Methodology for a GIS-based Land-use Map for Southland: A Review

Envirolink Advice Grant: ESRC274

Methodology for a GIS-based land-use map for Southland: a review

Daniel Rutledge, Andrew Manderson, Linda Lilburne, Anne-Gaelle Ausseil, Stella Belliss, Robbie Price

Landcare Research

Prepared for:

Environment Southland

Private Bag 90116
Invercargill
New Zealand

March 2016

Landcare Research, Gate 10 Silverdale Road, University of Waikato Campus, Private Bag 3127, Hamilton 3240, New Zealand, Ph +64 7 859 3700, Fax +64 7 859 3701, www.landcareresearch.co.nz

Reviewed by:	*Approved for release by:*
Fraser Morgan Senior Scientist Landcare Research	Suzie Greenhalgh Portfolio Leader – Supporting Business & Policy Landcare Research
Landcare Research Contract Report:	LC2491

Disclaimer

This report has been prepared by Landcare Research for Environment Southland. If used by other parties, no warranty or representation is given as to its accuracy and no liability is accepted for loss or damage arising directly or indirectly from reliance on the information in it.

Contents

Summary

Project and Client

- Methodology for a GIS-based Land-use Map for Southland: A Review
- Environment Southland

Objectives

- Undertake an urgent review of land-use mapping methods under active development by Environment Southland in support of policy formulation to meet the objectives of the National Policy Statement – Freshwater Management including modelling that links land-use with water quality
  - Review of suitability of the methods for determining land use
  - Review of limitations of the methods and data sources
  - Recommend potential improvements on the current methods
  - Recommendations on potential ground-truthing techniques.

Methods

  - Stage 1 (November-December 2015): Initial scoping, comprehensive review and provision of draft report providing review and recommendations of Version 1 land-use mapping methods
  - Stage 2 (January 2016): Workshop with Environment Southland staff to discuss the review and recommendations
  - Stage 3: (February–March 2016): Final check of Version 2 land-use mapping methods that began to incorporate Stage 1 recommendations and production of final report.

Recommendations

- Broad Principles (High-level Recommendations)
  - Suitability – start by outlining the target land-use classification that ideally meets the intended NPS-FM purposes and then develop the land-use methods that come as close as possible to realising the target classification
  - Repeatability – carefully and rigorously document all methods, processes, decisions, etc.; if possible avoid manual editing, and if not possible, implement a system to track any manual changes made
  - Accountability – form an internal Land Use Technical Advisory Group to oversee the process of land-use methodology development including formulating the land-use classification, executing and formally recording decisions, and serving as a conduit between the land-use mapping project team and the full set of internal end-users.

- Specific Recommendations

  Specific recommendations are too detailed and numerous to list here. Instead we outline some key recommendations that the Version 2 methods have implemented:

  - Provides a target land-use classification to address suitability
  - Reorders the layering of input data including using the LINZ Parcel Layer as the base layer
  - Uses the 2014 Protected Areas Network (PAN-NZ) database to help map protected and conservation areas
  - Identifies data sources in more detail when feasible or allowed
  - Simplifies and clarifies processing of AgriBase™ data
  - Provides several new figures to help illustrate the implementation of various rules
  - Expands discussion of validation, key assumptions and limitations
  - Includes detailed GIS methods as an Appendix to address repeatability.

Conclusions

- Overall, Environment Southland's approach to land-use mapping is sound and uses common data sources and mirrors similar efforts and methods undertaken by other regional councils across New Zealand. Improvements from Version 1 to Version 2 showed substantial progress towards and willingness to improve their approach.
- Going forward Environment Southland will continue to revise and refine both their land-use classification and their mapping methods, particularly as more detail on the modelling linking land use and water quality matures.
- Environment Southland and Landcare Research discussed the possibility of providing guidance on sensitivity analyses. We agreed to defer providing such advice pending availability of more details on the new interated land use-water quality modelling that they are currently developing.
- A key upcoming challenge for Environment Southland will be revising the classification and methods, or adapting them, to take advantage of the results of an industry-provided representative farm enterprise survey.
- Ultimately mapping farm systems, rather than generic or even slightly more detailed farm categories or reference farms, would provide the ideal foundation for robust modelling of land use-water quality relationships.

1 Introduction

In November 2015 Environment Southland approached Landcare Research to undertake an urgent review of new methods under development to map regional land use in support of activities related to implementing the National Policy Statement on Freshwater Management (NPS-FM), specifically activities within the associated NPS-FM Science Programme and Southland Economic Project.

Following initial discussions regarding the nature and scope of the project, Landcare Research recommended that Environment Southland apply for an Envirolink Medium Advice Grant to fund the project. Environment Southland subsequently prepared a successful application to Envirolink, which funded the project.

In December 2015 Landcare Research undertook a comprehensive review of the draft land-use methods and provided a draft report to Environment Southland. A 2-day workshop with Landcare Research was held in the Invercargill offices of Environment Southland in January 2016 to discuss the overall methods and specific draft recommendations. In mid-February 2016 Landcare Research under a final check of Version 2 of the land-use methods currently under development by Environment Southland. Version 2 implemented some, but not all, of the recommendations from the draft review report.

This report outlines the full set of recommendations and the results of the final check.

2 Background

Environment Southland requires knowledge of land use to respond to the NPS-FM and to be able to set limits. "Land-use Impacts" is a research theme with a series of projects to identify and fill critical knowledge gaps in the field of contaminant sources, loadings and transport, associated with rural and urban land use. The results of the projects will serve as key inputs to the policy development in Environment Southland's Water and Land 2020 & Beyond Strategy and into the Economic Programme associated with the NPS-FM, which will help formulate policy, planning, rules and guide management practices at finer scales (i.e. farm scales) needed to achieve the goals and objectives of the NPS-FM within the region. Environment Southland has developed a draft methodology for classifying and mapping land use based on the best available science and data available. Environment Southland now wants to ensure that the methods developed meet exacting standards that will withstand legal scrutiny and integrate any existing available information to strengthen the approach.

The land-use information, captured at regional and catchment levels, will feed into land-use based models for estimating nutrient losses, sediment and bacteria loadings and subsequent models through other aspects of the overall NPS-FM research programme. More spatially accurate and detailed land-use data and mapping will improve estimations of sources of threats to water quality and improve modelling of source-observation-impact processes and relationships. Collecting accurate land-use information will enable the Council to measure more accurately the effects of land use on water quality and the receiving ecosystems and the effects of policy actions on land use, land management, and the regional economy. The land-use information will be critical for conceptual and numerical model development for scenario testing. Making robust policy decisions affecting land use cannot be undertaken unless the council has the best land-use information possible to inform decisions. The land use information will have immediate application into other areas of research and be rolled out through limit setting processes in the Freshwater Management Units (FMUs) from 2016.

The peer reviewed land-use mapping methodology and resulting maps will be used across Council. A high quality land-use map will provide immediate and direct information for policy, planning and field staff. In addition, the land-use information and mapping will inform other research considering high nutrient loss land uses or landscapes. The results of these investigations will improve catchment models and a conceptual understanding of land-use losses and pathways of losses and assist with scenario modelling and limit-setting. To be as robust as possible, catchment-scale modelling should start with land-use information and maps that are as accurate as possible. The land-use information will also help determine the potential economic consequences of implementing proposed policies and rules for different land uses within catchment and FMUs. Having high-quality, peer-reviewed approaches to land-use mapping and development of land-use maps, based on best available defensible science, is an essential aspect of Environment Southland's response to the NPS-FM.

3 Objectives

The work plan consisted of five objectives:

1. Review of suitability of the methods for determining land use
2. Review of limitations of the methods and data sources
3. Recommend potential improvements on the current methods
4. Recommendations on potential ground-truthing techniques
5. Presentation and half-day workshop by the Project Lead with key client staff.

4 Methods

The project proceeded in three stages.

4.1 Stage 1: Comprehensive Expert Review

Landcare Research staff with relevant substantial experience in characterising and mapping land use (see Appendix 1), including as input into modelling that supports water resources management, reviewed the Environment Southland Version 1 methods report (Pearson 2015).

Each Landcare Research project team member independently reviewed the report from their own perspective and areas of expertise, although all staff focused on informing Objectives 1–4 outlined above. To aid the reviews, Landcare Research also prepared a visual work flow of the current methods and a spreadsheet that tracked the provenance of each primary and secondary land-use category by mapping all potential pathways between potential input data and sources and the output land-use categories.

Reviews from all Landcare Research project team members were collated and integrated into a draft review report transmitted to Environment Southland in December 2015. The visual work flows and provenance spreadsheet were also provided as additional supporting information electronically.

4.2 Stage 2: Two-day Workshop

Daniel Rutledge, the Project Lead, spent two days, 18–19 January 2016, workshopping the Stage 1 recommendations with the Environment Southland Project Lead and other key staff at council offices in Invercargill in fulfilment of Objective 5. The workshop covered both broad principles for robust land-use classification and mapping as well as robust discussions regarding the detailed comments and recommendations provided on the Version 1 report.

4.3 Stage 3: Final Check

Between Stage 2 and Stage 3, Environment Southland began to implement the recommendations from Stages 1 and 2 and produced a draft Version 2 methods report (Pearson & Couldrey 2016). Given the project scope and the evolving nature of the methods, Stage 3 consisted only of a final check of the Version 2 methods and not a comprehensive review as in Stage 1. The final check involved broadly assessing which recommendations Environment Southland adopted from Stage 1 and recommending, where appropriate and feasible, any additional areas for further improvement.

5 Recommendations

We present the results in a logical but slightly out-of-order sequence compared with the three project stages outlined above. We start with three high-level observations resulting from the Stage 1 review that were discussed at the workshop (Objective 5). Each of the three high-level observations served as the basis for a corresponding guiding principle for land-use mapping outlined later. We then present the results organised by the other four contracted objectives.

5.1 High-level Observations & Principles

This section briefly outlines three high-level principles derived from our overall observations regarding the process and procedures that Environment Southland is developing to classify and map land use in the region on an on-going basis. The specific land-use mapping methods reviewed and discussed later are a necessary, but not sufficient, component of the broader processes and procedures that Environment Southland must develop to maximise the value and minimise the risks going forward given the intended future uses of the land-use classification, spatial data layers and maps.

The issue of risk becomes particularly critical, given that the resulting land-use information will inform potentially sensitive questions of permissible land use and land-use management related to fresh water management as a result of implementing the NPS-FM. Given those potential sensitivities, Environment Southland must undertake a high level of due diligence to insure their overall approach to land-use classification and mapping will withstand scrutiny, including possibly legal tests such as Environment Court cases.

5.1.1 Principle 1: Suitability

Suitability is both a high-level principle and a specific assessment. As a high-level principle, suitability involves determining the ideal land-use classification that will best suit the intended purpose, goals, or objectives. As a specific assessment, suitability involves determining whether the particular data, methods, etc. being are appropriate for generating the ideal land-use classification. This section discusses suitability as a high-level principle. The section "Objective 1" below outlines the specific suitability assessment of the data and methods being developed by Environment Southland for the current project.

Developing a land-use classification requires an adaptive management approach. The exercise should proceed by:

- outlining the intended purpose, goals and objectives that the land-use classification must fulfil
- formulating an "ideal" or target classification in consultation with end-users (internal and external) that meets the intended purpose, goals, and objectives
- gathering and documenting all currently existing required data
- tracking and documenting all currently unavailable required data
- iteratively developing methods to transform and process the current data to generate the a realised classification.

The adaptive management process continues throughout methodological development. As different choices and compromises are made, the methods, realised classification, and even the target classification will likely undergo modification in a dynamic, integrated fashion.

One of the most challenging aspects of land-use classification and mapping involves deciding when to stop i.e. formulating and agreeing "stopping rules." Except for very simple purposes such as broadly reporting or mapping land-use, the realised land-use classification rarely meets the target classification. Having a robust and well-defined target classification helps minimise the challenge by providing a stable target at which to aim.

Relative to Environments Southland's current project and purpose, the target classification should specify in detail the land-use classification and resulting spatial data needed by the integrated land use-water quality modelling under development. From a modelling perspective, a good starting point is to conceptually outline the specific land-use inputs required by the water quality models.

At this stage, Environment Southland is exploring two options for water quality modelling that follow similar but somewhat conceptually different pathways:

1) *Loss modelling*: land-use information helps define and quantify inputs for Overseer-based modelling of nutrient losses from farms that would be part of the Environment Southland Flows Framework.

2) *Input modelling*: land-use information helps define and quantify estimates of nutrient inputs that will be then related via modelling to observed values of nutrients from extensive monitoring data. This approach may also use a regional physiographic framework under development by Environment Southland to help tailor the modelling to FMUs.

The level of detail of the land-use classification, particularly relative to primary production classes, will affect modelling under either pathway. For example, classifying land use at a high, generic level, e.g. dairy farm, sheep & beef farm, deer farm, arable farm, etc., will necessitate making generic assumptions. In other words, the same farm management practice, such as level of fertiliser application, would apply to each generic land-use class across the region.

A more detailed approach already implemented by other regions (Hill et al. 2012) uses the concept of representative farm enterprises. In that approach, several types of representative farm types exist under each higher-level category, i.e. several types of dairy farms, several types of sheep & beef farms, etc. The representative farms may also vary spatially across catchments or sub-regions as appropriate given a combination of environmental conditions and data availability. Each farm is then classified according to the best fitting representative farm type. This approach therefore provides more detailed matching between actual and modelling inputs and theoretically should improve overall model performance.

Finally, the ideal situation would involve having access to farm management plans and nutrient budgets to map and model each farm enterprise as accurately as possible as a farm system. Modelling based on farm systems would provide the best basis within each catchment for

- assessing each farm enterprise's relative contribution within different catchments
- understanding of the current overall nutrient budget and the gap, if any, relative to a target nutrient budget
- developing fair and equitable strategies for reducing any gaps to meet freshwater management outcomes under the NPS-FM, particularly via identification of similar farm systems and sharing of best management practices or development of novel management practices.

Of course, the relationship between level of desired detail achieved versus level of effort expended is positive and possibly even exponential, although the continued advancement of technology including "big data" and web-based technologies will continue to increase progress towards and feasibility of more farm system-type approaches.

Overall, Environment Southland has undertaken to date an adaptive management process to formulate and implement a target land-use classification. However, the detailed review below indicates that the process could be improved. Environment Southland would benefit by specifying at the outset a more detailed and desired target classification that would then guide their subsequent efforts and processes. On the other hand, the evolving nature of the water quality constrains Environment Southland's ability to outline a stable target classification, as the target in one sense keeps moving and evolving.

5.1.2 Principle 2: Repeatability

Repeatability is an essential characteristic of and objective for any land-use classification and mapping exercise. Repeatability is also a critical consideration for Environment Southland in this context. Undoubtedly they will need to regenerate the *exact* same classification and/or map in the future, including, for example, to map, monitor and report land-use change, cater for independent verification of results by outside parties in support of evidence-based policy development, or possibly give evidence in court proceedings.

At its simplest, repeatability means that someone other than the those involved in developing the methods and producing the land-use classification and maps can repeat the exercise and produce the same results. Repeatability is also clearly the foundation of sound scientific discourse and debate; conversely, lack of repeatability means the process is by nature unscientific.

Repeatability requires complete and thorough documentation. The Version 1 methods, which to our knowledge represented the extent of the formal documentation at the time of the Stage 1 review and subsequent discussions at the workshop, indicated a low probability that any party other than the Environment Southland land-use mapping project team could duplicate the results. Methods, when provided, were often vague, ambiguous and/or incomplete. Sometimes no methods were provided, or only general statements were made. For example, in several cases, the methods suggested that a change or edit was made manually, e.g. interactive editing of the boundaries of a polygon, manually resolving sliver issues, etc. Without any further documentation or edit control, there would be no way to repeat the exact same process reliably.

5.1.3 Principle 3: Accountability

Classifying and mapping land use requires making many choices, including sometimes compromises, given various considerations, either individually or in combination. Choices can result from consideration of the underlying purpose and objectives, data availability and limitations, and/or varying or evolving needs.

Robust classification and land-use mapping therefore requires robust accountability of what decisions were made, the process for making them, and who was/was not involved in making them, including what person or persons had final authority, etc. Operationalised properly robust accountability provides a solid and defensible evidence base, including understanding the origins and evolution of the process involved, clearly demonstrating the train of thinking and logic involved, and articulating how different decisions were made.

The Stage 1 comprehensive review indicated that overall accountability was relatively low. Documentation of methods was incomplete, as highlighted above. Decision-making processes were not clearly articulated relative to a number of considerations including

- how the land-use classification was formulated and agreed
- how changes or revisions were considered and decided
- how the resulting land-use data and maps would meet the needs of various internal and external stakeholders, etc.

The lack of accountability makes Environmental Southland highly vulnerable to future challenges. They would be hard-pressed to provide satisfactory and defensible evidence of the decision-making processes, work flows, or various procedures followed, as many of them appeared informal and unrecorded.

5.2 Objective 1: Review of Version 1 Suitability of Methods for Determining Land Use

The land-use classification being developed by Environment Southland has a key specific purpose of supporting modelling linkages between contaminant sources and surrounding receiving waters. The modelling will help Environment Southland better understand current and potential future relationships among contaminant sources, total loadings into receiving waters, key water-quality indicators, and, eventually, help them set appropriate targets for managing contaminants to meet water quality goals and objectives.

Ideally, a land-use classification would be flexible enough to cater to/for multiple, or at least more than one, purpose (Rutledge et al. 2009), but limitations of data and/or resources usually necessitate making choices in classification design to favour one particular purpose. Table 1 below illustrates that point. It provides examples of the land-use classifications developed for four studies involving different modelling efforts that assist with understanding linkages between source areas and water quality (Hill et al. 2012; Lilburne et al. 2012, 2013; Rutledge et al. 2012). For a broader overview see also Anastasiadis et al. (2013).

Table 1 not only yields insights into the variability in land-use characterisation commonly encountered but also highlights some useful common elements for land-use classifications related to water resources modelling.

The main differences among the examples in Table 1 relates to their relative breadth and degree of specificity. The first three examples provide classifications specifically targeted at water resources modelling using models such as OVERSEER®, SPAMSO, AquaiferSIM, etc., with a particular focus on nitrogen leaching at the farm scale. The fourth example is from an integrated model that includes the Sparrow model. For the first three, classification detail differs somewhat but common key features include specification of primary land use types (e.g. beef, cropping, dairy, dairy support, deer, sheep & beef), stocking rates and winter on/wintering off for dairy farming, use of irrigation or not, more detailed partitioning of sheep and/or beef farming systems, and, for the Canterbury land-use mapping exercise, more detailed land-uses for arable, fruit, etc.

The land-use classification for the integrated model (column four) is more general for primary production land-uses, but broader overall as the integrated model includes several components requiring additional land-use categories. In addition, the Sparrow model has simpler information requirements in terms of land use, given that it uses statistical modelling of relationships between percentages of land-use categories per sub-catchment and observed nutrient loads, whereas other models such as OVERSEER®, AquiferSim, etc., ideally rely on more detailed farm scale-inputs, such as what a farm plan would provide.

The methods being developed generate a map of land use for the Southland region using 25 land-use categories (Table 2). Two of the land-use categories (Native, Lakes and Ponds) actually represent land cover, not land use. Most, but not all, categories have at least one corresponding short, usually three-letter, code. Many of these codes are the same as the three-letter Farm Type codes found in the AgriBase™ database. Several categories that represent amalgamations of different AgriBase™ Farm Types have multiple corresponding codes. Four categories appear to have no corresponding code.

The first review comment relates to the inconsistency among different realisations of the primary land-use classification used throughout the Version 1 report. The inconsistencies relate to the number of categories, category names, and codes. The land-use classification appears in the report three times in tabular form: Table 1 (p. 8), Table 3 (p. 22), and in the table in Appendix 1.

Table 1 in the Version 1 report lists 15 primary land use categories, sub-categories, and a description for some, but not all, categories. Table 3 in the report lists 25 categories, "ES subcategories" that contain lists of corresponding codes, except for Conservation. The table in Appendix 1 lists the primary land use categories and "ES Codes," which in this case contain a single code, usually three-letters. The Appendix 1 table is titled as "Classification codes used in GIS layer" and seems to serve as summary table for showing how the original input data, primarily AgriBase™, maps to the primary land-use categories om the output.

The variability among the different representations causes confusion. The number of categories is inconsistent between Tables 1 and Table 3. A note in Table 1 indicates that "some of the sub-categories have become independent categories," which partly explains the discrepancy. However seeing as the sub-categories in Table 1 never reappear in the report, we question why they are mentioned at all. One category, Irrigation Area, also never appears again in the report after Table 1.

Category names are sometimes inconsistent: "Urban" vs "Urban and Residential", Plantation Forest vs "Plantation Forestry", etc. Similarly, the use of codes is inconsistent. Table 3 suggests that the codes relate to primary land-use subcategories, while the table in Appendix 1 suggests that each primary land-use category has a single corresponding code.

Beyond those inconsistencies, there is the question of the appropriateness of the categories themselves. Using the examples provided in Table 1 of the Version 1 report as a general guide, the list of 25 primary land-use categories seems to provide the minimum information needed to map land-use at the degree of detail needed to infer potential contaminant source areas.

However the categories as currently conceived do not provide the best representation possible for helping estimate and model contaminant sources and losses. In particular, the classification would benefit from more detailed land-use classes that reflect variation in stocking rates for dairy and sheep and/or beef. In addition, there should be some consideration of how to differentiate among farm types with/without irrigation and, for dairy farming, with or without wintering off.

The Conservation land-use category appears to be derived solely from the AgriBase™ Not Farmed (NOF) farm types. In our experience the Not Farmed farm type does not equate with conservation in most cases, although some records can be inferred as conservation. For example, the three largest Not Farmed farms in the March 2015 version of AgriBase™ appear to equate to areas of the public conservation estate managed by DOC based on the provided Stations Names. However, most other 'Not Farm' farm records indicate a purpose other than conservation or the purpose cannot be inferred from the information provided. There are numerous data sources for obtaining more robust delineation of protected areas and/or conservation that could be used and which we outline in the recommendations.

Table 1: Example land-use classifications developed for different modelling needs related to water resources management

AquiferSim Modelling of the mid-Mataura Basin (Lilburne et al. 2012)	Estimating nitrate-nitrogen leaching rates under rural land use in Canterbury (updated) (Lilburne et al. 2013)	Preparation of a GIS-based land-use map for the Canterbury region (Hill et al. 2012)	Waikato Integrated Scenario Explorer Technical Specifications 1.2 (Rutledge et al. 2012)
Cropping – Mixed – Irrigated – Winter Grazing	Arable	Arable – Continuous Irrigated & Dryland	Aquaculture
Cropping – Mixed – Non-irrigated – Winter Grazing	Beef – 100% - Dryland	Arable – Mixed Grazing Irrigated & Dryland	Airports
Cropping – Seasonal - Irrigated – No Winter Grazing	Beef – 100% - Irrigated	Arable – Seasonal Grazing Irrigate & Dryland	Bare Surfaces
Cropping – Seasonal – Non-irrigated – No Winter	Dairy – 3 cows per ha – winter-off	Beef – 100% Irrigated & Dryland	Biofuel Cropping
Dairy – <= 3.5 cows per ha – Winter Grazing	Dairy – 4 cows per ha – winter-on	Commercial Vegetable	Commercial
Dairy – > 3.5 cows per ha – Winter Grazing	Dairy – 4 cows per ha – winter-off	Dairy – 3 cows per ha, winter on	Community Services
Sheep & Beef – Intensive 80:20	Dairy – 5 cows per ha – winter-off	Dairy – 3 cows per ha, winter off	Dairy Farming
Sheep – Dryland	Dairy Support – irrigated	Dairy – 4 cows per ha, winter off	Forestry
	Dairy Support - Dryland	Dairy – 4 cows per ha, winter on	Freshwater
	Deer – 100% - Irrigated or Dryland	Dairy – 5 cows per ha, winter off	Horticulture
	Forestry – Exotic	Dairy Support	Indigenous Vegetation
	Forestry – Native	Deer – Irrigated & Dryland	Manufacturing
	Fruit Trees	Forestry – Exotic	Marine
	Golf	Forestry – Native	Mines and Quarries
	Lifestyle	Fruit – apple	Other Exotic Vegetation
	Pigs	Fruit – berry	Residential – High Density
	Vegetables	Fruit – summer fruit	Residential – Lifestyle Blocks
		Fruit – unspecified, mixed	Residential – Low Density
		Golf	Other Agriculture
		Grapes	Other Cropping
		Horses	Sheep, Beef or Deer Farming
		Lifestyle pastoral	Urban Parks & Recreation
		Pigs	Utilities
		Sheep/Beef – 10% beef, irrigated, dryland	Vegetable Cropping
		Sheep/Beef – 20% beef, irrigated dryland	Wetlands
		Urban or settlements	

Table 2: Environment Southland Primary Land Use Categories and Codes

Primary Land Use Category	Code
Arable	ARA
Beef	BEE
Conservation	NOF
Dairy	DAI
Dairy Support	DAISUP
Deer	DEE
Deer and/or Sheep and Beef	BND, SBD, SND
Flowers	FLO
Golf	-
Horticulture	FRU,VEG
Industry and Airports	IND, AIR
Lakes and Ponds	-
Lifestyle Blocks	LIF, LIF1, LIF2
Native	NAT
Nursery	NUR
Other Animals	OAN
Pasture Unspecified Type	PAS
Plantation Forestry	FOR
Rail	-
Recreation	REC, TOU
Road	-
Sheep	SHP
Sheep and Beef	SNB
Support Block	SUP
Urban and Residential	RES

5.3 Objective 2: Review of Version 1 Data Source and Methods Limitations

5.3.1 Data Source Limitations

Overview

Table 3 summarises the data sources currently being used to generate the Southland land-use map including comments on their key strengths and key limitations. The table, where possible, also estimates the update frequencies for the data sources. This is to help highlight the on-going challenge of trying to compile and map land use regularly using a range of data sources with varying vintages and update cycles.

The current data sources by-and-large consist of the handful of data sources commonly employed to map land use for a range of purposes including those related to water resource modelling and management. Efforts at mapping land use typically rely on several nationally-consistent, authoritative, and in most cases public, data sources: AgriBase™, Land Cover Database, LINZ topographic data, and cadastral property data, also maintained by LINZ (Hill et al. 2012; Lilburne et al. 2012, 2013; Morgan et al. 2010; Rutledge et al. 2009, 2010, 2012, 2015). In some cases, including for Environment Southland, third parties are contracted to augment property data for different purposes including resource management.

In addition to the common data sources used, the methods also rely on several Southland-specific data sources including information from Environment Southland's consent database for dairy farming and data provided in consultation with several industries. We compliment Environment Southland for taking advantage of such data sources, especially working directly with industry to help fill gaps and/or provide more accurate mappings of several key land use types within the region. Such data sources and interactions will be critical going forward to generate more robust land-use mapping to support a range of policy, planning, and resource management purposes.

A key missing data source, as outlined earlier, relates to protected areas and/or conservation. There are several existing data sources that could be used to map those areas, which we outline in the recommendations.

In addition to the overall assessment of data sources, below we provide comments on several specific data sources that we feel will improve the methods and resulting land-use mapping.

Table 3: Current data sources used to compile the Southland land-use map

Data Source	Organisation	Key Strength(s)	Key Limitation(s)	Version/ Date	Update Frequency
AgriBase™	AssureQuality	Comprehensive Consistent Often the only source of more detailed primary production land-use data	Not public Voluntary Some records substantially out of date Some overlaps, both legitimate and erroneous	April 2015 (October 2015 update received)	6-monthly
Arable	Foundation for Arable Research	Provides more detail on arable farming	Relies on industry cooperation	??	?? – possibly on request
Consent Database	Environment Southland	Comprehensive and historic record of dairy farms > 100 ha	Does not include dairy farms < 100 ha in size	??	Assume rolling updates as new consents are approved
Deer	Deer Farming Industry	Provides more detail on deer farming	Relies on industry cooperation	??	?? – possibly on request
Plantation Forestry	Forest Industry	Provides more detail on plantation forestry	Relies on industry cooperation	??	?? – possibly on request
Land Cover Database	Landcare Research	Authoritative Consistent Public	Inability to differentiate some land uses Infrequent updates and on-going uncertainty of support	4.1 2012	4-5 Years
Property Layer	LINZ	Authoritative Consistent Public		2015	Weekly rolling updates
Topographic Data	LINZ	Authoritative Consistent Public	Updates vary in frequency and coverage	2015	At least annually (varies)

AgriBase™

AgriBase™ is a particularly useful layer in that there are few readily available alternatives that capture agricultural land use and enterprises in such fine spatial detail and are, to some degree, accessible. However, AgriBase™ is also an imperfect database that relies on voluntary input from farmers, is incomplete in terms of coverage, and contains data configurations that capture real world land use situations that can be difficult to process from an analysis perspective (e.g. overlapping polygons).

As AgriBase™ represents the base data source on land use for the mapping exercise, we highlight some potential issues to flag when using that data, which the current methods being used by Environment Southland have partially addressed:

- Reported farm types may not readily agree with other attribute information provided, particularly for pastoral farming involving multiple types of stock.
- Farm boundaries may not represent the actual location of the farm. We know of some records where farm boundaries occur in urban areas and therefore seem to represent the property where the owner lives, rather than the farm itself.
- Records vary in terms of their currency. In the full AgriBase™ database, some records date from as early as 1993. From a monitoring and reporting standpoint, records greater than 5 years old would not qualify for the purposes of compiling Tier I national statistics.
- Enterprise area (reported Farm Size) and geographic areas (e.g. the Shape_area field in an ESRI GIS vector layer) often disagree, sometimes substantially.
- Some farms overlap, often for legitimate reasons such as leasing of grazing or cropping areas or perhaps two enterprises belong to the same owners, who treat them collectively for management purposes. However, some overlaps are errors, but there is currently no consistent method to separate legitimate for erroneous overlaps.

Land Cover Database

The description of the land cover database could be slightly improved. In particular, the scale of the data is referenced as 1:50 000 but the better reference would be to the nominal minimum mapping unit of 1 ha. Also, in practice the LCDB includes polygons with sizes smaller than 1 ha, particularly polygons that indicate areas of change.

LINZ Property Database

The description for the LINZ property database states that "there is no land use info associated with this resource". That is not entirely accurate. The LINZ property layer contains some land use information via both the "Parcel Intent" and "Statutory Actions" attributes. The current methods take advantage of both attributes to help map roads, rails, water via hydro parcels, and recreation areas.

In addition, property data usually have land use associated with them via valuation data. If the property information available to Environment Southland also contains that land-use information, it could be brought to bear to help map land use.

When using the LINZ property data, care is required because it relates to parcel and not ownership units. There could be a risk of misclassifying a parcel if the relationships among ownership become lost.

5.3.2 Version 1 Methods Limitations

Below we provided comments on the current methods, first generally in terms of the overall approach and work flow, and second specifically related to particular steps or rules.

Overview

Overall the approach taken seems mirrors that of similar efforts, particularly the approach taken to map land use in Canterbury (Hill et al. 2012). AgriBase™ provides the foundational data, which is then further refined and gap-filled by other data sources.

The current methods involve substantial spatial manipulation of input data sources via erasing, merging, unions, etc. In some cases the specific steps are outlined, in other cases the steps for annealing two different data sources remain implied. Also, given slight spatial discrepancies among data sources, the process of sequential steps often creates polygon slivers. In some cases those slivers are resolved using standard tools (e.g. the eliminate tool in ArcGIS) that provide some consistency in approach. In other cases slivers were resolved "manually" but in some cases no further detail was provided regarding what "manually" meant, e.g. what specific steps were taken, were any formal guidelines or protocols formulated and applied.

Overall we would recommend revising the methods to reduce the likelihood of creating situations that require manual editing. As much as possible the methods should be automated or semi-automated such that the provenance of all resulting polygons can be traced back to their original data sources. Any manual editing that remains should be carefully recorded and documented to ensure consistency going forward, as manual steps can often be the greatest source of variability and inconsistency, especially if undertaken by different operators at different points in time.

An alternative approach to consider would be to undertake a "combinatorial analysis" that first seeks to combine all the input data sources, including only relevant attributes, and then applies a series of rules for processing the combined data layer to produce the final map (Morgan et al. 2010; Price et al. 2010; Rutledge et al. 2011). The combinatorial analysis approach has some advantages and disadvantages. The main advantage is that the approach maintains all input data in a master data set such that the original data could be reconstituted if needed. In other words, no information is lost compared with the current methods. Second, the combinatorial analysis approach can help minimise the occurrence of slivers if implemented properly. Third, the approach can help ensure consistency going forward because all rules are explicitly captured in a documented workflow.

The main disadvantage of the combinatorial analysis approach is its complexity relative to a more traditional sequential work flow that the current methods follow. The complexity arises from the large amount of information, both spatial and attribute, contained in the master dataset. Most current spatial analysis software, however, can handle the complexity.

Finally, our review noted that the methods contain both "steps" and "rules." We could not determine the difference between them and in some cases they seemed interchangeable.

Specific Methods

1. Reclassification of AgriBase ™

Step1: Remove duplicate land parcels

As noted in the methods and earlier in this report, AgriBase™ contains overlapping farms, some of which are legitimate, while others are not. The current methods in the Version 1 report identify overlaps by intersecting AgriBase™ on itself, identifying the most recent record of any overlap, and applying the most recent record to any overlaps or duplicates.

That method has the potential to introduce errors by splitting up different non-contiguous parcels within the same enterprise. We would recommend an alternate approach that first selects the most recent farms based on the "SOURCE_DAT" attribute and saves those as the starting dataset. Rules will need to be developed to resolve any overlaps among farms with the same source date.

Farm enterprises are then added sequentially by decreasing date. If overlaps result from the addition of older farm enterprises, those portions of older farm enterprises that overlap newer farm enterprises, as determined by the source date, would be removed.

During this process, all original attribute data are retained for potential future use. This includes comparing the original geographic farm enterprise area with the final area to monitor which farm enterprises have lost area and how much area they have lost.

Step 2: Urban & Residential

In "All *properties* with a centroid within a TLA or RPMS town boundary…" does "property" refer to an AgriBase™ farm enterprise? If so, how many farm enterprises were reclassified as a result?

Similarly, Rule 2 reclassifies all "properties" less than 0.1 ha as residential. Does that apply to AgriBase™ farm enterprises?

A possible explanation is that the earlier intersect and dissolve separated some parcels that were part of a larger farm enterprise. If that is the case, these reclassification steps could erroneously classify some parcels as residential that are farms or support farming activities even though they occur within town boundaries and/or are very small, as some types of farming have small footprint requirements for individual management or production units, e.g. paddock sizes or minimum field sizes.

Step 3: Specialty land types

Overall, this process seems straightforward, although we are unclear why some farm enterprises types in Rule 3a were retained when they become amalgamated later in the process.

Also as discussed earlier, the "Natural" (NAT) farm enterprise often corresponds to public conservation estate. It would be worth considering removing that farm type entirely and replacing information on conservation and protected areas from other, more appropriate data sources as discussed earlier.

Step 4: Dry stock farm types (Sheep, Beef and Deer)

Numerous comments arose regarding the rules developed, which can be summarised as follows:

- “Dry stock” is now a widely but incorrectly used term to describe livestock types other than dairy. It is a corruption of a stock class (i.e. all non-breeding stock like wethers, dry cows, dry heifers, rams, dry ewes, unmated hoggets, etc.). Farms with breeding herds or flocks (i.e. most sheep and beef farms other than fattening/finishing) are not drystock farms.
- The method assumes that the more detailed attribute data provided by stock units are more accurate than the reported farm type. We can understand the rationale behind this but would recommend rethinking it.
- Why use this method for ‘drystock’ farm types and not all farm types, e.g., arable farms that also have livestock, dairy farms, etc.?
- The method would have to be used with caution for any farm enterprise whose original area was reduced due overlaps.
- When calculating stock units per hectare, was the area used the reported area for the farm enterprise, the geographic area, or the effective grazing area?
- How did you treat highly improbable results when calculating stock units per hectare? Simple calculations in AgriBase™ sometimes yield clearly nonsensical results. e.g. hundreds or thousands of stock units, dairy cows, sheep, etc.. per hectare.
- The rules reference stock numbers (e.g. “Shp_no”), when they should reference the stock units (e.g. “SheepStockUnits”).
- Stock Units
    - The comments about stock class details being unknown did not make sense. The common definition of a stock unit is 1 ewe raising 1 lamb, consuming 550 kg DM/ha/yr.
      See http://portal.beeflambnz.com/tools/benchmarking-tool/definitions
    - Enterprise livestock numbers within AgriBase™ are totals. They include all livestock on the farm at a given time within a given class (e.g. sheep include ewes, wethers, rams, hoggets, 2ths, etc.). As different stock types have different stock units, the resulting calculations may not be accurate.
    - Where does the minimum stock factor of 47.5 come from?
    - Note that OVERSEER uses Revised Stock Units (RSUs).
- The rules could be simplified such that any farm enterprise with DeerStockUnits > 47.5 AND (SheepStockUnit > 47.5 OR BeefStockUnit > 47.5) is classified as Deer and Sheep and/or Beef, rather than creating three separate classes that ultimately get amalgamated.

Step 5: Lifestyle blocks

- Does the *property* in "…*property* between 0.1 and 5.0 ha." refer to an AgriBase™ farm enterprise? If so, this rule likely misclassifies numerous viable farms smaller than 5 hectares in size, especially given that some farm types can range in size well below 5 hectares (Rutledge et al. 2015).

Step 6: AgriBase™ support blocks

- Rule 6a presumably reclassifies all dairy farms in AgriBase™ as Dairy Support. Presumably, many farms are later replaced by the dairy farms represented in the consent database. However, not all dairy farms may be replaced, in which case a legitimate dairy farm may become dairy support.
- Why is the 100 dairy cow threshold important? Is a dairy farm with 100 cows or less not still a dairy farm?
- DRY and GRA are classified as "Support" but support of what? The DRY farm type in AgriBase™ is specifically "Dairy dry stock" so should DRY not be reclassified to Dairy Support?
- The approach described is likely to underestimate Dairy Support, which is only identified as those AgriBase™ farms that claim to be Dairy but are not represented in the consents derived dairy layer (mostly the < 100 ha DAI farms), and DRY and GRA farms (these AgriBase™ categories are not widely used). The only method used to identify winter grazing of animals as a secondary land use is via the AgriBase™ fodd_ha field. A field survey by Environment Southland in 2011 in the mid-Mataura showed winter grazing on non-dairy farms to be quite extensive (Lilburne et al. 2012). Also Monaghan et al. (2010) noted that winter grazing is usually part (2–3 years) of a pasture renewal cycle.
- Another approach would be to use the recently completed mapping of 'Winter livestock forage map – Southland Region 2014' (North & Belliss 2015).

Step 7: Remaining AgriBase™ Sheep, Beef, and Deer farm type

No comments.

2. *Modification of Environment Southland Property Layer*

The preface to this section states that the property layer was used to identify roads, rail, hydro, conservation land, and recreation parcels. However, the subsequent steps outlined related only to the identification of urban/residential areas and recreation areas. General comments on the use of the property layer to map roads, etc., are as follows:

- Using LINZ classifications (e.g. Purpose, Statutory Intent, etc.), visual inspection is strongly recommended, because legal Purpose or Intent does not necessarily correspond to actual land use. It is not uncommon for schools, councils, Māori, etc., to lease land that it is not being used for its designated purpose/intent back to the neighbouring farmer.

- Similarly, with road parcels, paper roads are numerous, many of which are used as farmland and are not captured in AgriBase™. The areas involved are potentially significant.
- The HYDRO purpose parcels are possibly highly inaccurate because rivers channels change frequently. An alternative would be to use the LCDB and/or LINZ topographic information to map rivers, lakes, and ponds.

Step 1: Creation of polygon to fill unknown areas

Generally, we found this section hard to follow because we are not familiar with the details of the particular data source used. We are most familiar with the publically available cadastral data available from the LINZ data source, the basis for some of our comments.

Our general impression is that the base property layer contains all geometries found in the full parcel dataset – primary, secondary, and tertiary – that creates the potential for overlaps. If we want to avoid overlaps when undertaking similar analyses, we use the primary parcel dataset that by design does not contain any overlapping parcels. We wonder if the methods used by Environment Southland could take a similar approach.

Step 2: Urban and Residential Areas

Using the centroid method to determine whether a parcel is inside or outside an urban area is suitable for polygons with regular shapes. Problems can occur with irregular or elongated polygons or with multi-part polygons, as is the case with the unmodified AgriBase™ data. Therefore using a centroid approach can be risky, or in the case of AgriBase™, multipart polygons (e.g. some farms have satellite or support blocks located a distance away from the main parcel). In these cases, the result of using a centroid approach can be risky.

We usually retain the original farm areas, including those outside the catchment boundaries, to model nutrient losses, then clip the results when finished.

Step 3: Recreation areas

No comments

Step 4: Combine with reclassified AgriBase™ layer

In principle there should be a good match between AgriBase™ and LINZ parcel data (the former is based on the latter). It may have been better to add the LINZ parcels at the start.

Joining the reclassified AgriBase™ to the non-erased modified property layer would likely create gaps as well as slivers. Was that the case and, if so, how were gaps filled?

3. Environment Southland and Industry Information

Step 1: Forestry layer

Is this actual forestry blocks (i.e. footprints) or parcels/farms where forestry is the dominant land use?

Was there an attempt to match the forestry boundaries with parcel/farm boundaries? If you update a more accurate boundary into AgriBase™, there will be slivers, etc., and/or legacy.

Step 2: Industry layer

No comments

Step 3: Golf course layer

No comments

Step 4: Environment Southland Dairy layer

Does the consent database contain stock numbers from which stock units per hectare could be calculated?

4. Classification of unknown parcels by LCBD v4

Overall, it would be useful to know what proportion of the resulting land use map came from LCDB v4 as opposed to the other data sources with more explicit land-use information

Step 1: Pasture

We would recommend that all pasture from LCDB gets reclassified as Pasture Unspecified Use (PAS) for consistency. In particular, we did not agree with classifying pasture between 5 and 40 hectares as LIF2, because those areas ultimately got mapped as Lifestyle and likely included many viable farming enterprises.

Step 2: Native

As much native vegetation will likely occur in conservation or protected areas, if these areas are included in updated methods we expect that very little native vegetation will remain. Any remaining areas of native vegetation should be classified as "Unknown." It can then be a simple overlay exercise with the LCDB to determine the proportions of different land cover occurring on Unknown land uses across the region.

Step 3: Exotic forest

As with native vegetation, we recommend areas of exotic vegetation (or weeds) that otherwise do not get captured by other data sources be classified as "Unknown."

Step 4: Arable

Overall, we recommend trying to differentiate the arable land uses derived here into true arable (mixed farming with livestock, dominant crops tend to be grains), annual cropping (very intensive – vegetables, etc.), and perennial crops (grapes, orchards). Differentiating among these types can be important for estimating nutrient loadings.

5. Secondary Land Use/Land Cover

- The 'Forest harvested' is assumed to be replanted as exotic forest. That is probably a fair assumption in Southland but easy to check using recent satellite imagery or aerial photos.
- The approach of relying on the AgriBase™ ara_ha, fodd_ha and veg_ha fields to identify secondary land uses is a problematic as these fields are often empty.
- Rule 10a is not clear. As the SBD, SND, and BND categories already include deer, there seems no need to have deer as a secondary land use.

5.3.3 Validation and Errors

It was not clear why the inclusion of industry data on deer and arable farming was treated as validation data, as opposed to input data like the other industry supplied data sources.

5.3.4 Aerial photography

How was it determined that a property potentially contained errors requiring correction by inspection of aerial photography? What rules or guidelines were followed to decide when/when not to make changes?

As before, were the properties/polygons that changed as a result of comparison with aerial photography tracked in any way?

5.3.5 Sliver Errors

Was the manual correction of sliver errors tracked in any way?

5.3.6 Historic Land Use Map

Our understanding is that the maps have been developed for information purposes only and will not be used for modelling purpose. We therefore reviewed the section on the historic land use map but did not spend substantial on it.

How accurate are the points from the council database? What are the dates?

Provide more description on the process to intersect the LCDB and the council database, which should clearly state the assumption that the regional council database is not changing over time.

We were not convinced that "mixed drystock" can be inferred from the LCDB.

5.4 Objective 3: Recommend Potential Improvements on Current Methods

Below we list our recommendations for the overall principles and for improving the Version 1 land-use classification and methodology, both overall and as fit-for-purpose in terms of helping map sources of potential water quality contaminants. We have numbered the recommendations to make it easier for Environment Southland to track which recommendations they and do not accept.

5.4.1 High Level Principles

Based on the high-level observations, we recommend using three high-level principles that we discussed with Environment Southland at the Objective 5 workshop.

1. Suitability – start by outlining the target land-use classification that meets the intended NPS-FM purposes and needs of associated water quality models, then iteratively develop methods that deliver an agreed realised classification. The Version 1 report provided several incompatible variations on land-use classification that generated both confusion and ambiguity.

2. Repeatability – carefully and rigorously document all methods, processes, decisions, etc. We also strongly recommended avoiding as much as possible any manual editing. In cases where manual editing is necessary, Environment Southland must implement a system to track the evolution of affected polygons. Documentation and data archiving should be robust enough that someone in the future, for example 10 or even 20 years from now, could source the same data and repeat the analysis to reproduce the land-use classification and associated spatial data layers and maps. Ten or 20 years constitute a reasonable period of time for consideration of land-use change, and even represent a shorter period of time relative to some key processes operating on a much longer time scale such as climate change,.

3. Accountability – form an internal Land Use Technical Advisory Group to oversee the process of land-use methodology development including formulating the land-use classification, executing and formally recording decision making, and serving as a conduit between the land-use mapping project team and the full set of internal end-users. The mapping project team, in consultation with the end-users and higher-level management, should develop formal terms of reference for the Advisory Group, including purpose, responsibilities and expectations, frequency of meetings, and processes for considering and deciding different questions or making different choices, etc.

5.4.2 Land-use Classification

1. Introduce the target land-use classification, both primary and secondary, at the start, including a consistent code, category name, and description and reference it consistently throughout the report.

2. Map only land use, rather than having a mixture of land use and land cover.

3. Include an “Unknown” category to identify where robust land-use information is lacking. Over time the goal will be to eliminate any areas of unknown land use.

4. Consider changing Conservation” to “Protected Areas” as the latter category is broader and more suitable for a wider range of purposes than the former. Alternatively, there could be Conservation and Other Protected Areas categories.

5. Consider mapping pig farming as a separate primary land-use category, as such activities can also be significant sources of contaminants.

6. Irrigation appears in Table 1 in the Version 1 report but not thereafter. Irrigation is an important factor to consider for when modelling potential land-use impacts on water resources, including a key point of differentiation for some farming types (e.g. irrigated beef versus non-irrigated beef). If possible, revise the primary land-use classification to include irrigated farms.

7. AgriBase™ is, as noted, a farm-level classification and very open to variations due to operator/preference/bias. However, to move from farm to paddock level analyses is feasible but is a big step and would involve either:

8. Incorporate individual farm maps/plans into the land use mapping data input stream, or regular paddock level analyses from medium resolution imagery. This could follow methods similar to those used by North and Belliss (2015) to map Southland region winter forage in winter 2014.

9. Consider including extensive high country livestock as a land use category, as it is very likely to generate lower contaminant loads compared with some other land-use categories.

5.4.3 Data Sources

1. Source data on conservation and/or protected areas from appropriate sources including the Protected Areas Network (PAN-NZ) database, the Public Conservation Areas Layer via koordinates.com, and/or the two main covenant schemes, QE II National Trust and Ngā Whenua Rāhui.
2. Develop a series of consistency checks to run on AgriBase™ before using it to help minimise errors or misclassifications.
3. Consider using the primary parcel layer (or equivalent) as the starting point for the property level analyses to avoid the need to remove overlaps.
4. Include electoral address points as a data source and adapt the recently developed guidelines for monitoring the effects of land fragmentation to help refine treatment of lifestyle blocks and/or diffuse rural residential development.

5.4.4 Methods

1. Clarify the difference between a step and a rule. Otherwise, make all steps as rules and all rules as sub-rules or similar.
2. Consider using a combinatorial approach outlined in the review section to foster consistency, transparency, and repeatability.
3. Add “Source” and “SourceClass” attributes to each input data source and populate them with the name of the data source and the original land use/land cover category,

respectively. Carry the attributes through the whole method such that the final land-use could be symbolised by the source or source class to demonstrate provenance. Maintaining the link to the source input data will also maintain the ability to map the dates of the various data sources and therefore provide an indication of the vintage of the ultimate source data, which could serve also as a simple measure of uncertainty, e.g. older data are more uncertain.

4. Document all the steps and rules in as much detail as possible, and include an appendix with any GIS rules or commands used for each step and/or rule.
5. Implement a consistent numbering system for all rules and steps to make them easier to reference.
6. A section on the key limitations and assumptions of the land use mapping procedure would be useful. We note there has been no attempt to verify the resulting map – or to assess its accuracy. Some discussion on likely errors in the layer should be made, although we recognise that it is difficult/expensive to obtain an independent dataset of land use for a robust quantification of the error.
7. Rather than identifying urban/residential properties by centroid, consider using an alternative method that may not be as sensitive to parcel shape (e.g., >50% of parcel area is located in a given zone).
8. Develop facilities and methods to record all manual changes.
9. Retain original farm areas, including those outside catchment boundaries or regional boundaries, for analysis and modelling. Clip results when finished.

5.4.5 Other – Land Use Symbology

We recommend applying and publishing for broad use a standard land-use symbology based on a standard colour model such as Red-Green-Blue (RGB) or Cyan-Magenta-Yellow-Black (CMYK) rather than specifying colour names. Colour names are ambiguous and usually system or application specific, and can therefore produce significant variability when implemented by different parties across different systems. Defining a standard symbology based on a standard colour model minimises potential variability. For example, the Land Environments of New Zealand (LENZ) classification (Leathwick et al. 2003a, 2003b) includes a standard set of RGB values for each land environment in each of the four levels of the LENZ classification.

5.5 Objective 4: Recommendations for Potential Ground-truthing Techniques

Ground-truthing can take many different meanings and forms. In the context of developing methods to map land use it means undertaking an independent assessment to gauge the accuracy of the resulting land-use map, in terms of both classification and location.

Classification accuracy asks whether the land use represented was the actual land use on the ground at the time indicated, e.g. was the plantation forest shown on the 2012 land use map really a plantation forest? Location accuracy asks whether the point, lines or polygon boundaries delineated match the corresponding real world conditions, also at the time indicated. The two considerations are intimately related in that errors in classification and location can affect each other.

Ground-truthing of land use can fundamentally occur in one of two ways: direct observation or indirect observation, e.g. via inference. Below we outline common options for both.

5.5.1 Direct Observation

Direct observation means what it implies: an observer goes to a location and documents the observed land use or uses. The documentation could be descriptive, photographic, or both. The main advantage of direct observation is that it theoretically limits or at least reduces the potential for error. On the other hand, it is well known the two or more observers may come to different conclusions, e.g. classifications, unless consistently trained.

The main disadvantages of direct observation relate to the relationship between effort and scale of observation. Directly observing all land uses in Southland would entail a huge effort. To provide a consistent snapshot of land use at one time would require a reasonably large number of observers to travel the entire region during a relative short period of time, i.e. short enough relative to typical rates of land-use change to minimise temporal mismatches.

Most direct observation campaigns, including most monitoring activities, involve robust sub-sampling designed to represent the broader full population under consideration, which in this case means land use.

Given historic rates of land-use change, it seems quite likely Environment Southland could design a robust sub-sampling scheme that includes a large enough sample size to provide a statistically significant result. The details of such a scheme would require further investigation and assessment.

5.5.2 Indirect Observation (Inference)

For practicality, we outline two basic types of indirect observations, both of which rely on inference.

Second-hand Observations

Second-hand observations consist of collecting data from another (human) observer, rather than doing it yourself. Examples include surveys, voluntary self-reporting, mandatory self-reporting (e.g. for a consent), or other mechanisms whereby the observer reports their findings. Conclusions from second-hand observations infer that the observations reported are accurate, i.e. the observer knows what s/he is doing and is telling the truth.

The main advantage of second-hand observations lies in strength of numbers, including the ability to mobilise large numbers of observers to carry out a common purpose, e.g. Landcare Research's Garden Bird Survey[1] or Survey of Rural Decision Makers.[2] Such efforts can be

[1] http://www.landcareresearch.co.nz/science/plants-animals-fungi/animals/birds/garden-bird-surveys

quite powerful if implemented correctly and provide data and analyses that would otherwise not be possible.

The main disadvantages relates to logistics and variability among observers. Some observers may be very knowledge and experienced and more reliable, others less so. Nonetheless, there are ways to design such efforts to help minimise the variability, including simply having enough observers to generate ample statistical power to suit the purpose.

It is worth noting that many data sources used in the current land-use mapping methods, and indeed used commonly throughout New Zealand, utilise second-hand observations, e.g. AgriBase™ is survey based.

Remote Sensing

The advantages of (satellite) remote sensing in a ground-truthing regime are:

- Acquisition of information over very large areas at a single point in time
- Fixed orbits and sensor characteristics in satellite series enable repeatable imaging and semi-automated processing from image to information
- Very cost effective now that a number of medium-resolution services are free-to download.

The Land Cover Database (LCDB), which is already used in the current methods, maps land cover at 5-yearly intervals dating back to 1996 – plus a 1990 baseline for Kyoto forest information, all derived primarily from satellite imagery.

There is also the Land Use and Carbon Analysis System, which consists of 12 land use classes derived from the same original imagery as the LCDB. These databases are produced by semi-automated processing, have a minimum mapping unit of 1 ha and, now that the time series is building up, can be used as a broad-brush mechanism for land cover and land-use-change mapping. However, be aware that the margins of error in these datasets can sometimes be as great as or greater than many small-scale changes in land uses.

Regions often need more detailed land-use mapping – more classes and higher spatial resolutions – and thus it can be better to use LCDB/LUCAS information as a baseline on which to build a more tailored land-use map. For example, LCDB could be used to indicate areas where tussocklands are changing to high producing grasslands and efforts could then be concentrated on the areas in these transition zones.

[2] http://www.landcareresearch.co.nz/science/portfolios/enhancing-policy-effectiveness/srdm

Although many analyses are carried out by computer routines using pre-determined spectral signatures and profiles for the identifications, the best results are obtained when some ground truth "tunes" the mapping to the region and year under investigation.

A representative ground truth data set could be, for example, 100 paddock histories, with information on planting and harvest dates of the major land uses required, along with some information on major treatments (spraying, direct drilling, ploughing, grazing, irrigation). These are then checked against the spectral signature "library", which is adjusted, if required, before the mapping is completed.

Environment Southland could benefit from using remote sensing for ground-truthing in the following possible ways:

- Use existing remote sensing-derived land-use change databases to identify transitioning areas in order to optimise fieldwork or other investigations
- Use free-to-download image sources (for example, Landsat-8 at 30m spatial resolution, Sentinel-2A at 10 and 20 m) to carry out land-use change investigations tailored to the timing and classification needs of the region
- Assuming the Awarua reception facilities planned local downloading of medium-high resolution satellite imagery proceeds, team with Venture Southland to absorb daily/weekly image information updates into the Environment Southland land-use work stream.

5.6 Stage 3: Final Check

At the time of Stage 3, Environment Southland had begun incorporating some of the initial recommendations. Our final check of the Version 2 methods indicated that Environment Southland had already implemented some of our recommendations as outlined below.

5.6.1 High-level Principles

Environment Southland implemented the recommendations related to suitability and repeatability. The methods now provide a clear target land-use classification at the outset of the Version 2 methods. The target classification is clearer and more understandable than the previous version.

Our recommendations regarding repeatability have also been implemented in that the Version 2 report contains detailed methods outlined in the appendix. As indicated earlier, we did not have the resources to double-check the methods in detail.

We cannot comment whether Environment Southland has yet had an opportunity to implement our recommendations on accountability. Nonetheless we strongly encourage Environment Southland to form the recommended Technical Land Use Advisory Group and implement appropriate processes to document that various decision-making processes to help build a solid evidence base for their subsequent analyses and consultations.

5.6.2 Specific Recommendations

Environment Southland has begun to implement many of the specific recommendations that we outlined in our draft December 2015 report. We outline some key recommendations that the Version 2 methods (Pearson & Couldrey 2016) have already implemented:

- Provided a target land-use classification (Table 1) to address suitability
- Reordered the layering of input data including using the LINZ Parcel Layer as the base layer
- Used the 2014 Protected Areas Network (PAN-NZ) database developed by Landcare Research as part of the Envirolink Regional Council Biodiversity Indicators Tool Project (Rutledge and Price 2015) to help map protected and conservation areas
- Identified data sources in more detail (e.g. Plantation Forestry in Figure 4) when feasible or allowed
- Simplified and clarified processing of AgriBase™ data
- Provided several new figures to help illustrate the implementation of various rules
- Expanded discussion of validation, key assumptions and limitations
- Included detailed GIS methods as an Appendix to address repeatability.

6 Conclusions

Overall, the data sources and methods Environment Southland is using to map land use in the Southland region are sound and mirror similar efforts and methods undertaken by other regional councils across New Zealand. Most limitations result, as is usually the case, from deficiencies or limitations in the input data sources, rather than issues with methodologies. Improvements from Version 1 to Version 2 showed substantial progress towards and willingness to improving their land-use mapping approach.

Going forward we expect that Environment Southland will continue to revise and refine both their land-use classification and mapping methods, particularly as more detail on the associated water-quality modelling becomes available to guide further efforts.

When scoping the project, Environment Southland and Landcare Research discussed the possibility of providing guidance on sensitivity analyses. Given that the methods for modelling the relationship between sources and observed levels of contaminants remain under active development, Environment Southland and Landcare Research agreed to defer considerations of sensitivity analysis until the water quality modelling matures.

An upcoming challenge for Environment Southland will be revising the classification and methods, or adapting them, to take advantage of the results of an industry-provided representative farm enterprise survey. Currently the land-use classification and mapping remain very high-level in terms of eventual use for water-quality modelling. Other regions in New Zealand, e.g. Canterbury, developed methods to classify and map farms in more detail, such as more detailed classes of stocking rates (see Table 1). Without knowing the details of what information ultimately become available, based on our discussions with Environment

Southland project staff to date, the resulting dataset will lend itself to development of representative farm types or classes, perhaps by catchment or at least FMUs. The representative farms take one step from a more generic classification scheme, e.g. dairy farm, towards a more ideal or target classification scheme that maps farm systems.

Ultimately, mapping farm systems, rather than even slightly more detailed farm categories or classes, would provide the ideal foundation for robust modelling of land use–water quality relationships. Access to such information would benefit everyone by helping outline potential sources of observed impacts and facilitate common exploration of how and to what degree each player should contribute to achieve common freshwater management outcomes and goals.

7 Acknowledgements

The authors thank Fraser Morgan for his review, Anne Austin for her editorial oversight, and Environment Southland staff for their contributions and time during the project. The report was prepared under Envirolink Medium Advice Grant C09X1515 1645-ESRC274 between the New Zealand Ministry for Business, Innovation and Employment, on behalf of Environment Southland Regional Council, and Landcare Research.

8 References

Anastasiadis S, Kerr S, Arbuckle C, Elliott S, Hadfield, J, Keenan B, McDowell R, Webb T, Williams R 2013. Understanding the practice of water quality modelling, Motu Economic and Public Policy Research Working Paper prepared for the Parliamentary Commission for the Environment. 28 p.

Hill Z, Lilburn L, Elley R, Ford R, Pairman D 2012. Preparation of a GIS-based land use map for the Canterbury region: background report. Environment Canterbury Report No. R12/119. 29 p.

Lilburne L, Moore C, Bidwell V, Burberry L 2012, AquiferSim Modelling of the mid-Mataura Basin. Envirolink Project 1004-ESRC238. Landcare Research Report prepared for Environment Southland. 46 p.

Leathwick J, Wilson G, Rutledge D, Wardle P, Morgan F, Johnston K, McLeod M, Kirkpatrick R 2003a. Land environments of New Zealand – Nga Taiao o Aotearoa. Auckland, David Bateman. 184 p.

Leathwick J, Morgan F, Wilson G, Rutledge D, McLeod M, Johnston K. 2003b. Land environments of New Zealand: a technical guide. Auckland, David Bateman. 237 p.

Lilburne L, Webb T, Robson M, Watkins N 2013. Estimating nitrate-nitrogen leaching rates under rural land uses in Canterbury (updated). Environment Canterbury Report No. R14/19. 45 p.

North H, Belliss H 2014. Winter livestock forage map – Southland Region 2014. Lincoln, Landcare Research.

Morgan F, Rutledge D, Price R 2010. Envirolink tools - Land Use Database Project: gap analysis and prioritisation report. Landcare Research report LC0910/147 for Envirolink and the Regional Council Land Monitoring Forum. 75 p.

Pearson L 2015. Methodology for a GIS-based land use map for Southland, Version 1 (draft). Environment Southland unpublished report. 35 p.

Pearson L, Couldrey M 2016. Methodology for a GIS-based land use map for Southland Technical Report, Version 2 (draft). Environment Southland unpublished report. 71 p.

Price, R, Rutledge DT, Morgan F 2010. New Zealand Land Use Database Environlink Project Draft Database Design Report. Landcare Research Report LC85 for Envirolink and the Regional Council Land Monitoring Forum. 34 p.

Rutledge DT, Briggs C, Lynn I, Price R 2010. Land use trends in Marlborough District: consequences for soil resources. Landcare Research report for Marlborough District Council. 50 pp.

Rutledge DT, Cameron M, Hurkens J, MacDonald G, McBride G, Poot J, Price R, Phyn D, Schmidt J, Tait A, van Delden H, Woods R 2012. Creating futures Waikato integrated scenario explorer technical specifications, Version 1.2. Landcare Research Contract Report LC117 for Environment Waikato Regional Council. 103 p.

Rutledge D, Price R 2015. Regional council biodiversity indicators. Indicator M18: area and type of legal biodiversity protection achieved. Methodology – brief overview. Landcare Research unpublished technical report. 5 p.

Rutledge D, Price R, Briggs C, Cowell S 2009. Geospatial land use classification for New Zealand: Review and recommendations. Statistics New Zealand Official Statistics Research Series Vol. 5. 54 p.

Rutledge D, Price R, Hart G 2015. National guidelines for monitoring and reporting effects of land fragmentation. Landcare Research Report LC2144 prepared for the Regional Council Land Monitoring Forum. 76 p.

Rutledge DT, Price R, Morgan F, Hill R, Taylor A, Cuff J 2011. LANDCLASS: Landscape Classification and Analysis Support System user guide. 25 p.

Appendix 1 – Landcare Research Project Team

Dr Daniel Rutledge (Project Lead): senior scientist and research priority area leader with 20+ years' experience in landscape ecology and spatial analysis including landscape (land use, land cover, land tenure, etc.) classification, characterisation, and change analysis. Key relevant projects that Dr Rutledge led or contributed to include graduate research on geospatial modelling of forest cover distributions, studying long-term trends and changes in land cover and wildlife habitats, the Land Environments of New Zealand classification, review and recommendations for a geospatial land-use classification for New Zealand in collaboration with Statistics New Zealand, develop of the LANDCLASS landscape classification support system in collaboration with the Regional Council Land Monitoring Forum as part of the Envirolink Land-use Database Tools project, development of national guidelines for monitoring and reporting effects of land fragmentation also in collaboration with the Regional Council Land Monitoring Forum as part of an Envirolink Tools project, championing the on-going development and maintenance of the national and comprehensive Protected Areas Network New Zealand (PAN-NZ) database including developing a first generation indicator (M18) on the area and type of legal biodiversity protection as part of the Regional Council Biodiversity Indicators Envirolink Tools project in collaboration with the Regional Council Biodiversity Monitoring Forum , and on-going development of a comprehensive he New Zealand Landscape Database (NZLD) to support research on landscape characterisation, mapping and change, ecosystem services, climate change, food security, etc.

Dr Anne-Gaelle Ausseil: research priority area leader with 15 years' experience in spatial modelling and remote sensing. Key relevant projects include the ecosystem services programme (assessing state and trends of ecosystem services in New Zealand), and the latest Innovative Data Analysis MBIE programme (2014–2018), which she leads. The project aims to support environmental reporting by developing techniques to characterise the provenance, quality and uncertainties of combining data sources in the soil, land use, and biodiversity domains.

Ms Stella Belliss is a remote sensing specialist with over 35 years' experience. She has been involved in researching practical New Zealand applications for optical and radar sensors and systems. Currently she is working on land use and crop type mapping for both Canterbury and Hawkes Bay Regional Councils and recently completed a project to map the winter forage crops in Southland.

Dr Linda Lilburne is a spatial modelling specialist with over 20 years' experience. Relevant projects include land use, nutrient load, and groundwater modelling in the mid-Mataura, estimating catchment nutrient loads for a number of nutrient management zones in Canterbury, using a temporal sequence of images to detect bare ground in the winter, i.e. with the potential for high nitrate losses. She is also involved in the Matrix for Good Management (MGM) project, completing a matrix of nitrate leaching losses under a range of farm systems, climates and soils.

Dr Andrew Manderson is a senior researcher with 12 years' experience in agriculture, soils, GIS, and nutrient management related science and policy development. Relevant projects include land-use change prediction for GHGs, regeneration of national and regional land use methodologies (Innovative Data Analysis project – current), spatial modelling of catchment nutrient loads (Rotorua, Mangatainoka), land use mapping with SPASMO modelling to inform the TRIM freshwater model (Ruataniwha), and currently engaged in the development of new methods for estimating and validating catchment nutrient loads and land use for Freshwater Accounting.

Mr Robbie Price is a senior technician with over 20 years practical experience in land use, land cover, and vegetation research and mapping. His relevant recent work includes the Protected Areas Network New Zealand (PAN-NZ) database, the Envirolink Land-use Database Tool project, and the Vital Sites conservation prioritisation framework, and collaboration on developing threatened environment classifications in New Zealand and Australia. His research also includes conservation and resource management related publications in a range of fields including ecosystems services, impacts of long-term climate fluctuations, and rare species management. His expertise is in spatial data management and analysis, and automation of data processing systems.