someone other than the researcher(s), please contact Professor John O'Neill, Director, Research Ethics, telephone 06 350 5249, email humanethics@massey.ac.nz.

Megan Cushnahan
PhD Candidate
Massey University
Private Bag 11 222
Palmerston North
Phone: +64 27 838 1631
M.Z.Cushnahan@massey.ac.nz

Dr Brennon Wood
Senior Lecturer in Agricultural Systems
Massey University
Private Bag 11 222
Palmerston North
Phone: +64 21 138 9340
B.A.Wood@Massey.ac.nz

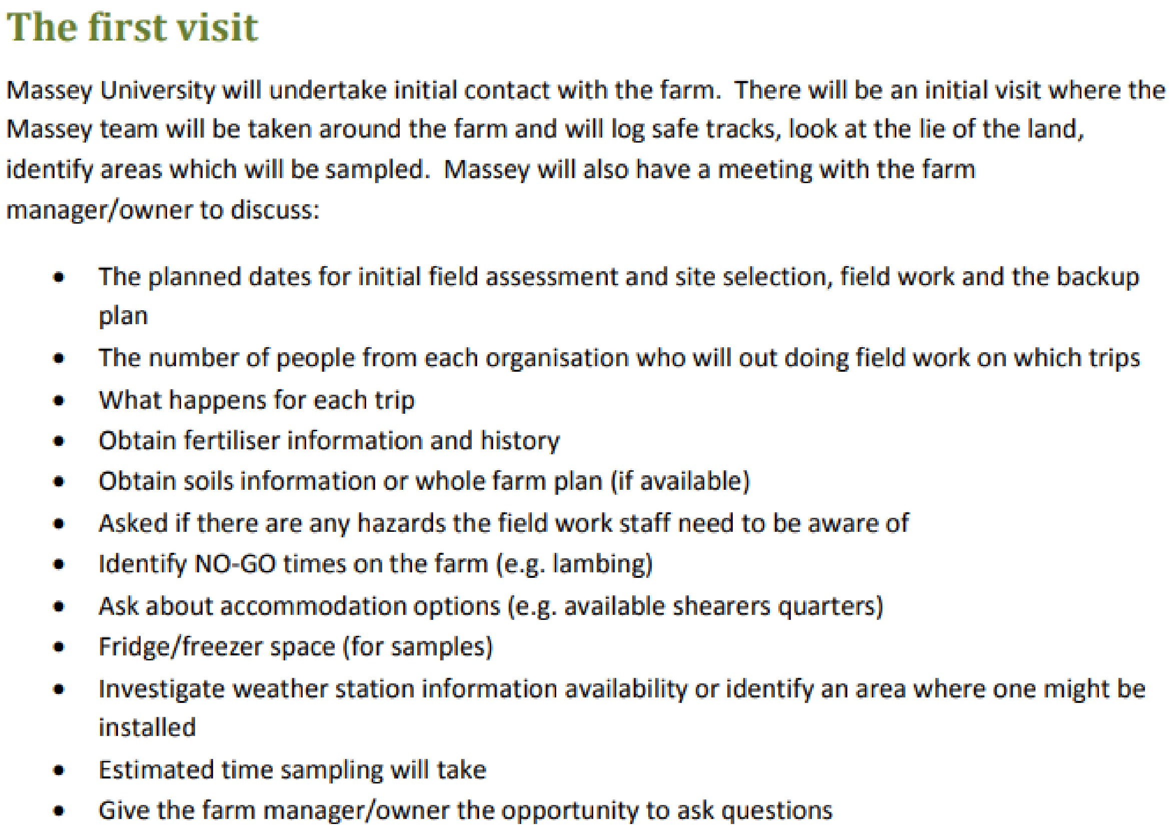
[redacted]

Te Kunenga
Ki Pūrehuroa

Institute of Agriculture and Environment
Private Bag 11222, Palmerston North 4442, New Zealand T +64 6 356 9099 extn 83626 F +64 6 350 5680 www.massey.ac.nz

Appendix E: Excerpt from the Pioneering to Precision project protocol document.

The following artefact is an excerpt from the P2P project protocol document, which explains what the P2P researchers would discuss with the farm manager/ owner at their initial visit.



The first visit

Massey University will undertake initial contact with the farm. There will be an initial visit where the Massey team will be taken around the farm and will log safe tracks, look at the lie of the land, identify areas which will be sampled. Massey will also have a meeting with the farm manager/owner to discuss:

- The planned dates for initial field assessment and site selection, field work and the backup plan
- The number of people from each organisation who will out doing field work on which trips
- What happens for each trip
- Obtain fertiliser information and history
- Obtain soils information or whole farm plan (if available)
- Asked if there are any hazards the field work staff need to be aware of
- Identify NO-GO times on the farm (e.g. lambing)
- Ask about accommodation options (e.g. available shearers quarters)
- Fridge/freezer space (for samples)
- Investigate weather station information availability or identify an area where one might be installed
- Estimated time sampling will take
- Give the farm manager/owner the opportunity to ask questions

References

Adams, R. J., Smart, P., & Huff, A. S. (2017). Shades of Grey: Guidelines for Working with the Grey Literature in Systematic Reviews for Management and Organizational Studies. *International Journal of Management Reviews, 19*(4), 432-454. doi:10.1111/ijmr.12102

Adão, T., Hruška, J., Pádua, L., Bessa, J., Peres, E., Morais, R., & Sousa, J. J. (2017). Hyperspectral imaging: A review on UAV-based sensors, data processing and applications for agriculture and forestry. *Remote Sensing, 9*(11), 1110.

Adrados, C., Girard, I., Gendner, J.-P., & Janeau, G. (2002). Global positioning system (GPS) location accuracy improvement due to selective availability removal. *Comptes Rendus Biologies, 325*(2), 165-170.

Adrian, A. M., Norwood, S. H., & Mask, P. L. (2005). Producers' perceptions and attitudes toward precision agriculture technologies. *Computers and Electronics in Agriculture, 48*(3), 256-271.

AgGateway. (2014). Data privacy and use white paper [White paper]. Retrieved from http://www.aggateway.org/eConnectivityActivities/Committees/DataPrivacySecurity.aspx

Agrawal, A. (1995). Dismantling the divide between indigenous and scientific knowledge. *Development and change, 26*(3), 413-439.

AgriOptics New Zealand Ltd. (2019). EM Surveying. Retrieved from https://www.agrioptics.co.nz/portfolio/em-survey/

Aken, J. E. v. (2004). Management research based on the paradigm of the design sciences: the quest for field-tested and grounded technological rules. *Journal of management studies, 41*(2), 219-246.

Akpo, E., Crane, T. A., Vissoh, P. V., & Tossou, R. C. (2014). Co-production of Knowledge in Multi-stakeholder Processes: Analyzing Joint Experimentation as Social Learning. *The Journal of Agricultural Education and Extension*(ahead-of-print), 1-20.

Allen, R. C. (1991). The two English agricultural revolutions 1450 - 1850. In B. M. Campbell & M. Overton (Eds.), *Land, Labour and Livestock: Hisotrical Studies in European Agricultural Productivity*. Manchester: Distributed exclusively in the USA and Canada by St. Martin's Press.

Allen, R. C. (1992). Enclosure and the yeoman: the agricultural development of the South Midlands 1450-1850. *OUP Catalogue*.

Alrøe, H. F., & Noe, E. (2014). Second-order science of interdisciplinary research: A polyocular framework for wicked problems.

Andersen, H., & Hepburn, B. (2016). Scientific method. In E. N. Zalta (Ed.), *The Stanford Encyclopedia of Philosophy (Summer 2016 Edition)*. URL = <https://plato.stanford.edu/archives/sum2016/entries/scientific-method/>.

Anderson, G., & Yang, C. (1996). *Multispectral videography and geographic information systems for site-specific farm management.* Paper presented at the Proceedings of the Third International Conference on Precision Agriculture.

Anderson, N., De Dreu, C. K., & Nijstad, B. A. (2004). The routinization of innovation research: A constructively critical review of the state-of-the-science. *Journal of organizational Behavior, 25*(2), 147-173.

Apiculture New Zealand. (2016). About Apiculture New Zealand. Retrieved from https://apinz.org.nz/about/

ASD Inc. (2016). FieldSpec 4 Hi-Res NG Spectroradiometer. Retrieved from http://www.asdi.com/products/fieldspec/fieldspec-4-hi-res-ng-spectroradiometer

Ausseil, A., Dymond, J., Dynes, R., Shepherd, J., DeVantier, B., & Sutherland, A. (2011). *Estimating pasture quality using Landsat ETM+: application for the greenhouse gas inventory of New Zealand.* Paper presented at the International Symposium on Remote Sensing for the Environment, Sydney, Australia.

Back, S.-W., Yu, S.-H., Kim, Y.-J., Chung, S.-O., & Lee, K.-H. (2014). An Image-Based Application Rate Measurement System for a Granular Fertilizer Applicator. *Transactions of the ASABE, 57*(2), 679-687.

Bagheri, N. (2017). Development of a high-resolution aerial remote-sensing system for precision agriculture. *International journal of remote sensing, 38*(8-10), 2053-2065.

Balafoutis, A., Beck, B., Fountas, S., Vangeyte, J., Wal, T. V. d., Soto, I., Eory, V. (2017). Precision agriculture technologies positively contributing to GHG emissions mitigation, farm productivity and economics. *Sustainability, 9*(8), 1339.

Ballance Agri-Nutrients Ltd. (2019). https://ballance.co.nz/mapping. Retrieved from https://ballance.co.nz/mapping

Barham, B. L., Foltz, J. D., Jackson-Smith, D., & Moon, S. (2004). The dynamics of agricultural biotechnology adoption: Lessons from series rBST use in Wisconsin, 1994–2001. *American Journal of Agricultural Economics, 86*(1), 61-72.

Barrowclough, P. (2013, February 2013). *Precision Agriculture Association of New Zealand.* Paper presented at the 26th Annual FLRC Workshop, Massey University, Palmerston North, New Zealand.

Bates, J., Lin, Y.-W., & Goodale, P. (2016). Data journeys: Capturing the socio-material constitution of data objects and flows. *Big Data & Society, 3*(2), 2053951716654502.

Battell, D. (2018). *Review of the Ministry for Primary Industries' Primary Growth Partnership* New Zealand: Ministry for Primary Industries Retrieved from https://www.mpi.govt.nz/dmsdocument/30402/loggedIn

Bauman, Z., Beck, U., Beck-Gernsheim, E., Benhabib, S., Burgess, R., Chamberlain, M., Wengraf, T. (2002). Qualitative interviewing: Asking, listening and interpreting. *Qualitative Research in Action. 1st ed. London: SAGE Publications*, 226-241.

Becker, H., Berger, P., Luckmann, T., Burawoy, M., Gans, H., Gerson, K., Inciardi, J. (2002). Observation and interviewing: Options and choices in qualitative research. *Qualitative research in action, 6*, 200-224.

Beef + Lamb New Zealand. (2018). *Compendium of Farm Facts 2018* (Vol. 42).

Beef + Lamb New Zealand. (2020a). Hill country sheep and beef farms. Retrieved from https://beeflambnz.com/knowledge-hub/PDF/hill-country-sheep-and-beef-farms.pdf

Beef + Lamb New Zealand. (2020b). *New season outlook 2020-21*. Retrieved from New Zealand: https://beeflambnz.com/sites/default/files/data/files/B%2BLNZ%20New%20Season%20Outlook%202020-21.pdf

Bell, G., Hey, T., & Szalay, A. (2009). Beyond the data deluge. *Science, 323*(5919), 1297-1298.

Bennett, J. M. (2015). Agricultural big data: utilisation to discover the unknown and instigate practice change. *Farm Policy Journal, 12*(1), 43-50.

Berreman, G. D. (2012). Behind many masks: ethnography and impression management. *Ethnographic fieldwork: an anthropological reader*, 152-174.

Bewley, J. M., & Russell, R. A. (2010). *Reasons for slow adoption rates of precision dairy farming technologies: evidence from a producer survey.* Paper presented at the Proceedings of the First North American Conference on Precision Dairy Management.

Blommaert, J., & Jie, D. (2010). *Ethnographic fieldwork: A beginner's guide*: Multilingual Matters.

Bollen, F., Praat, J.-P., & Yule, I. J. (2001). *Precision horticulture: Progress, opportunities and requirements in the supply chain.* Paper presented at the Proceedings of the First Australian Geospatial Information and Agriculture Conference.

Box, J. F. (1980). RA Fisher, the Life of a Scientist.

Bradbury, S., Mackenzie, C., Mackenzie, J., & Hedley, C. B. (2013). *Precision irrigation as a tool to reduce nutrient leaching and runofff.* Paper presented at the Fertilizer and Lime Research Centre, Massey University, Palmerston North, New Zealand. http://flrc.massey.ac.nz/publications.html.

Bramley, R. G. V. (2009). Lessons from nearly 20 years of Precision Agriculture research, development, and adoption as a guide to its appropriate application. *Crop and Pasture Science, 60*(3), 197-217.

Bramley, R. G. V., & Trengove, S. (2013). Precision agriculture in Australia: present status and recent developments. *Engenharia Agrícola, 33*(3), 575-588.

Brewer, G., Gajendran, T., & Hilaire, T. (2013). ICT-mediated project team collaboration and issues of trust: an Actor Network Theory perspective.

Brown, R. M., Dillon, C. R., Schieffer, J., & Shockley, J. M. (2015). The carbon footprint and economic impact of precision agriculture technology on a corn and soybean farm. *Journal of Environmental Economics and Policy*, 1-14.

Burger, B. (1996). ADAR Digital Aerial Photography Applications In Precision Farming. *Precision agriculture*(precisionagricu3), 905-905.

Callaghan Innovation. (2016). New Zealand AgTech: Transforming Farming. In.

Callon, M. (1986). The sociology of an actor-network: The case of the electric vehicle. In *Mapping the dynamics of science and technology* (pp. 19-34): Springer.

Cameron, D. (2016). Sustaining the productivity of New Zealand's hill country–A land manager’s view. *Hill Country–Grassland Research and Practice Series, 16*, 151-156.

Campbell, B. M., & Overton, M. (1993). A new perspective on medieval and early modern agriculture: six centuries of Norfolk farming c. 1250-c. 1850. *Past & Present, 141*(1), 38-105.

Campbell, J. B., & Wynne, R. H. (2011). *Introduction to remote sensing*: Guilford Press.

Carbonell, I. M. (2016). The ethics of big data in big agriculture. *Internet Policy Review, 5*(1), 13.

Carletto, C., de Janvry, A., & Sadoulet, E. (1999). Sustainability in the Diffusion of Innovations: Smallholder Nontraditional Agro-Exports in Guatemala. *Economic Development and Cultural Change, 47*(2), 345-369.

Carolan, M. S. (2017). Publicising Food: Big Data, Precision Agriculture, and Co-Experimental Techniques of Addition. *Sociologia Ruralis, 57*(2), 135-154.

Carolan, M. S. (2020). Automated agrifood futures: robotics, labor and the distributive politics of digital agriculture. *The Journal of Peasant Studies, 47*(1), 184-207.

Carpi, A., & Egger, A. E. (2011). The Nature of Scientific Knowledge. *Visionlearning*(2).

Cayley, D. (2009). Ideas on the Nature of Science.

Chalmers, A. F. (2013). *What is this thing called science?* : Hackett Publishing.

Chang, C.-I. (2013). *Hyperspectral data processing: algorithm design and analysis*: John Wiley & Sons.

Charmaz, K., & Mitchell, R. G. (2001). Grounded theory in ethnography. *Handbook of ethnography, 160*, 174.

Chen, W., Bell, R., Brennan, R., Bowden, J., Dobermann, A., Rengel, Z., & Porter, W. (2009). Key crop nutrient management issues in the Western Australia grains industry: a review. *Soil Research, 47*(1), 1-18.

Chi, M., Plaza, A., Benediktsson, J. A., Sun, Z., Shen, J., & Zhu, Y. (2016). Big data for remote sensing: Challenges and opportunities. *Proceedings of the IEEE, 104*(11), 2207-2219.

Chibnik, M. (1981). Small Farmer Risk Aversion: Peasant Reality or Policymakers' Rationalization? *Culture & Agriculture, 1*(10), 1-7.

Chok, S., Grafton, M. C., Yule, I. J., & White, M. (2016). *Accuracy of Differential Rate Application Technology for Aerial Spreading of Granular Fertilizer within New Zealand.* Paper presented at the Proceedings of the 13th International Conference on Precision Agriculture, St. Louis, MO, USA.

Christensen, L. B., Johnson, B., Turner, L. A., & Christensen, L. B. (2011). Research methods, design, and analysis.

Christiansen, M. P., Laursen, M. S., Mikkelsen, B. F., Teimouri, N., Jørgensen, R. N., & Sørensen, C. A. G. (2018). Current potentials and challenges using Sentinel-1 for broadacre field remote sensing. *arXiv preprint arXiv:1809.01652*.

Clarke, D., Litherland, A., Mata, G., & Burling-Claridge, R. (2006). *Pasture monitoring from space.* Paper presented at the Proceedings of the South Island dairy event (SIDE) conference.

Clunies-Ross, T., & Hildyard, N. (2013). *The politics of industrial agriculture*: Routledge.

Collier, J., & Collier, M. (1986). *Visual anthropology: Photography as a research method*: UNM Press.

Cook, S., Adams, M., & Bramley, R. G. V. (2000). *What is obstructing the wider adoption of precision agriculture technology.* Paper presented at the Proceedings of the 5th International Conference on Precision Agriculture, Bloomington, Minnesota, USA, 16-19 July, 2000.

Cook, S., & Bramley, R. (1998). Precision agriculture—opportunities, benefits and pitfalls of site-specific crop management in Australia. *Australian Journal of Experimental Agriculture, 38*(7), 753-763.

Corbin, J., Strauss, A., & Strauss, A. L. (2015). *Basics of qualitative research*: Sage.

Corden, A., & Sainsbury, R. (2006). Exploring 'quality': Research participants' perspectives on verbatim quotations. *International journal of social research methodology, 9*(2), 97-110.

Corwin, D. L. (2013). Site-specific management and delineating management zones. In M. A. Oliver (Ed.), *Precision Agriculture for Sustainability and Environmental Protection* (pp. 135-157).

Cracknell, A. P. (2018). The development of remote sensing in the last 40 years. In: Taylor & Francis.

Craighead, M., & Yule, I. J. (2001). *Opportunities for increased profitability from precision agriculture.* Paper presented at the First Australian Geospatial Information and Agriculture Conference, Sydney, Australia.

Crane, T. A. (2014). Bringing science and technology studies into agricultural anthropology: Technology development as cultural encounter between farmers and researchers. *Culture, Agriculture, Food and Environment, 36*(1), 45-55.

Creswell, J. W., & Clark, V. L. P. (2017). *Designing and conducting mixed methods research*: Sage publications.

Cropanzano, R. (2009). Writing nonempirical articles for Journal of Management: General thoughts and suggestions. In: SAGE Publications Sage CA: Los Angeles, CA.

Cushnahan, M. Z., Wood, B. A., & Yule, I. J. (2017). *Is big data driving a paradigm shift in precision agriculture?* Paper presented at the The International Tri-Conference for Precision Agriculture in 2017, Hamilton, New Zealand.

Cushnahan, M. Z., Wood, B. A., Yule, I. J., & Wilson, R. (2016). *Discovering the value of hyperspectral data for New Zealand Agriculture.* Paper presented at the Integrated Nutrient and Water Management for Sustainable Farming,

Massey University, Palmerston North.

Cushnahan, T. A., Yule, I. J., Grafton, M. C., Pullanagari, R. R., & White, M. (2017). The classification of hill country vegetation from hyperspectral imagery.

Daberkow, S. G., & McBride, W. D. (2003). Farm and operator characteristics affecting the awareness and adoption of precision agriculture technologies in the US. *Precision agriculture, 4*(2), 163-177.

Daberkow, S. G., McBride, W. D., Robert, P., Rust, R., & Larson, W. (2000). *Adoption of precision agriculture technologies by US farmers.* Paper presented at the Proceedings of the 5th International Conference on Precision Agriculture, Bloomington, Minnesota, USA, 16-19 July, 2000.

De Laet, M., & Mol, A. (2000). The Zimbabwe bush pump: Mechanics of a fluid technology. *Social studies of science, 30*(2), 225-263.

de Leon, F. L. L., & McQuillin, B. (2020). The role of conferences on the pathway to academic impact evidence from a natural experiment. *Journal of Human Resources, 55*(1), 164-193.

Demirbaş, N. (2018). Precision Agriculture in Terms of Food Security: Needs for The Future. *Precis. Agric, 27*.

Desai, A., Zoccatelli, G., Adams, M., Allen, D., Brearley, S., Rafferty, A. M., Donetto, S. (2017). Taking data seriously: the value of actor-network theory in rethinking patient experience data. *Journal of health services research & policy, 22*(2), 134-136.

DeWalt, K. M., & DeWalt, B. R. (1998). Participant Observation. In H. R. Bernard & C. C. Gravlee (Eds.), *Handbook of methods in cultural anthropology* (pp. 259-299): Rowman & Littlefield.

Dilley, R. S. (1991). *Agricultural Change and Common Land in Cumberland 1700-1850.*

Dixon, J., & McCann, M. (1997). Precision Agriculture in the 21st century. *Geospatial and information technologies in crop management. Washington DC*.

Dodd, M., McDowell, R., & Quinn, J. (2016). A review of contaminant losses to water from pastoral hill lands and mitigation options. *Hill Country–Grassland Research and Practice Series, 16*, 137 - 148.

DOMO. (2019). *Data never sleeps 6.0*. Retrieved from https://www.domo.com/solution/data-never-sleeps-6.

Douthwaite, B., Keatinge, J., & Park, J. (2001). Why promising technologies fail: the neglected role of user innovation during adoption. *Research policy, 30*(5), 819-836.

Drew, C. J., Hardman, M. L., & Hosp, J. L. (2007). Ethical issues in conducting research. In.

Eastwood, C. R. (2008). Innovative precision dairy systems: a case study of farmer learning and technology co-development.

Eastwood, C. R., Chapman, D., & Paine, M. (2012). Networks of practice for co-construction of agricultural decision support systems: Case studies of precision dairy farms in Australia. *Agricultural Systems, 108*, 10-18.

Eastwood, C. R., Jago, J., Edwards, J., & Burke, J.-L. (2015). Getting the most out of advanced farm management technologies: Roles of technology suppliers and dairy industry organisations in supporting precision dairy farmers. *Animal Production Science*.

Eastwood, C. R., Klerkx, L., Ayre, M., & Rue, B. D. (2019). Managing socio-ethical challenges in the development of smart farming: from a fragmented to a comprehensive approach for responsible research and innovation. *Journal of Agricultural and Environmental Ethics, 32*(5-6), 741-768.

Edwards, P., Fleming, A., Lacey, J., Lester, L., Pinkard, E., Ruckstuhl, K., Williams, T. (2019). Trust, engagement, information and social licence—insights from New Zealand. *Environmental Research Letters, 14*(2), 024010.

Edwards, P. N. (2010). *A vast machine: Computer models, climate data, and the politics of global warming*: Mit Press.

Elsevier, B. (2017). *Scopus: content coverage guide*. Retrieved from www.elsevier.com

Erickson, B., & Widmar, D. A. (2015). Precision Agricultural Services Dealership Survey Results. *Sponsored By CropLife Magazine And The Center For Food And Agricultural Business. Purdue University*.

Ernle, R. E. P. B., & Prothero, R. E. (1936). *The Pioneers and Progress of English Farming* (A. D. Hall Ed. Fifth ed.). London: Longmans, Green.

Falzon, M.-A. (2016). *Multi-sited ethnography: Theory, praxis and locality in contemporary research*: Routledge.

Farmer, B. (1986). Perspectives on the'Green Revolution'in South Asia. *Modern Asian Studies, 20*(1), 175-199.

Fenwick, T., & Nimmo, G. R. (2015). Making visible what matters: Sociomaterial approaches for research and practice in healthcare education. *Researching medical education*, 67-80.

Fielke, S., Nelson, T., Blackett, P., Bewsell, D., Bayne, K., Park, N., Small, B. (2017). Hitting the bullseye: Learning to become a reflexive monitor in New Zealand. *Outlook on AGRICULTURE, 46*(2), 117-124.

Fisher, R. A. (1918). The correlation between relatives on the supposition of Mendelian inheritance. *Transactions of the. Royal Society Edinburgh*, 399-433.

Fisher, R. A. (1925). Statistical methods for research workers Oliver and Boyd. *Edinburgh, Scotland, 6*. Retrieved from http://www-history.mcs.st-and.ac.uk/Extras/Fisher_Statistics_History.html

Fisher, R. A. (1937). *The design of experiments*: Oliver And Boyd; Edinburgh; London.

Fleming, K., Westfall, D., Wiens, D., & Brodahl, M. (2000). Evaluating farmer defined management zone maps for variable rate fertilizer application. *Precision agriculture, 2*(2), 201-215.

Flick, U. (2004). Triangulation in qualitative research. *A companion to qualitative research*, 178-183.

Flowers, M., Weisz, R., & Heiniger, R. (2000). *Aerial photographic determination of nitrogen application timing and rate recommendations in winter wheat.* Paper presented at the Proceedings of the 5th International Conference on Precision Agriculture, Bloomington, Minnesota, USA, 16-19 July, 2000.

Fokoué, E. (2015). A taxonomy of Big Data for optimal predictive machine learning and data mining. *arXiv preprint arXiv:1501.00604*.

Foltz, J. D., & Chang, H.-H. (2002). The adoption and profitability of rbST on Connecticut dairy farms. *American Journal of Agricultural Economics, 84*(4), 1021-1032.

Fonseca, F. (2007). The double role of ontologies in information science research. *Journal of the American Society for Information Science and Technology, 58*(6), 786-793.

Forbes, M. (2011). Pair co-ordinates for global success. *Stuff*. Retrieved from http://www.stuff.co.nz/business/farming/4949695/Pair-co-ordinates-for-global-success

Franzen, D. W., & Mulla, D. J. (2015). A History of Precision Agriculture. In *Precision Agriculture Technology for Crop Farming* (pp. 1-20): CRC Press.

Franzen, D. W., Reitmeier, L., Giles, J. F., & Cattanach, A. C. (1999). Aerial photography and satellite imagery to detect deep soil nitrogen levels in potato and sugarbeet. *Precision agriculture*(precisionagric4a), 281-290.

Fraser, A. (2018). Land grab/data grab: precision agriculture and its new horizons. *The Journal of Peasant Studies*, 1-20.

Fujisaka, S. (1994). Learning from six reasons why farmers do not adopt innovations intended to improve sustainability of upland agriculture. *Agricultural Systems, 46*(4), 409-425.

Funk, C., & Kennedy, B. (2017). Public confidence in scientists has remained stable for decades. *Pew Research Center*.

Gao, L., Yang, G., Yu, H., Xu, B., Zhao, X., Dong, J., & Ma, Y. (2016). Retrieving winter wheat leaf area index based on unmanned aerial vehicle hyperspectral remote sensing. *Transactions of The Chinese Society of Agricultural Engineering, 32*(22), 113-120.

Garud, R. (2008). Conferences as venues for the configuration of emerging organizational fields: The case of cochlear implants. *Journal of management studies, 45*(6), 1061-1088.

Gauch, H. G. (1988). Model selection and validation for yield trials with interaction. *Biometrics*, 705-715.

Gauch, H. G. (2012). *Scientific method in brief*: Cambridge University Press.

Geertz, C. (1974). "From the native's point of view": On the nature of anthropological understanding. *Bulletin of the American Academy of Arts and Sciences*, 26-45.

Gemtos, T. A., Fountas, S., & Aggelopoulou, K. (2011). *Precision Agriculture Applications In Horticultural Crops In Greece and Worldwide.* Paper presented at the HAICTA.

Gill, P. (2008). *The Everyday Lives of Men: An Ethnographic Investigation of Young Adult Male Identity.* Victoria,

Gillingham, A. G. (1973). *Influence of physical factors on pasture growth on hill country.* Paper presented at the Proceedings of the New Zealand Grassland Association.

Gillingham, A. G., & During, C. (1973). Pasture production and transfer of fertility within a long-established hill pasture. *New Zealand journal of experimental agriculture, 1*(3), 227-232.

Gillispie, C. C. (2004). *Science and polity in France: The end of the old regime*: Princeton University Press.

Glaser, B. G., Strauss, A. L., & Strutzel, E. (1968). The discovery of grounded theory; strategies for qualitative research. *Nursing research, 17*(4), 364.

Glesne, C. (2016). *Becoming qualitative researchers: An introduction*: ERIC.

Goldfinch, S., Dale, T., & DeRouen, K. (2003). Science from the periphery: Collaboration, networks and'Periphery Effects' in the citation of New Zealand Crown Research Institutes articles, 1995-2000. *Scientometrics, 57*(3), 321-337.

Goldshleger, N., Chudnovsky, A., & Ben-Binyamin, R. (2013). Predicting salinity in tomato using soil reflectance spectra. *International journal of remote sensing, 34*(17), 6079-6093.

Gotway, C. A., Ferguson, R. B., Hergert, G. W., & Peterson, T. A. (1996). Comparison of kriging and inverse-distance methods for mapping soil parameters. *Soil Science Society of America Journal, 60*(4), 1237-1247.

Gower, B. (2012). *Scientific method: A historical and philosophical introduction*: Routledge.

Grafton, M. C. (2010). *Flow of particulate material from a topdressing aircraft: a thesis presented in partial fulfilment of the requirements of the degree of Doctor of Philosophy in Agricultural Engineering at Massey University, New Zealand.* Massey University,

Grafton, M. C. (2015). Food Science - The case for using emerging technologies to boost crop yields in developing countries. *Policy Forum.* 6 May 2015. Retrieved from https://www.policyforum.net/food-science/

Grafton, M. C., Willis, L., McVeagh, P., & Yule, I. J. (2016). *Measuring pasture mass and quality indices over time using remote and proximal sensors.* Paper presented at the Proceedings of the 13th International Conference on Precision Agriculture (Unpaginated, Online), St. Louis, MO, USA.

Grafton, M. C., Yule, I. J., Davies, C., & Jones, J. (2009). *Incidents and accidents in aerial topdressing in New Zealand: Causes and approaches to mitigation.* Paper presented at the 6th International Conference for Conveying and Handling of Particulate Solids: 3-7 August 2009, Brisbane Convention & Exhibition Centre, Queensland, Australia.

Grafton, M. C., Yule, I. J., & Lockhart, J. (2010). *An economic analysis of the topdressing industry.* Paper presented at the Proc. 23 rd Annual FLRC Conf., Massey University.

Gray, D. I., Parker, W., Kemp, E., Kemp, P. D., Brookes, I., Horne, D., . . . Reid, J. (2003). *Feed planning–alternative approaches used by farmers.* Paper presented at the Proceedings of the New Zealand Grassland Association.

Griffin, T. W. (2011). Precision Agriculture Adoption and Profitability: Fact vs Myth Beyond Pretty Maps. Retrieved from http://past.infoag.org/abstract_papers/papers/abstract_103.pdf

Griffin, T. W., Lambert, D. M., & Lowenberg-DeBoer, J. (2005). Economics of lightbar and auto-guidance GPS navigation technologies. *Precision agriculture, 5*.

Griffin, T. W., Lowenberg-DeBoer, J., Lambert, D. M., Peone, J., Payne, T., & Daberkow, S. G. (2004). Adoption, profitability, and making better use of precision farming data. *Staff paper*(04-06).

Guillemin, M., & Gillam, L. (2004). Ethics, reflexivity, and “ethically important moments” in research. *Qualitative inquiry, 10*(2), 261-280.

Gunther, F., & March, R. (1956). Instruments in agriculture - the pesticides field. *JOURNAL OF AGRICULTURAL AND FOOD CHEMISTRY, 4*(7), 598-602.

Haboudane, D., Miller, J. R., Pattey, E., Zarco-Tejada, P. J., & Strachan, I. B. (2004). Hyperspectral vegetation indices and novel algorithms for predicting green LAI of crop canopies: Modeling and validation in the context of precision agriculture. *Remote sensing of environment, 90*(3), 337-352.

Haboudane, D., Miller, J. R., Tremblay, N., Zarco-Tejada, P. J., & Dextraze, L. (2002). Integrated narrow-band vegetation indices for prediction of crop chlorophyll content for application to precision agriculture. *Remote sensing of environment, 81*(2-3), 416-426.

Hall Jamieson, K., McNutt, M., Kiermer, V., & Sever, R. (2019). Signaling the trustworthiness of science. *Proceedings of the National Academy of Sciences, 116*(39), 19231-19236.

Hammersley, M. (2018). What is ethnography? Can it survive? Should it? *Ethnography and Education, 13*(1), 1-17.

Hammersley, M., & Atkinson, P. (2007). *Ethnography: Principles in practice*: Routledge.

Hanseth, O., & Monteiro, E. (1998). *Changing irreversible networks.* Paper presented at the ECIS.

Harrin, E. (2007). Project management in the real world. *British Computer Society in association with CAPDM.*

Harris, J. A. (1920). Practical universality of field heterogeneity as a factor influencing plot yields. *Journal of Agricultural Research, 19*(7), 279-314.

Hasu, M., Leitner, K.-H., Solitander, N., & Varblane, U. (2012). 5 Accelerating the Innovation Race. *Challenging the Innovation Paradigm, 9*, 87.

Hedley, C. B. (2009). *The developments of proximal sensing methods for soil mapping and monitoring, and their application to precision irrigation.* (Doctor of Philosophy in Soil Science Doctoral Thesis). Massey University, Palmerston North New Zealand. Retrieved from https://mro.massey.ac.nz/bitstream/handle/10179/1217/02whole.pdf?sequence=2&isAllowed=y

Hedley, C. B. (2015). The role of precision agriculture for improved nutrient management on farms. *Journal of the Science of Food and Agriculture, 95*(1), 12-19.

Hedley, C. B., Bradbury, S., Ekanayake, J., Yule, I. J., & Carrick, S. (2010). *Spatial irrigation scheduling for variable rate irrigation.* Paper presented at the Proceedings of the New Zealand Grassland Association.

Hedley, C. B., & Yule, I. J. (2012). *Farmer uptake of variable rate irrigation technologies in New Zealand.* Paper presented at the 10th international conference on precision agriculture.

Hermanowicz, J. C. (2002). The great interview: 25 strategies for studying people in bed. *Qualitative sociology, 25*(4), 479-499.

Hey, T., Tansley, S., & Tolle, K. M. (2009). *The fourth paradigm: data-intensive scientific discovery* (Vol. 1): Microsoft research Redmond, WA.

Hillison, J. (1996). The origins of agriscience: Or where did all that scientific agriculture come from? *Journal of Agricultural Education, 37*, 8-13.

Hoffmann, V. (2011). *Knowledge and Innovation Management. Module Reader. Hohenheim University*.

Huberty, M. (2015). Awaiting the second big data revolution: from digital noise to value creation. *Journal of Industry, Competition and Trade, 15*(1), 35-47.

Hudson, B. D. (1992). The soil survey as paradigm-based science. *Soil Science Society of America Journal, 56*(3), 836-841.

Hugo, W. (2009). Spiraling reference: a case study of apprenticeship into an academic community of practice. *South African journal of higher education, 23*(4), 703-721.

IEEE. (2019, 5/06/2019). Koen Langendoen. Retrieved from https://ieeexplore.ieee.org/author/37326885400

International Society of Precision Agriculture. (2019). Precision Ag Definition. Retrieved from https://www.ispag.org/about/definition

Iphofen, R. (2013). Research ethics in ethnography/anthropology. *European Commission (http://ec. europa. eu/research/participants/data/ref/h2020/other/hi/ethics-guide-ethnog-anthrop_en. pdf)*.

Irrigation New Zealand. (2017). Why We Irrigate. Retrieved from https://www.irrigationnz.co.nz/KnowledgeResources/Category?Action=View&Category_id=210

Irwin, K. (2006). Into the dark heart of ethnography: The lived ethics and inequality of intimate field relationships. *Qualitative sociology, 29*(2), 155-175.

Isik, M., & Khanna, M. (2003). Stochastic technology, risk preferences, and adoption of site-specific technologies. *American Journal of Agricultural Economics, 85*(2), 305-317.

Jacobs, N., & McFarlane, A. (2005). Conferences as learning communities: some early lessons in using 'back-channel'technologies at an academic conference–distributed intelligence or divided attention? *Journal of Computer Assisted Learning, 21*(5), 317-329.

Jago, J., Eastwood, C. R., Kerrisk, K., & Yule, I. J. (2013). Precision dairy farming in Australasia: adoption, risks and opportunities. *Animal Production Science, 53*(9), 907-916.

Johnson, R., Watkinson, A., & Mabe, M. (2018). The STM report. *An overview of scientific and scholarly publishing. 5th edition October*.

Jones, D. B., & Tadajewski, M. (2016). *The Routledge companion to marketing history*: Routledge.

Jones, E. L. (1965). Agriculture and economic growth in England, 1660–1750: agricultural change. *The Journal of Economic History, 25*(1), 1-18.

Jones, J. W., Antle, J. M., Basso, B., Boote, K. J., Conant, R. T., Foster, I., Janssen, S. (2017). Brief history of agricultural systems modeling. *Agricultural Systems, 155*, 240-254.

Jouanjean, M.-A. (2019). Digital opportunities for trade in the agriculture and food sectors.

Kamilaris, A., Kartakoullis, A., & Prenafeta-Boldú, F. X. (2017). A review on the practice of big data analysis in agriculture. *Computers and Electronics in Agriculture, 143*, 23-37.

Kaye-Blake, B. N., Chris; Parker, Chris; Dunn, Michael. (2014). *Economic contribution of PGP - A cost-benefit analysis of potential impacts*. Retrieved from Wellington, New Zealand: https://www.mpi.govt.nz/dmsdocument/4622/loggedIn

Kell, D. B., & Oliver, S. G. (2004). Here is the evidence, now what is the hypothesis? The complementary roles of inductive and hypothesis-driven science in the post-genomic era. *Bioessays, 26*(1), 99-105.

Kelly, A. (2009). In defence of anonymity: rejoining the criticism. *British Educational Research Journal, 35*(3), 431-445.

Kemp, P., & Lopez, I. (2016). *Hill country pastures in the southern North Island of New Zealand: an overview*. Paper presented at the Hill Country Symposium. Grassland Research and Practice. Series Nº16.

Kerridge, E. (1969). The agricultural revolution reconsidered. *Agricultural History, 43*(4), 463-476.

Khosla, R. (2010). *Precision agriculture: challenges and opportunities in a flat world.* Paper presented at the 19th World Congress of Soil Science.

Khun, K., Vigneault, P., Tremblay, N., Bouroubi, M. Y., Cavayas, F., & Codjia, C. (2016). *Comparative benefits of drone imagery for nitrogen status determination in corn* Paper presented at the 13th International Conference on Precision Agriculture, St. Louis, Missouri, USA.

Kitchin, R. (2014). *The data revolution: Big data, open data, data infrastructures and their consequences*: Sage.

Kitchin, R., & Lauriault, T. P. (2015). Small data in the era of big data. *GeoJournal, 80*(4), 463-475.

Kitchin, R., & Lauriault, T. P. (2018). Digital data and data infrastructures. In J. Ash, Kitchin, R. and Leszczynski, A. (Ed.), *Digital Geographies* (pp. pp. 83-94). London: Sage.

Kitchin, R., & McCardle, G. (2014). Big Data, new epistemologies and paradigm shifts. *Big Data & Society, 1*(1), 2053951714528481.

Klinkenborg, V. (2012, 4 November 2012). Did Farmers of the Past Know More Than We Do? *The New York Times*, p. SR12. Retrieved from https://www.nytimes.com/2012/11/04/opinion/sunday/crop-rotation-and-the-future-of-farming.html

Knook, J., & Turner, J. A. (2020). Reshaping a farming culture through participatory extension: An institutional logics perspective. *Journal of Rural Studies, 78*, 411-425.

Knorr Cetina, K. (2001). Laboratory studies: The cultural approach to the study of science. *Handbook of science and technology studies*.

Knorr, K. D. (1981). The Manufacture of Knowledge An Essay on the Constructivist and Contextual Nature of Science.

Kritikos, M. (2017). *Precision agriculture in Europe: Legal, social and ethical considerations*. Brussels: European Unition Retrieved from www.europarl.europa.eu/RegData/etudes/STUD/.../EPRS_STU(2017)603207_EN.pdf

Lamb, D. W., Frazier, P., & Adams, P. (2008). Improving pathways to adoption: Putting the right P's in precision agriculture. *Computers and Electronics in Agriculture, 61*(1), 4-9.

Lamb, J., Anderson, J., Malzer, G., Vetch, J., Dowdy, R., Onken, D., & Ault, K. (1995). *Perils of monitoring grain yield on-the-go*. Paper presented at the Site-specific management for agricultural systems.

Lambert, D. M., Paudel, K. P., & Larson, J. A. (2015). Bundled Adoption of Precision Agriculture Technologies by Cotton Producers. *Journal of Agricultural and Resource Economics, 40*(2), 325-345.

Lambert, M. G., & Roberts, E. (1978). Aspect differences in an unimproved hill country pasture: II. Edaphic and biotic differences. *New Zealand Journal of Agricultural Research, 21*(2), 255-260.

Langendoen, K., Baggio, A., & Visser, O. (2006). *Murphy loves potatoes: Experiences from a pilot sensor network deployment in precision agriculture*. Paper presented at the Parallel and Distributed Processing Symposium, 2006. IPDPS 2006. 20th International.

Lark, R. (2001). Some tools for parsimonius modelling of within-field variation of soil and crop systems. *Soil & Tillage Res, 58*(3-4), 99-111.

Larsen, W., Nielsen, G., & Tyler, D. (1994). Precision navigation with GPS. *Computers and Electronics in Agriculture, 11*(1), 85-95.

Larson, J. A., Roberts, R. K., English, B. C., Cochran, R. L., & Wilson, B. S. (2005). A computer decision aid for the cotton yield monitor investment decision. *Computers and Electronics in Agriculture, 48*(3), 216-234.

Latour, B. (1987). *Science in action: How to follow scientists and engineers through society*: Harvard university press.

Latour, B. (1993). *The pasteurization of France*: Harvard University Press.

Latour, B. (1996). On actor-network theory: A few clarifications. *Soziale welt*, 369-381.

Latour, B., & Woolgar, S. (1986). Laboratory life: the construction of scientific facts. In: Princeton, NJ: Princeton University Press.

Law, J. (2009). Actor network theory and material semiotics. *Social theory*, 141.

Lawrence, H. G. (2007). Adoption of precision agriculture technologies for fertiliser placement in New Zealand: a thesis presented in partial fulfilment of the requirements of the degree of Doctor of Philosophy in Agricultural Engineering at Massey University, Palmerston North, New Zealand.

Lawrence, S. (2001). Online or invisible. *Nature, 411*(6837), 521.

Lelong, C. C., Pinet, P. C., & Poilvé, H. (1998). Hyperspectral imaging and stress mapping in agriculture: a case study on wheat in Beauce (France). *Remote sensing of environment, 66*(2), 179-191.

Levy, E. B., & Madden, E. A. (1933). The point method for pasture analysis. *NZ Journal of Agriculture, 46*, 267-279.

Li, F., Mistele, B., Hu, Y., Chen, X., & Schmidhalter, U. (2014). Optimising three-band spectral indices to assess aerial N concentration, N uptake and aboveground biomass of winter wheat remotely in China and Germany. *ISPRS Journal of Photogrammetry and Remote Sensing, 92*, 112-123.

Lile, J., Blackwell, M., Thomson, N., Penno, J., Macdonald, K., Nicholas, P., Coulter, M. (2001). *Practical use of the rising plate meter (RPM) on New Zealand dairy farms.* Paper presented at the Proceedings of the Conference-New Zealand Grassland association.

Lindblom, J. L., C.; Ljung, M.; Jonsson, A. (2014). *Adoption Of Precision Agriculture In Sweden – The Case Of Soil Maps*. Paper presented at the Proceedings of the 12th International Conference on Precision Agriculture., Sacramento, California, USA.

Lindsay Corporation. (2019). Field MAP. Retrieved from https://zimmaticanz.com/fieldmap

Linnenluecke, M. K., Marrone, M., & Singh, A. K. (2020). Conducting systematic literature reviews and bibliometric analyses. *Australian Journal of Management, 45*(2), 175-194.

Lissaman, W., Casey, M., & Rowarth, J. (2013). Innovation and technology uptake on farm. *Proceedings of the New Zealand Grassland Society, 75*, 27-32.

Liu, A., Guo, Y., & Guan, R. (2010). *The research status on precision agriculture by use of bibliometric analysis from three databases.* Paper presented at the World Automation Congress (WAC), 2010.

Livestock Improvement Corporation. (2017). Pastures from SPACE™: Satellite technology signals game changer for kiwi farmers [Press release]. Retrieved from https://www.lic.co.nz/news/pastures-space-satellite-technology-signals-game-changer-kiwi-farmers/

Lofland, J. (1976). Doing social life: The qualitative study of human interaction in natural settings.

Lomax, A., Memon, A., & Painter, B. D. (2010). The Canterbury Water Management Strategy as a collaborative planning initiative: a preliminary assessment.

Lowenberg-DeBoer, J., & Boehlje, M. (1996). Revolution, evolution or dead-end: Economic perspectives on precision agriculture. *Precision agriculture*(precisionagricu3), 923-944.

Lynch, M., & Amann, K. (1996). Scientific practice and ordinary action: Ethnomethodology and social studies of science. *ISIS-International Review Devoted to the History of Science and its Cultural Influence, 87*(4), 766-766.

Maat, H. (2011). The history and future of agricultural experiments. *NJAS-Wageningen Journal of Life Sciences, 57*(3), 187-195.

MacLeod, A., Cameron, P., Ajjawi, R., Kits, O., & Tummons, J. (2019). Actor-network theory and ethnography: Sociomaterial approaches to researching medical education. *Perspectives on medical education, 8*(3), 177-186.

Malinowski, B. (1967). *A Diary in the Strict Sense of the Term; Pref. by Valetta Malinowska; Introd. by Raymond Firth; Tr. by Norbert Guterman; Index of Terms by Mario Bick*: Routledge & K. Paul.

Marchant, B. P., Daley, A. G., & Webster, R. (2013). Rational management decisions from uncertain spatial data. In *Precision Agriculture for Sustainability and Environmental Protection* (pp. 79-99): Routledge.

Marcus, A. I. (1985). *Agricultural science and the quest for legitimacy: Farmers, agricultural colleges, and experiment stations, 1870-1890*: Iowa State University Press.

Mariotto, I., Thenkabail, P. S., Huete, A., Slonecker, E. T., & Platonov, A. (2013). Hyperspectral versus multispectral crop-productivity modeling and type discrimination for the HyspIRI mission. *Remote sensing of environment, 139*, 291-305.

Market Study Report. (2020). *Global precision agriculture market size to record substantial growth by 2025*. Retrieved from USA: https://www.globenewswire.com/news-release/2020/08/04/2072232/0/en/Global-precision-agriculture-market-size-to-record-substantial-growth-by-2025.html

Martín-Martín, A., Orduña-Malea, E., Ayllón, J. M., & López-Cózar, E. D. (2016). The counting house: Measuring those who count. Presence of bibliometrics, scientometrics, informetrics, webometrics and altmetrics in the Google Scholar citations, Researcherid, ResearchGate, Mendeley & Twitter. *arXiv preprint arXiv:1602.02412*.

Marton, F. (1981). Phenomenography—describing conceptions of the world around us. *Instructional science, 10*(2), 177-200.

Massey University. (2015). *Code of Ethical Conduct for Research, Teaching and Evaluations Involving Human Participants*

Massey University. (2016). New imaging tool to revolutionise agriculture [Press release]. Retrieved from https://www.massey.ac.nz/massey/about-massey/news/article.cfm?mnarticle_uuid=F249C860-91D5-3693-72A5-890346C3FDB2

Mata, G., Clark, D., Edirisinghe, A., Waugh, D., Minneé, E., & Gherardi, S. (2007). *Predicting accurate paddock-average pasture cover in Waikato dairy farms using satellite images.* Paper presented at the Proceedings of the Conference - New Zealand Grassland Association.

Matese, A., Toscano, P., Di Gennaro, S. F., Genesio, L., Vaccari, F. P., Primicerio, J., Gioli, B. (2015). Intercomparison of UAV, aircraft and satellite remote sensing platforms for precision viticulture. *Remote Sensing, 7*(3), 2971-2990.

Matheron, G. (1963). Principles of geostatistics. *Economic geology, 58*(8), 1246-1266.

Matthew, C. (2019). Levy Oration 2018: Our industry through the eyes of a pasture agronomist. *Journal of New Zealand Grasslands, 81*, 11-14.

Mayer, J. F. (1768). *Die Lehre vom Gyps als einem vorzüglich guten Dung zu allen Erd-Gewächsen auf Aeckern und Wiesen, Hopfen-und Weinbergen*: Posch.

Mayer, J. F. (1773). *Lehrbuch für die Land-und Haußwirthe in der pragmatischen Geschichte der gesamten Land-und Haußwirthschafft des Hohenlohe Schillingsfürstischen Amtes Kupferzell*: Zeh.

McAfee, A., Brynjolfsson, E., Davenport, T. H., Patil, D., & Barton, D. (2012). Big data. *The management revolution. Harvard Bus Rev, 90*(10), 61-67.

McBratney, A. B., Whelan, B. M., Ancev, T., & Bouma, J. (2005). Future directions of precision agriculture. *Precision agriculture, 6*(1), 7-23.

Medyckyj-Scott, D., Stock, K., R., G., M., G., Dzierzon, H., J., S., & Collins, A. (2016). *A Data Ecosystem for Land and Water Data to Achieve the Challenge Mission*. Retrieved from New Zealand: https://ourlandandwater.nz/wpcontent/uploads/2019/04/OLWDataEcosystemWhitePaper-Final-1.pdf

Merriam, S. B., & Tisdell, E. J. (2015). *Qualitative research: A guide to design and implementation*: John Wiley & Sons.

Miller, H. J. (2010). The data avalanche is here. Shouldn't we be digging? *Journal of Regional Science, 50*(1), 181-201.

Ministry for Primary Industries. (2014). New programme set to transform hill country farms. Retrieved from https://www.mpi.govt.nz/news-and-resources/media-releases/new-programme-set-to-transform-hill-country-farms-/

Ministry for Primary Industries. (2019). *Apiculture: Ministry for Primary Industries 2019 Apiculture Monitoring Programme*. Wellington, New Zealand Retrieved from https://www.mpi.govt.nz/dmsdocument/40849/direct#:~:text=The%20value%20of%20New%20Zealand%27s,volume%20(down%207%20percent).

Mintert, J., Widmar, D., Langemeier, M., Boehlje, M., & Erickson, B. (2016). The Challenges of Precision Agriculture: Is Big Data the Answer?

Mogues, T. (2015). Political Economy Determinants of Public Spending Allocations: A Review of Theories, and Implications for Agricultural Public Investment. *The European Journal of Development Research, 27*(3), 452-473. doi:10.1057/ejdr.2015.35

Montgomery, M. (Producer). (2019, 1/12/2020). Image Bands. *Geospatial Online -Introduction to Remote Sensing*. Retrieved from https://www.google.com/search?q=humboldt+state+university&rlz=1C1CHBF_enNZ804NZ804&oq=humboldt+state+university&aqs=chrome..69i57j46j0l6.16745j0j7&sourceid=chrome&ie=UTF-8

Morgan, J., & Connolly, E. (2013). Plant-soil interactions: nutrient uptake. *Nature Education Knowledge, 4*(8), 2.

Morton, J., Stafford, A., Gillingham, A. G., Old, A., & Knowles, O. (2016). The development of variable rate application of fertiliser from a fixed wing topdressing aircraft. *Grassland Research and Practice Series, 16*, 163-167.

Mueller, G. (1977). Catholics and Protestants: Agricultural Modernization in Two German Villages by Guenter Golde. New York: Academic Press, 1975, xxxii+ 198 pp. $12.95. In: Oxford University Press.

Mulla, D. J. (1993). *Mapping and managing spatial patterns in soil fertility and crop yield*. Paper presented at the Proceedings of Soil Specific Crop Management.

Mulla, D. J. (2013). Twenty five years of remote sensing in precision agriculture: Key advances and remaining knowledge gaps. *Biosystems Engineering, 114*(4), 358-371.

Mulla, D. J., Bhatti, A., Hammond, M., & Benson, J. (1992). A comparison of winter wheat yield and quality under uniform versus spatially variable fertilizer management. *Agriculture, ecosystems & environment, 38*(4), 301-311.

Muller, R. (2020). Orthorectification. Retrieved from https://www.dlr.de/eoc/en/desktopdefault.aspx/tabid-6144/10056_read-20918/

Murray, R. (2007). *Variable rate application technology in the New Zealand aerial topdressing industry: a thesis presented in partial fulfilment of the requirements of the degree of Doctor of Philosophy in Agricultural Engineering at Massey University, New Zealand.* Massey University,

Murray, R., Yule, I. J., & Gillingham, A. G. (2007). Developing variable rate application technology: modelling annual pasture production on hill country. *New Zealand Journal of Agricultural Research, 50*(1), 41-52.

Natural Resources Canada. (2015, 25/11/2020). Glossary of Remote Sensing Terms. Retrieved from https://www.nrcan.gc.ca/maps-tools-publications/satellite-imagery-air-photos/satellite-imagery/satellite-imagery-products/educational-resources/glossary-remote-sensing-terms/9483

Nawar, S., Corstanje, R., Halcro, G., Mulla, D., & Mouazen, A. (2017). Delineation of Soil Management Zones for Variable-Rate Fertilization. In *Advances in Agronomy* (Vol. 143, pp. 175-245): Elsevier.

Nelson, W. (2017, 16-18 October 2017). *Proceedings of the 7th Asian-Australasian Conference on Precision Agriculture.* Paper presented at the 7th Asian-Australasian Conference on Precision Agriculture.

Neményi, M., Mesterhazi, P., Pecze, Z., & Stépán, Z. (2003). The role of GIS and GPS in precision farming. *Computers and Electronics in Agriculture, 40*(1-3), 45-55.

Nimmo, R. (2011). Actor-network theory and methodology: Social research in a more-than-human world. *Methodological Innovations Online, 6*(3), 108-119.

Nola, R., & Irzik, G. (2005). Naive Inductivism as a Methodology in Science. *Philosophy, Science, Education and Culture*, 207-230.

Norton, B. J. (1978). Karl Pearson and Statistics The Social Origins of Scientific Innovation. *Social studies of science, 8*(1), 3-34.

Nowak, P. (1997). A sociological analysis of site-specific management. *The State of Site-Specific Management for Agriculture*(thestateofsites), 397-422.

Nowak, P. (1998). Agriculture and change: the promises and pitfalls of precision. *Communications in Soil Science & Plant Analysis, 29*(11-14), 1537-1541.

Nuthall, P., & Bishop-Hurley, G. (1999). Feed planning on New Zealand farms. *Journal of International Farm Management, 2*, 100-112.

NZIER. (2019). Water policy is top of mind… [Press release]. Retrieved from https://nzier.org.nz/media/water-policy-is-top-of-mind

OECD. (2007). *OECD Reviews of Innovation Policy New Zealand 2007*. Retrieved from http://www.oecd-ilibrary.org/science-and-technology/oecd-reviews-of-innovation-policy_19934211#

Oester, S., Cigliano, J. A., Hind-Ozan, E. J., & Parsons, E. (2017). Why Conferences Matter—An Illustration from the International Marine Conservation Congress. *Frontiers in Marine Science, 4*, 257.

Office of the Auditor General. (2015). *Primary Growth Partnership Financial Management: Assurance on use of PGP funding - Ravensdown, Pioneering to Precision*. Wellington, New Zealand

Oliver, D. M., Fish, R. D., Winter, M., Hodgson, C. J., Heathwaite, A. L., & Chadwick, D. R. (2012). Valuing local knowledge as a source of expert data: farmer engagement and the design of decision support systems. *Environmental Modelling & Software, 36*, 76-85.

Oliver, M. A. (2010). *Geostatistical applications for precision agriculture*: Springer Science & Business Media.

Oliver, M. A., Bishop, T., & Marchant, B. (2013). *Precision agriculture for sustainability and environmental protection*: Routledge.

Olson, K. D., & Elisabeth, P. (2003). *An economic assessment of the whole-farm impact of precision agriculture.* Paper presented at the International Conference on Precision Agriculture.

Olykan, S. T., Lucas, R. J., & Moot, D. J. (2019). Sampling methods for clover species in grazed pastures to diagnose mineral deficiencies. *Journal of New Zealand Grasslands, 81*, 241-248.

Orduña-Malea, E., Ayllón, J. M., Martín-Martín, A., & López-Cózar, E. D. (2015). Methods for estimating the size of Google Scholar. *Scientometrics, 104*(3), 931-949.

Ortiz, B., Balkcom, K., Duzy, L., Van Santen, E., & Hartzog, D. (2013). Evaluation of agronomic and economic benefits of using RTK-GPS-based auto-steer guidance systems for peanut digging operations. *Precision agriculture, 14*(4), 357-375.

Overton, M. (1996). *Agricultural revolution in England: the transformation of the agrarian economy 1500-1850* (Vol. 23): Cambridge University Press.

Paddock, W. C. (1970). How green is the green revolution? *BioScience*, 897-902.

Pallottino, F., Biocca, M., Nardi, P., Figorilli, S., Menesatti, P., & Costa, C. (2018). Science mapping approach to analyze the research evolution on precision agriculture: world, EU and Italian situation. *Precision agriculture*, 1-16.

Passioura, J. (1973). Sense and nonsense in crop simulation. *J. Aust. Inst. Agric. Sci, 39*(3), 181-183.

Pesce, M., Kirova, M., Soma, K., Bogaardt, M., Poppe, K., Thurston, C., Urdu, D. (2019). Research for AGRI Committee—Impacts of the Digital Economy on the Food Chain and the CAP. *European Parliament, Policy Department for Structural and Cohesion Policies: Brussels, Belgium*, 80.

Pfister, H., & Blitzstein, J. (2013). Course: CS109 Data Science. In.

Pickering, A. (1992). *Science as practice and culture*: University of Chicago Press.

Pretty, J. N. (1994). Alternative systems of inquiry for a sustainable agriculture. *IDS bulletin, 25*(2), 37-49.

Provost, F., & Fawcett, T. (2013). *Data Science for Business: What you need to know about data mining and data-analytic thinking*: " O'Reilly Media, Inc.".

Provost, L. (2015). *Ministry for Primary Industries: Managing the Primary Growth Partnership*. Retrieved from https://www.oag.govt.nz/2015/primary-growth

Pullanagari, R. R., Yule, I. J., King, W., Dalley, D., & Dynes, R. A. (2011). The use of optical sensors to estimate pasture quality. *International Journal on Smart Sensing and Intelligent Systems, 4*(1), 125-137.

Pullanagari, R. R., Yule, I. J., Tuohy, M., Hedley, M. J., Dynes, R. A., & King, W. (2012). In-field hyperspectral proximal sensing for estimating quality parameters of mixed pasture. *Precision agriculture, 13*(3), 351-369.

Racelogic. (2018, 23 August 2018). How does DGPS (Differential GPS) work? Retrieved from https://racelogic.support/01VBOX_Automotive/01General_Information/Knowledge_Base/How_Does_DGPS_(Differential_GPS)_Work%3F

Radcliffe, J., Dale, W., & Viggers, E. (1968). Pasture production measurements on hill country. *New Zealand Journal of Agricultural Research, 11*(3), 685-700.

Rao, N. (2017). Big Data and Climate Smart Agriculture-Review of Current Status and Implications for Agricultural Research and Innovation in India.

Rao, N. R. (2008). Development of a crop-specific spectral library and discrimination of various agricultural crop varieties using hyperspectral imagery. *International journal of remote sensing, 29*(1), 131-144.

Ravensdown Fertiliser Co-operative Ltd. (2014). Transforming hill country farming returns with remote soil scanning and variable rate aerial application. Retrieved from http://www.ravensdown.co.nz/media/1879/pgp-brochure.pdf

Ravensdown Fertiliser Co-operative Ltd. (2017). Spreading - Aerial. Retrieved from https://www.ravensdown.co.nz/media/4326/aerowork-brochure.pdf

Ravensdown Fertiliser Co-operative Ltd. (2018). Ravensdown PGP gets new funding boost [Press release]. Retrieved from https://www.ravensdown.co.nz/expertise/ravensdown-pgp-gets-new-funding-boost

Ravensdown Fertiliser Co-operative Ltd. (2019a). Ravensdown to reinvest in innovation and environment [Press release]. Retrieved from https://www.scoop.co.nz/stories/BU1908/S00355/ravensdown-to-reinvest-in-innovation-and-environment.htm

Ravensdown Fertiliser Co-operative Ltd. (2019b). Smart Maps. Retrieved from https://www.ravensdown.co.nz/services/my-ravensdown/smart-maps

Rennie, G., King, W., Puha, M., Dalley, D., Dynes, R., & Upsdell, M. (2009). *Calibration of the C-DAX Rapid pasturemeter and the rising plate meter for kikuyu-based Northland dairy pastures.* Paper presented at the Proceedings of the New Zealand Grassland Association.

Rhoads, B. L., Wilson, D., Urban, M., & Herricks, E. E. (1999). Interaction between scientists and nonscientists in community-based watershed management: emergence of the concept of stream naturalization. *Environmental management, 24*(3), 297-308.

Richter, R., & Schläpfer, D. (2002). Geo-atmospheric processing of airborne imaging spectrometry data. Part 2: Atmospheric/topographic correction. *International journal of remote sensing, 23*(13), 2631-2649.

Rijswijk, K., Bewsell, D., Small, B., & Blackett, P. (2015). Reflexive monitoring in New Zealand: Evaluation lessons in supporting transformative change. *Evaluation Journal of Australasia, 15*(4), 38-43.

Robert, P., & Stafford, J. (1999). *Precision Agriculture: research needs and status in the USA.* Paper presented at the Precision agriculture'99, Part 1. Papers presented at the 2nd European Conference on Precision Agriculture, Odense, Denmark, 11-15 July 1999.

Roberts, T. (2007). Right product, right rate, right time and right place… the foundation of best management practices for fertilizer. *Fertilizer Best Management Practices, 29*.

Robertson, M., Isbister, B., Maling, I., Oliver, Y., Wong, M., Adams, M., Tozer, P. (2007). Opportunities and constraints for managing within-field spatial variability in Western Australian grain production. *Field Crops Research, 104*(1), 60-67.

Robertson, M., Llewellyn, R., Mandel, R., Lawes, R., Bramley, R. G. V., Swift, L., . O'Callaghan, C. (2012). Adoption of variable rate fertiliser application in the Australian grains industry: status, issues and prospects. *Precision agriculture, 13*(2), 181-199.

Roche, D. G., Kruuk, L. E., Lanfear, R., & Binning, S. A. (2015). Public data archiving in ecology and evolution: how well are we doing? *PLoS Biology, 13*(11).

Rossini, M., Fava, F., Cogliati, S., Meroni, M., Marchesi, A., Panigada, C., Amaducci, S. (2013). Assessing canopy PRI from airborne imagery to map water stress in maize. *ISPRS Journal of Photogrammetry and Remote Sensing, 86*, 168-177.

Roumasset, J. A. (1976). Rice and risk; decision making among low-income farmers.

Rutto, E., & Arnall, D. B. (2004). The History of the GreenSeeker™ Sensor. In: Sillwater, Oklahoma: Oklahoma State University Cooperative Extension Service.

Ryan, M. (2019). Agricultural Big Data Analytics and the Ethics of Power. *Journal of Agricultural and Environmental Ethics*, 1-21.

Sagan, C. (2011). *The demon-haunted world: Science as a candle in the dark*: Ballantine Books.

Sanches, I. D. A. (2009). *Hyperspectral proximal sensing of the botanical composition and nutrient content of New Zealand pastures.* (Doctor of Philosophy in Earth Science Doctoral). Massey University, Palmerston North. Retrieved from https://mro.massey.ac.nz/bitstream/handle/10179/1194/02whole.pdf?sequence=1&isAllowed=y

Saunders, B., Kitzinger, J., & Kitzinger, C. (2015). Anonymising interview data: Challenges and compromise in practice. *Qualitative Research, 15*(5), 616-632.

Schellberg, J., Hill, M. J., Gerhards, R., Rothmund, M., & Braun, M. (2008). Precision agriculture on grassland: Applications, perspectives and constraints. *European Journal of Agronomy, 29*(2), 59-71.

Schieffer, J., & Dillon, C. (2015). The economic and environmental impacts of precision agriculture and interactions with agro-environmental policy. *Precision agriculture, 16*(1), 46-61.

Schimmelpfennig, D., & Ebel, R. (2016). Sequential Adoption and Cost Savings from Precision Agriculture. *Journal of Agricultural and Resource Economics, 41*(1).

Schmoldt, D. L. (2001). Precision agriculture and information technology. *Computers and Electronics in Agriculture, 30*(1), 5-7.

Schoengold, K., & Sunding, D. L. (2014). The impact of water price uncertainty on the adoption of precision irrigation systems. *Agricultural Economics, 45*(6), 729-743.

Schön, D. A. (1983). *The reflective practitioner: How professionals think in action* (Vol. 5126): Basic books.

Schrijver, R. (2016). Precision Agriculture and the Future of Farming in Europe. *Scientific Foresight Study*.

Schueller, J. K., & Bae, Y. H. (1987). Spatially attributed automatic combine data acquisition. *Computers and Electronics in Agriculture, 2*(2), 119-127.

SCImago Research Group (CSIC) OECD. (2016). *Compendium of Bibliometric Science Indicators*. Paris: OECD Retrieved from http://oe.cd/scientometrics.

Searcy, S., Schueller, J., Bae, Y., Borgelt, S., & Stout, B. (1989). Mapping of spatially variable yield during grain combining. *Transactions of the ASAE, 32*(3), 826-0829.

Shapin, S. (2018). *The scientific revolution*: University of Chicago Press.

Shaw, D. R., & Willers, J. L. (2006). Improving pest management with remote sensing. *Outlooks on Pest Management, 17*(5), 197-201.

Sismondo, S. (2011). *An introduction to science and technology studies*: John Wiley & Sons.

Sonka, S., & Cheng, Y.-T. (2015a). Big Data: Alive and Growing in the Food Sector! *farmdoc daily*.

Sonka, S., & Cheng, Y.-T. (2015b). Precision Agriculture: Not the Same as Big Data But *farmdoc daily*.

Sonka, S., & IFAMR, I. (2014). Big data and the ag sector: more than lots of numbers. *International Food and Agribusiness Management Review, 17*(1), 1.

SPECIM Spectral Imaging Ltd. (2016). Aisa FENIX Hyperspectral Sensor. Retrieved from http://www.specim.fi/downloads/AisaFenix_ver1-2015.pdf

Spiegelhalter, D. (2020). Should we trust algorithms? *Harvard Data Science Review, 2.1*. Retrieved from https://doi.org/10.1162/99608f92.cb91a35a

Sposito, G., & Reginato, R. (1992). Pedology: The science of soil development. *G. Sposito*, 9-25.

Staenz, K., Szeredi, T., & Schwarz, J. (1998). ISDAS–A system for processing/analyzing hyperspectral data. *Canadian Journal of Remote Sensing, 24*(2), 99-113.

Statistics New Zealand. (2019, 18 April 2019). Irrigated Land. Retrieved from https://www.stats.govt.nz/indicators/irrigated-land

Sternberg, R. J. (1991). Robert J. Sternberg. *Psychological Bulletin, 109*(1), 3-4.

Stolzy, L., Cahoon, G., & Szuszklewicz, T. (1957). Measurement of soil moisture: Accurate instruments for measuring soil moisture conditions practical means of determining proper timing of irrigation. *California Agriculture, 11*(4), 21-37.

Sullivan, M. (2012). A brief history of GPS. *PC World*. Retrieved from URL: http://www. pcworld. com/article/2000276/a-brief-history-of-gps. html

Sunding, D., & Zilberman, D. (2001). The agricultural innovation process: research and technology adoption in a changing agricultural sector. *Handbook of agricultural economics, 1*, 207-261.

Sveiby, K.-E., Gripenberg, P., & Segercrantz, B. (2012). *Challenging the innovation paradigm*: Routledge.

Swinton, S. M., & Lowenberg-DeBoer, J. (1998). Evaluating the profitability of site-specific farming. *Journal of production agriculture, 11*(4), 439-446.

Swinton, S. M., & Lowenberg-DeBoer, J. (2001). *Global adoption of precision agriculture technologies: Who, when and why*. Paper presented at the Proceedings of the 3rd European Conference on Precision Agriculture.

Talebpour, B., Türker, U., & Yegül, U. (2015). The role of precision agriculture in the promotion of food security. *International Journal of Agricultural and Food Research, 4*(1).

Tey, Y. S., & Brindal, M. (2012). Factors influencing the adoption of precision agricultural technologies: a review for policy implications. *Precision agriculture, 13*(6), 713-730.

Thompson, F. M. L. (1968). The second agricultural revolution, 1815-1880. *The Economic History Review, 21*(1), 62-77.

Thompson, N., & Pascal, J. (2012). Developing critically reflective practice. *Reflective practice, 13*(2), 311-325.

Thomson, N. (1983). Factors influencing the accuracy of herbage mass determinations with a capacitance meter. *New Zealand journal of experimental agriculture, 11*(2), 171-176.

Thomson, N., Upsdell, M., Hooper, R., Henderson, H., Blackwell, M., McCallum, D., Bishop-Hurley, G. (2001). *Development and evaluation of a standardised means for estimating herbage mass of dairy pastures using the rising plate meter*. Paper presented at the Proceedings of the conference - New Zealand Grassland Association.

Timmer, C. P. (1969). The turnip, the new husbandry, and the English agricultural revolution. *The Quarterly Journal of Economics*, 375-395.

Tolich, M. (2004). Internal confidentiality: When confidentiality assurances fail relational informants. *Qualitative sociology, 27*(1), 101-106.

Tolle, K. M., Tansley, D. S. W., & Hey, A. J. G. (2011). The Fourth Paradigm: Data-Intensive Scientific Discovery [Point of View]. *Proceedings of the IEEE, 99*(8), 1334-1337. doi:10.1109/jproc.2011.2155130

Toth, C., & Jóźków, G. (2016). Remote sensing platforms and sensors: A survey. *ISPRS Journal of Photogrammetry and Remote Sensing, 115*, 22-36.

Toynbee, A., & Milner, A. (1894). *Lectures on the industrial revolution of the eighteenth century in England* (Fourth ed.). London: Longmans, Green and Co.

Tranfield, D., Denyer, D., & Smart, P. (2003). Towards a methodology for developing evidence-informed management knowledge by means of systematic review. *British journal of management, 14*(3), 207-222.

Tremblay, N. (2017). *Confronting the challenges of big data for precision agriculture*. Paper presented at the Presentation at International Society of Precision Agriculture Annual Meetings, February.

United States House of Representatives. (1997). Precision Agriculture Research, Education, and Information Dissemination Act of 1997. *1*, 16. Retrieved from https://www.congress.gov/105/bills/hr725/BILLS-105hr725ih.pdf

University of Hawaiʻi. (2020, 2020). Practices of Science: Precision vs. Accuracy. Retrieved from https://manoa.hawaii.edu/exploringourfluidearth/physical/world-ocean/map-distortion/practices-science-precision-vs-accuracy#:~:text=Accuracy%20refers%20to%20how%20close,Precision%20is%20independent%20of%20accuracy.&text=The%20best%20quality%20scientific%20observations%20are%20both%20accurate%20and%20precise.

Van Maanen, J. (2011). *Tales of the field: On writing ethnography*: University of Chicago Press.

Vázquez, F. (2018, 8 Dec 2018). Ontology and Data Science. Retrieved from https://towardsdatascience.com/ontology-and-data-science-45e916288cc5

Verdouw, C. N., Beulens, A. J. M., Reijers, H. A., & van der Vorst, J. G. A. J. (2015). A control model for object virtualization in supply chain management. *Computers in industry, 68*, 116-131. doi:10.1016/j.compind.2014.12.011

Von Bueren, S., & Yule, I. J. (2013). Multispectral aerial imaging of pasture quality and biomass using unmanned aerial vehicles (UAV). *Accurate and Efficient Use of Nutrients on Farms, Occasional Report*(26).

Walton, J. C., Lambert, D. M., Roberts, R. K., Larson, J. A., English, B., Larkin, S. L., Reeves, J. M. (2008). Adoption and abandonment of precision soil sampling in cotton production. *Journal of Agricultural and Resource Economics*, 428-448.

Walton, J. C., Roberts, R. K., Lambert, D. M., Larson, J. A., English, B. C., Larkin, S. L., Reeves, J. M. (2010). Grid soil sampling adoption and abandonment in cotton production. *Precision agriculture, 11*(2), 135-147.

Wanninger, L. (2004). *Introduction to network RTK*. Paper presented at the International Association to Geodesy meeting, Working Group.

Whelan, B. M. (2011). A review of the history of Precision Agriculture in Australia and some future opportunities. *Proceedings of the Precision Agriculture Australia Conference 2011, 2011*, 4.

Whelan, B. M., & McBratney, A. B. (2000). The “null hypothesis” of precision agriculture management. *Precision agriculture, 2*(3), 265-279.

White, M. (2021). *Ravensdown PGP Programme Quarterly Progress Report*. Retrieved from

Wiles, R., Crow, G., Heath, S., & Charles, V. (2008). The management of confidentiality and anonymity in social research. *International journal of social research methodology, 11*(5), 417-428.

(2020). The Oprah Conversation with Isabel Wilkerson [Television series episode]. In Winfrey, O. (Executive producer), *The Oprah Conversation*: Apple TV+.

Winstead, A. T., Norwood, S. H., Griffin, T. W., Runge, M., Adrian, A. M., Fulton, J. P., & Kelton, J. (2010). *Adoption and use of precision agriculture technologies by practitioners.* Paper presented at the Proceedings of the 10th International Conference of Precision Agriculture. Denver, Colorado: CDROM.

Wolf, S. A., & Wood, S. D. (1997). Precision farming: Environmental legitimation, commodification of information, and industrial coordination. *Rural sociology, 62*(2), 180-206.

Wolfert, S., Ge, L., Verdouw, C., & Bogaardt, M.-J. (2017). Big data in smart farming–a review. *Agricultural Systems, 153*, 69-80.

Wolinsky, H. (2009). The crash reaches the universities. *EMBO reports, 10*(3), 209-211.

Wu, C., Niu, Z., Tang, Q., & Huang, W. (2008). Estimating chlorophyll content from hyperspectral vegetation indices: Modeling and validation. *Agricultural and forest meteorology, 148*(8-9), 1230-1241.

Wu, M., Huang, W., Niu, Z., Wang, Y., Wang, C., Li, W., Yu, B. (2017). Fine crop mapping by combining high spectral and high spatial resolution remote sensing data in complex heterogeneous areas. *Computers and Electronics in Agriculture, 139*, 1-9.

Young, S. L., & Carter, P. G. (2013). Applications of remote sensing in precision agriculture for sustainable production. In *Precision Agriculture for Sustainability and Environmental Protection* (pp. 100-116): Routledge.

Yule, I. J. (2011). Precision Agriculture-A view to the Future. Conference presentation. *Precision Agriculture–A View to the Future [Online]. Christchurch, Thursday, 11*.

Yule, I. J. (2015). *Sensors - What Have We Got And Where Are We Headed?* Paper presented at the In: Moving farm systems to improved attenuation.

Yule, I. J. (2016). *Precision Agriculture in New Zealand*. Paper presented at the InfoAg 2016, St Louis, Missouri, USA.

Yule, I. J., Grafton, M. C., & Pullanagari, R. R. (2014, July 2014). *Exploiting the Variability in Pasture Production on New Zealand Hill Country.* Paper presented at the 12th International Symposium on Precision Agriculture, Sacramento.

Yule, I. J., Lawrence, H., Hedley, C. B., Pullanagari, R. R., Draganova, I., Betteridge, K., & Dynes, R. A. (2010). *Technologies to improve productivity and reduce environmental impact of dairy systems.* Paper presented at the EcoTechs Conference: International workshop: Eco-design Methods and Tools for Agricultural Technologies, Clermont-Ferrand, France.

Yule, I. J., Lawrence, H., & Murray, R. (2010). *Pasture yield measurement with the CDAX pasture meter.* Paper presented at the Proceedings of the International Conference on Precision Agriculture.

Yule, I. J., Pullanagari, R. R., Irwin, M., McVeagh, P., White, M., Kereszturi, G., & Manning, M. (2015). Mapping nutrient concentration in pasture using hyperspectral imaging. *Journal of New Zealand Grasslands, 77*, 47-50.

Yule, I. J., & Wood, B. A. (2014). *Precision Agriculture as Bricolage: Understanding the Site Specific Farmer.* Paper presented at the 12th International Conference on Precision Agriculture, Sacramento, California USA.

Yule, I. J., & Wood, B. A. (2015). How do we transform our industry to achieve a modern vision for New Zealand Agriculture? *, In: Moving farm systems to improved attenuation. (Eds L.D. Currie and L.L Burkitt)*.

Zarco-Tejada, P., Hubbard, N., & Loudjani, P. (2014). Precision Agriculture: An Opportunity for EU Farmers—Potential Support with the CAP 2014-2020. *Joint Research Centre (JRC) of the European Commission*.

Zhang, B., & Tillman, R. (2007). A decision tree approach to modelling nitrogen fertiliser use efficiency in New Zealand pastures. *Plant and Soil, 301*(1), 267-278.

Zhang, B., Tillman, R., Gillingham, A., & Gray, M. (2009). Fine-scale spatial modelling of pasture production and fertiliser responses under variable climate: Assessing a decision support tool in pasture management. *New Zealand Journal of Agricultural Research, 52*(4), 455-470.

Zhang, N., Wang, M., & Wang, N. (2002). Precision agriculture—a worldwide overview. *Computers and Electronics in Agriculture, 36*(2), 113-132.