Mapping the extent of artificial drainage in New Zealand

Prepared for: Lincoln Agritech

September 2018

Mapping the extent of artificial drainage in New Zealand

Contract Report: LC3325

Andrew Manderson

Manaaki Whenua – Landcare Research

Reviewed by:	*Approved for release by:*
Michael Blaschek Pedometrician Manaaki Whenua – Landcare Research	Linda Lilburne Research Priority Area Leader Manaaki Whenua – Landcare Research

Contents

Summary

Project and Client

- Lincoln Agritech has contracted Manaaki Whenua Landcare Research to develop a GIS layer of *artificially drained land in New Zealand* (NZ).

Objectives

- Develop a national GIS layer that identifies land that is *suitable* for artificial drainage (i.e. land where artificial drainage may be useful or necessary), and then refine according to land that is most *likely* to be artificially drained according to land-use intensity and other factors.
- Validate the result with known observations of artificially drainage.

Scope

- The scope is restricted to farmland artificial drainage types that have the most direct link to surface freshwater bodies through surface channelling (e.g. drains, hump/hollow) or subsurface conduits (e.g. tiles, moles, perforated pipe).
- Scale of analysis is constrained by available national datasets (mostly 1:50,000 scale).

Methods

- Literature review was undertaken to identify soils known or recommended to be suitable for artificial drainage, and to identify soil properties that influence suitability.
- A suitability model has been developed in ESRI ArcGIS 10.3 using fuzzy set theory, whereby semantic statements are constructed to scale relevant datasets (soils, soil properties) into membership functions. These are combined via fuzzy overlay to produce a novel rating of land suitability for artificial drainage.
- Drainage scheme data held by various regional and territorial authorities have been sourced and compiled into a new national dataset of surface drain networks (i.e. drains managed by councils as assets).
- A contemporary land-use intensity map has been developed, representing agricultural intensity at the block level (i.e. within farm).
- A second fuzzy logic model to estimate the likelihood of land being artificially drained has been produced, including the development of scaled relationship functions and fuzzy membership classes for land use intensity and surface drain density.
- A validation dataset has been compiled, drawing from the National Soils Database (NSD), resource consents (regional authorities), aerial photography interpretation (surface drainage), and new observations from a variety of sources. A simple validation is performed by overlay.

Results and conclusions

- A fuzzy inference method was developed for estimating both the suitability of land for artificial drainage, and the likelihood of artificial drainage in New Zealand.

- New input datasets were developed with value beyond this project, including a land intensity layer, a line network of artificial surface drains maintained by councils, a new dataset of land under statutory protections, and a validation point dataset of observed artificial drainage.
- Two national 'suitability for artificial drainage' estimates were produced, based on the mean-area weighted suitability of multiple soil types occurring within single soil polygons, and the maximum value suitability of compound soil units. The former is best used for reporting areas, while the latter is more appropriate for validation purposes. At the 0.5 fuzzy probability threshold, we estimate the area of land suitable for drainage in NZ at 5.4 million ha or 20% of NZ's total land area (mean-area weighted method).
- Two national 'likelihood of being artificially drained' estimates were produced. Likelihood modifies suitability by taking into account current land cover, land-use type and intensity, slope, and the proximity of receiving drain networks (to transfer water away from drained areas).
- We estimate that 2.5 million hectares of land is currently artificially drained at the 0.55 fuzzy probability level (moderate confidence). Confidence classes are provided.
- The estimate achieves a 90% validation accuracy using 8,000 observation points. Accuracy could be improved with more rigorous cleaning of the observations, and a proportion of the inaccuracy arises from scale limitations associated with the soil input data.

Recommendations

We have developed a national GIS layer and map that predicts the current extent of artificially drained land in NZ at different levels of confidence. Our estimate achieves a prediction accuracy of 90%.

We recommend that the artificial drainage layer be made available for use in national and regional modelling applications. In particular, those applications involving contaminant source and flow modelling that haven't yet been able to fully account for the contribution from landscapes with artificial drainage.

We also recommend the continued growth and enhancement of artificial drainage datasets, namely the surface drain dataset, and the point observation (validation) data. Understanding of artificial drainage and its importance to surface water quality will continue to grow, and improvements in these datasets will contribute to better estimates and certainties.

1 Introduction

Artificial drainage is the modification of natural pathways of water flow to improve the removal of excess water from otherwise wet land. Such modifications are increasingly recognised as important pathways for the transport of contaminants from land to freshwater (Stenger et al. 2016; Monaghan et al. 2016).

Historically, little has been done to monitor the installation or expansion of artificial drainage in NZ. Few estimates of the extent of artificially drained land have been made, and those that have tend to focus on soils and soil properties, and areas of land that may benefit from having artificial drainage installed. These are suitability evaluations. To our knowledge, none has yet made an estimate of actual artificial drainage extent – land that is actually drained, rather than an estimate of land that may be suitable for drainage.

Manaaki Whenua – Landcare Research was previously contracted by Lincoln Agritech to develop a proposal for mapping artificial drainage in NZ. The report (Manderson & Belliss 2016) contains much of the background material used in this report. This study focuses on detailing the method and presenting results.

1.1 Objectives

The aim is to develop a national GIS layer that identifies land that is *suitable* for artificial drainage (i.e. land where artificial drainage would be useful or necessary), and then refine the result into land that is *likely* to be artificially drained. The final map will be validated using observations.

2 Method

The method involved three key stages. First, *suitability for artificial drainage* was identified from fuzzy inference applied to soil permeability, drainage class, and expert-recommended soils for drainage. Two suitability outputs were produced (maximum and mean weighted average) to accommodate multiple soils recorded for single polygons. Second, the *likelihood of being artificially drained* is predicted by combining suitability with a drainage likelihood layer developed from land cover, legal protection, surface drainage networks, and a specially constructed land use intensity layer. Last, results are validated from a new spatial dataset of observed drainage. Maps of key input variables used in fuzzy inference are included as Appendix 2.

2.1 Suitability for artificial drainage.

Suitability describes whether artificial drainage would be useful or necessary. For example, it would be neither useful nor necessary to install drainage for a permeable soil that has no evidence of internal drainage restriction. Conversely, a slowly permeable and 'poorly drained' soil would qualify as being highly suitable.

The method for estimating soil suitability for artificial drainage (Fig. 1) is applied to all soils in both the Fundamental Soils Layers (FSLs) and S-Map datasets. The FSLs provides near-national coverage (99%) but is of low quality, while S-Map is of better quality but has limited national coverage (34%, September 2018). For this project, The Soil Map of Stewart Island (Leamy 1974) was manually digitised, attributed, and added to the FSLs dataset to provide national coverage (upgraded FSL coverage = 100%). In the final stages of the process, S-Map results were overwritten into the FSLs to generate a hybrid national soil map.

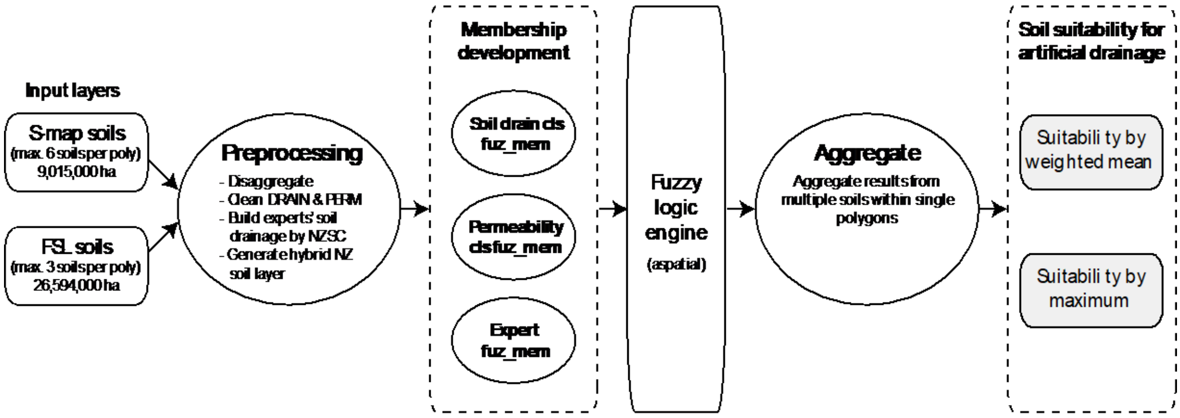


Figure 1 Process of estimating soil suitability for artificial drainage.

Both the FSLs and S-Map contain polygons that record more than one soil. These compound soil codes were disaggregated for analysis. Up to three and six soils per polygon can be recorded in the FSLs and S-Map, respectively. S-map has an attribute describing the proportion (as a percent) of a soil within a given polygon. The proportion for FSLs is assumed at 60:40 for two compound soils (a common ratio used in early soil and land resource surveys), and 60:20:20 for three compound soils.

2.1.1 Fuzzy membership development

Three factors were used to develop inputs for the fuzzy inference of soil suitability for artificial drainage: Soil drainage class, soil permeability class, and expert-recommended soils for drainage.

The *soil drainage class fuzzy membership* is based on *soil drainage class* recorded in both the FSLs and S-Map. Drainage class (Table 1) describes the natural wetness condition of a soil, inferred from the observation of redox segregations and low chroma colours at different depths of the soil profile (Milne et al. 1995). These features are the result of differences in waterlogging and reducing conditions over long periods of time, and thus provide a strong inference for determining artificial drainage suitability.

Table 1 Soil drainage class

Class	Description/interpretation	FSL code	Smap code	Fuzzy value
Very poor	Organic soils, or organically-enriched mineral soils with grey subsoil	1	vp	0.999
Poor	Significant redox features occur up to the base of the topsoil	2	p	0.999
Imperfect	Significant redox features below 40cm depth	3	i	0.5
Moderate	Minor redox features evident within the upper profile	4	mw	0.2
Well	No significant redox features within 80cm depth	5	w	0.05

The small number of classes available limits options to develop a continuously scaled fuzzy relationship. Hence, we assign fuzzy values directly to the classes (Table 1). Very poorly drained and poorly drained soils are highly suitable for artificial drainage, while well drained soils – by definition – do not require artificial drainage and are therefore unsuitable. The greatest uncertainty is with imperfectly drained soils, which may or may not be suitable for artificial drainage depending on other soil properties. Moderately well drained soils are similar, but we assign a slight suitability in recognition that artificial drainage may be a feasible proposition under certain soil conditions and intensive cattle stocking rates.

Soil permeability class describes how freely water moves down through the soil profile (irrespective of drainage class). Permeability is an important indicator of wetness potential for soils that do not have seasonally fluctuating water tables. Under such conditions, soils with rapid permeability are unsuitable for artificial drainage, while slowly permeable soils that generate excess water (e.g. ponding) are more suitable. The *soil permeability fuzzy membership* is based on aggregated classes from the FSLs and S-Map (Table 2).

Table 2 Soil permeability classes

Class	FSL code	Smap code	Description	Fuzzy value
Slow	S	s	slow	0.999
Med/slow	M/S	m/s	slow	0.999
Slow/med	S/M	s/m	slow	0.999
Slow/rapid	S/R	s/r	slow	0.999
Rapid/slow	R/S	r/s	slow	0.999
Medium	M	m	medium	0.5
Rapid/med	R/M	r/m	medium	0.5
Medium/rapid	M/R	m/r	rapid	0.15
Rapid	R	r	rapid	0.15

The development of drainage and permeability fuzzy memberships was broadly calibrated against the drainage and permeability class combinations used by Pearson (2015).

Development of the *expert-recommended fuzzy membership* is based on the identification of soils that experts and commentators have recommended as requiring, or being suitable for, artificial drainage. Recommendations are drawn mostly from literature (see Manderson & Belliss 2016). In addition, we manually extracted drainage recommendations from 170 Soil Information Sheets for the Southland Region.[1]

Recommendations from literature use different soil naming and classification systems. These were unified through standardisation to the NZ Soil Classification (NZSC). The fuzzy membership was developed on the strength of recommendations (e.g. 'requires' vs. 'may respond well' to artificial drainage), and the 'degree of wetness' suggested by NZSC descriptions to the subgroup level as outlined by Hewitt (2010). Fuzzy values for over 250 soil classifications are listed as Appendix 1.

2.1.2 Fuzzy inference modelling

Lookup tables were created and used to assign the previously developed fuzzy membership values to individual soils within polygons for both the FSLs and S-Map. Fuzzy modelling was implemented within a database environment using the same equations used in ESRI's gamma Fuzzy Overlay. Gamma values were tested to ensure an equally calculated midpoint (i.e. where the three input variables have the same input of 0.5, the result will always equal 0.5). The calculations were implemented for each soil within a polygon. A rule was applied to the results to force all poor and very poorly drained soils to a high suitability. This is because – irrespective of any other criteria – poor and very poorly drained soils will always have the highest suitability for artificial drainage.

[1] http://venturesouthland.co.nz/resources/land-use-information/soil-information-sheets

Polygons with multiple soils produced multiple fuzzy probabilities that needed to be aggregated to a single prediction. Aggregating by the **maximum** probability recognises that some part of a polygon will be suitable for artificial drainage, but distorts the suitability of the whole polygon. In contrast, **area weighted average** produces a better overall summary, but can smooth large values (e.g. 50% highly unsuitable + 50% highly suitable produces an indeterminate result). This method may also create problems with validation, when validation data are more detailed than the soils data. A third option was developed, whereby a linear weighting function was developed and applied as a **scaled area weighted average** to emphasis high suitability inclusions (Fig. 2).

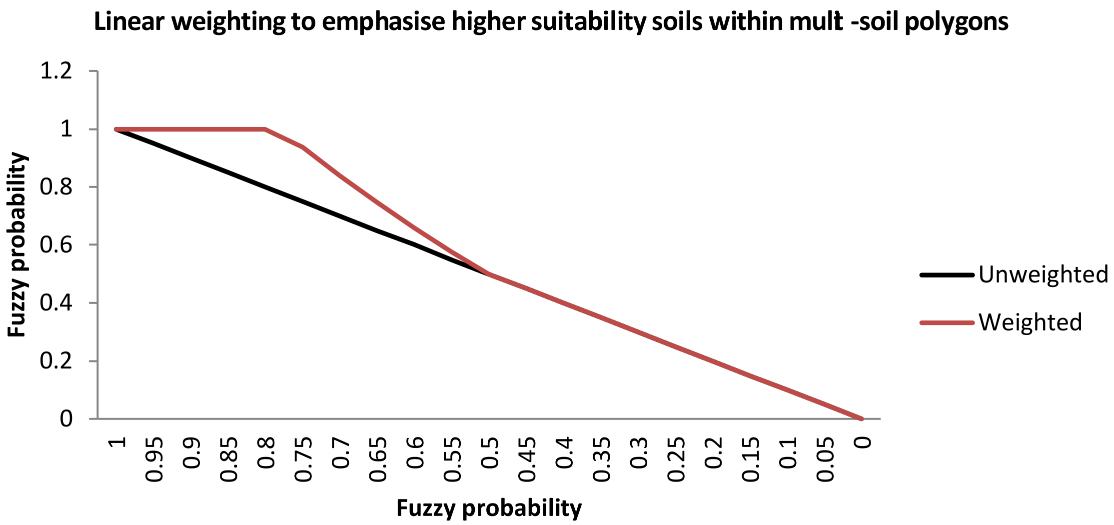


Figure 2 Weighting function applied to the area weighted average to emphasise higher suitability soils in multi-soil polygons.

Other methods of aggregation were considered, including the unweighted average, and a non-aggregation technique of using the first probability (of the dominant soil). An initial implementation check was performed by using S-Map results, comparing the total count of polygons that would qualify according to the different methods (Table 2) (an eligibility of 25% indicates that 25% of S-Map qualifies as being suitable for artificial drainage at the 0.5 probability threshold). Other than the maximum, most methods produced a similar net result, but the pattern of qualifying polygons was different between the techniques.

Table 3 Comparison of different aggregation methods for S-Map multi-soil polygons

Method	Description	Total count	Count of polygons >0.5 prob.	Eligibility (%)
Dominant soil	Probability for the first sibling (i.e. the dominant soil only)	533759	133785	25%
Average	The mean probability of all siblings (unweighted). Does not account for proportion	533759	134284	25%
Maximum	Maximum probability. Recognises that a sibling may require drainage, but ignores the proportion of that sibling.	533759	164528	31%
Area weighted average	Average probability weighted by area	533759	127076	24%
Scaled area weighted average	Scaled/weighted sibling probabilities, then weighted by area	533759	133944	25%

A provisional validation check was also performed using a small set of consent data provided by Environment Southland. The dataset describes 800 km of subsurface tile drains. Convex hull polygons were constructed around clusters of tile drains, and intersected with S-Map and FSL outputs. Land suitable for drainage is reported by percent of the total convex hull footprint (10,370 ha) at different probabilities (Table 4).

Table 4 Percent of land suitable for artificial drainage (within Southland tile drain footprint) at different fuzzy probability levels, for S-Map and the FSLs

Fuzzy value	S-Map results					FSL results	
	Dom soil	Mean	Maximum	Area wgt mean	Scaled area wgt mean	Maximum	Area wgt mean
0.4	77%	89%	90%	88%	89%	60%	59%
0.45	75%	86%	88%	79%	84%	60%	59%
0.5	**75%**	**77%**	**88%**	**77%**	**78%**	60%	59%
0.55	74%	72%	87%	71%	76%	60%	59%
0.6	74%	71%	87%	71%	71%	60%	59%
0.65	74%	68%	86%	69%	71%	60%	58%
0.7	74%	65%	86%	68%	71%	60%	58%

These are positive results, although they do not achieve 100% accuracy. The FSLs produced a notable lower accuracy. A small part of the inaccuracy is explained by the convex hull method. Further, we also encountered a little uncertainty with some of the tile drain data itself (see Section 3.2.2). However, the largest inaccuracies are apparent with soils that do not qualify as being suitable for artificial drainage because they are well drained. This, we believe, is a scale-related error, in that not every small area of different soil can be recorded in a given soil polygon. However, it is occasionally possible that otherwise well-drained soils in Southland do actually have some drainage installed, mainly to capture excess water from localised seeps (Ross Monaghan, pers. comm.).

We opted to include both the maximum and the scaled area weighted average methods (i.e. two suitability layers were produced). Both layers were implemented through the likelihood model to produce two equally valid results, but with applications for different purposes.

2.2 Likelihood of artificial drainage

Suitability describes the potential for artificial drainage according to soils and their properties. However, while a soil may qualify as being highly suitable, other considerations are necessary to determine if it is likely to be drained. For example, a highly suitable soil located in native forest is unlikely to have artificial drainage.

Four key factors are considered in the development of a likelihood classification: land cover, land use intensity, slope, and proximity to surface drain networks.

2.2.1 Land cover

A land cover fuzzy membership was developed using non-agricultural covers from the 2012 Land Cover Database (LCDB 4.2) (Table 5), wetlands from the Freshwater Ecosystems of New Zealand (FENZ) database, non-agricultural covers from the Topo50 dataset, and a specially constructed layer of land under statutory protection.

Table 5 Reclassification of LCDB4 covers

Class	LCDB4 2012 land cover classes	Fuzzy value
Likely	High Producing Exotic Grassland, Low Producing Grassland, Orchard, Vineyard or Other Perennial Crop, Short-rotation Cropland	0.99
Unlikely	Alpine Grass/Herbfield, Broadleaved Indigenous Hardwoods, Built-up Area (settlement), Deciduous Hardwoods, Depleted Grassland, Estuarine Open Water, Exotic Forest, Fernland, Flaxland, Forest – Harvested, Gorse and/or Broom, Gravel or Rock, Herbaceous Freshwater Vegetation, Herbaceous Saline Vegetation, Indigenous Forest, Lake or Pond, Landslide, Mangrove, Mānuka and/or Kānuka, Matagouri or Grey Scrub, Mixed Exotic Shrubland, Permanent Snow and Ice, River, Sand or Gravel, Sub Alpine Shrubland, Surface Mine or Dump, Tall Tussock Grassland, Transport Infrastructure, Urban Parkland/Open Space	0.01

Drainage of wetlands for agricultural purposes is considered as unlikely. While such drainage does occur, it is increasingly rare, as many remaining wetlands are afforded either explicit protection or restricted development through regional plans and policies. For this analysis, current wetlands recorded in the FENZ database are considered unlikely to be drained, and are thus assigned a fuzzy membership value of 0.1.

Non-agricultural covers were also extracted from the NZ Topo50 Landcover Dataset.[2] Features included airports, buildings, canals, cemeteries, gravel pits, ice, lagoon, lake, landfill, mine, mud, pond, pumice pit, quarry, residential areas, shingle, snow, and scree. All are considered to have low likelihood of being artificially drained for farming purposes.

A new layer depicting land under statutory protection was constructed for 2018. This involved sourcing and combining spatial data describing Queen Elizabeth II[3] open space covenants, Ngā Whenua Rāhui protection,[4] and Department of Conservation Public Conservation Areas.[5] Other statutory protection mechanisms were identified by reviewing relevant Acts (e.g. Reserves Act 1977, National Parks Act 1990, Wildlife Act 1953) and extracting qualifying parcels through string query from LINZ datasets. Artificial drainage is considered unlikely for all land managed under a statutory protection.

The resulting land cover was largely binary in character, in that only two primary classes could be constructed (likely or unlikely to be drained). Accordingly, the implication of land cover was added to the analysis as a conditional Boolean statement that overwrites any other fuzzy value (i.e. land under non-agricultural covers will always have a low likelihood of being artificially drained).

2.2.2 Slope

Soils that require artificial drainage are more likely to be encountered on flat land (Hudson et al. 1962; Bowler 1980). However, we found few references for guidelines regarding maximum slope for artificial drainage. We suspect this is because even steep slopes can be drained with the appropriate system of drainage (e.g. contour drains). Further, it is feasible to drain moderately steep slopes by installing tiles or perforated pipe horizontal to the slope itself (David Horne, pers. comm.). Generally, however, we expect that drainage of steep slopes is for the most part avoided to minimise risks associated with blowout and scouring, and because steeper land is often less productive than flat land, and thus less favourable for drainage investment.

Analysis was undertaken to identify potential slope thresholds. First, we identified 54 soils in the National Soils Database that noted artificial drainage as part of the site description. The majority of these soils had a slope of 0°, and the maximum slope recorded was 4°.

Second, tile drain data supplied by Environment Southland was intersected with slope (derived from a 15-m resolution DEM) to calculate the minimum, maximum, and average slope along drain lines. While some very steep slopes were implicated (red crosses in Fig. 3) they tended to be extreme outliers, perhaps explained as data inaccuracies (in the DEM or the tile drain data). Results indicated that artificial drainage is most likely to occur on flat to undulating slopes (88% of tile lines occurred on slopes 0–7°), and least likely for

---

[2] https://data.linz.govt.nz/set/4786-nz-topo50-landcover-data/

[3] https://qeiinationaltrust.org.nz/

[4] https://www.doc.govt.nz/get-involved/funding/nga-whenua-rahui/

[5] https://catalogue.data.govt.nz/dataset/doc-public-conservation-areas

strongly rolling country where slopes approached 20°. In-between slopes (rolling, 7–15°) carry the most uncertainty.

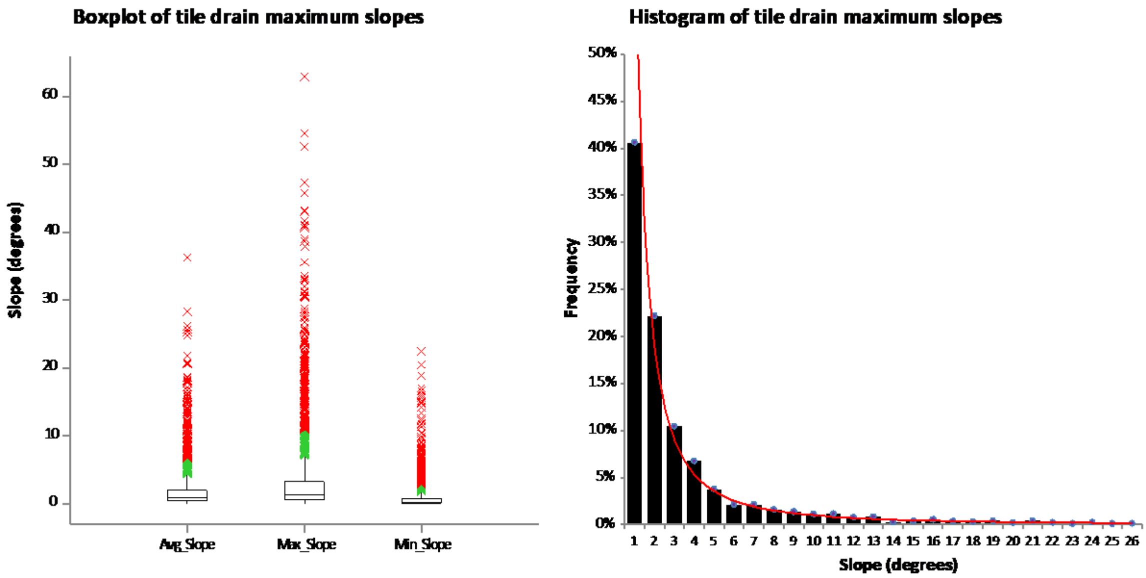


Figure 3 Left: Boxplot of average, maximum, and minimum slopes calculated along Southland tile drain lines (tile drain data supplied by Environment Southland). Right: Frequency distribution (as percent of total count) of maximum slope with a fitted power function (red line).

The frequency distribution modelled well as a power function for direct translation into a fuzzy membership. However, the Southland tile drain dataset is only a small sample of NZ drainage types and drainage types. After consultation with other expertise (Drs D Horne and R Monaghan), the eligibility of slopes <=7° was increased by fitting a sigmoid function. However, initial testing produced fuzzy outputs that were strongly biased by the slope fuzzy membership. Subsequent attempts at weighting also produced erroneous results, so we adopted a scaled transformation method whereby the fuzzy membership function was converted to a weighted scaling factor (Fig. 4). Implementation to the soil suitability layers, results in nil or little transformation to fuzzy values at low slopes, but a strong transformation to lower likelihood values in the 7–20° slope range, and a uniform low likelihood beyond 20° slopes.

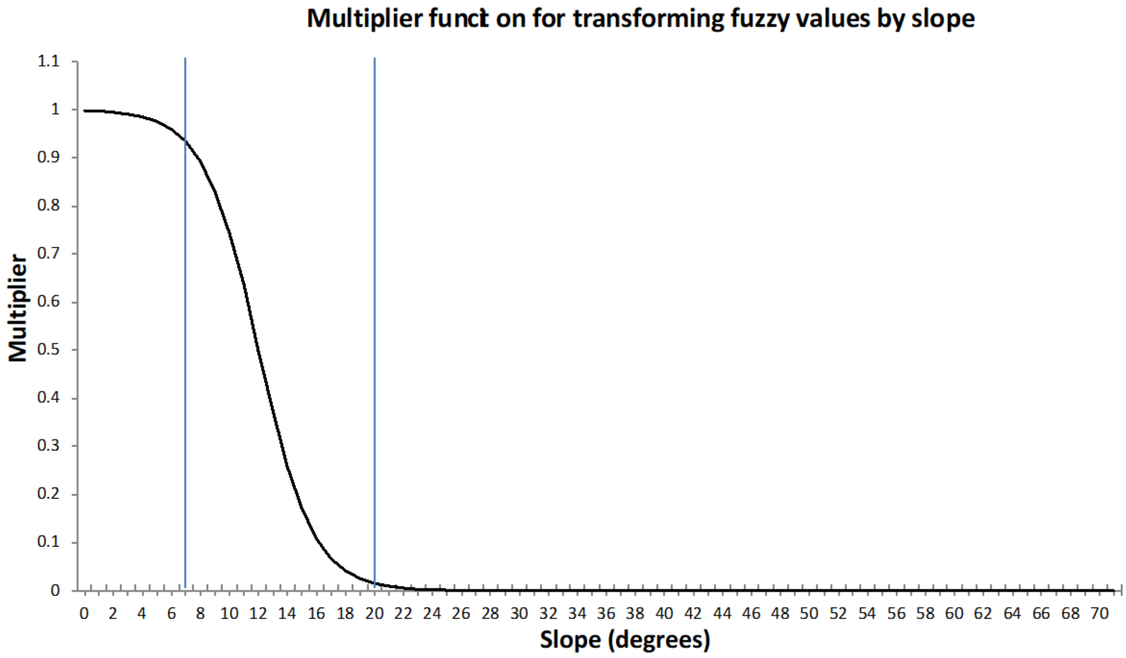


Figure 4 Multiplier function for accommodating the effect of slope on drainage likelihood. Likelihood is slightly decreased as slope increases to 7°, and then rapidly decreased out to a maximum slope of 20° where artificial drainage is least likely.

2.2.3 Distance from drains

Surface drains are either a type of drainage installed in situations where subsurface drains are inappropriate (e.g. draining peat soils), or higher order drain networks that transport water away from drained areas. The occurrence and density of surface drains is a good indicator (i.e. carries high likelihood) that land has been artificially drained.

A new artificial surface drain layer has been compiled, and used to generate a distance from drains adjustment factor. The drain layer includes approximately 16,260 km of network based on the LINZ Topo50 dataset,[6] and a further 8,680 km sourced from ten regional and district authorities (Fig. 5). The LINZ Topo50 drains were cleaned to remove man-made waterways that are least likely to associate with artificial drainage (e.g. canals, water races), while council data are assumed to be of high quality as council drains are managed as assets as part of drainage schemes.

A distance from drains layer was calculated using Euclidean distance at a 15m pixel resolution. We propose that likelihood of artificial drainage is high for farms that contain or neighbour surface drain lines. The average dairy farm size is 185 ha according to the Agribase farm dataset, which broadly translates to a farm length of 1,360 m assuming a basic geometry. Hence, land within 1,360 m is modelled as being most likely. We propose

[6] https://data.linz.govt.nz/layer/50262-nz-drain-centrelines-topo-150k/

a maximum distance of 2,000 m (~400 ha farm), beyond which the surface drain dataset becomes an unreliable predictor for likelihood.

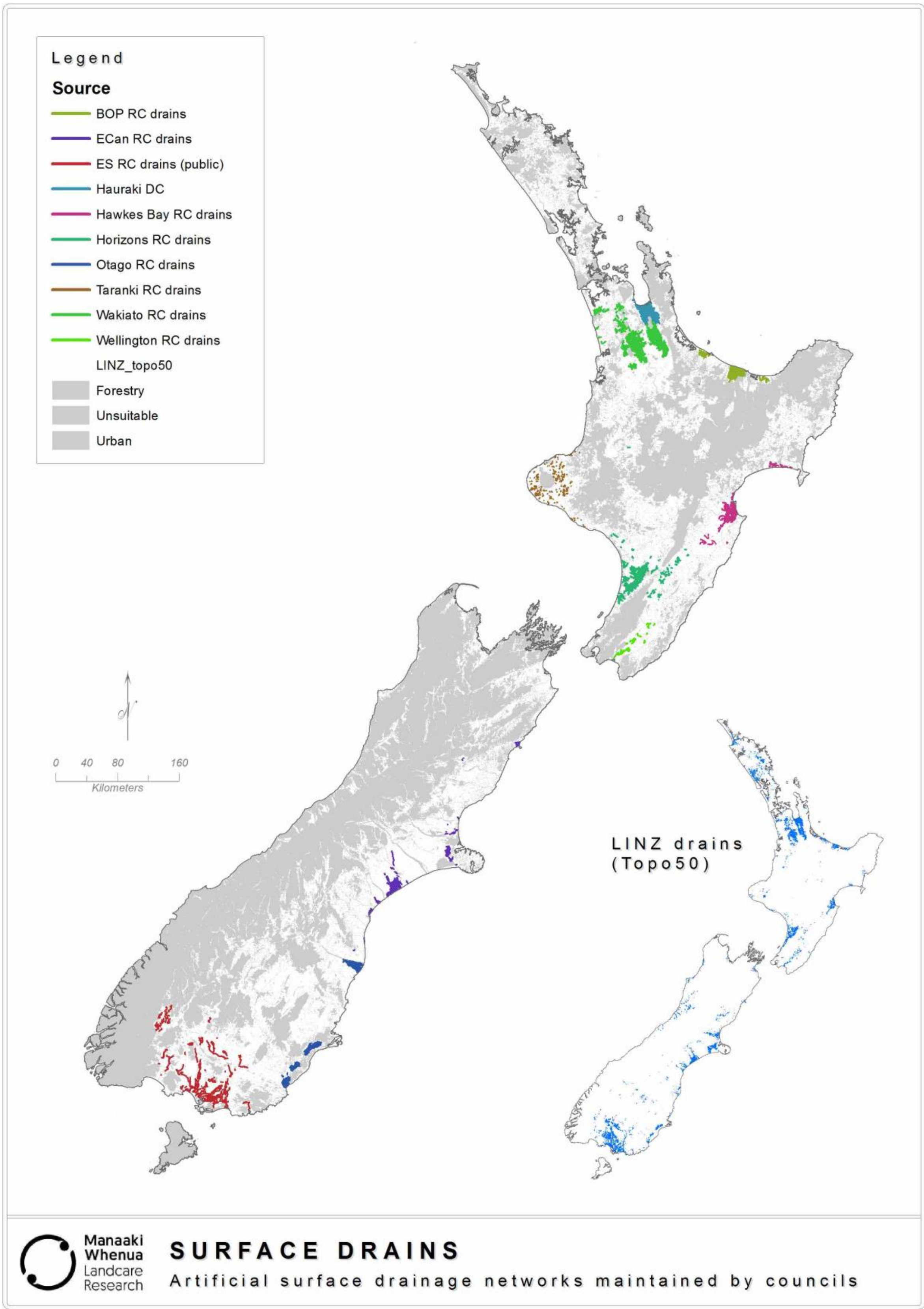


Figure 5 Sources of surface drain network data.

The surface drain layer is an incomplete dataset. Based on comparisons with council and Topo50 data, there are likely drains that have yet to be mapped, or were not able to be obtained from a particular district authority. Any national fuzzy membership generated with incomplete data will produce misleading results for those areas where data are not available.

To accommodate this problem, the distance from drains is used as a fuzzy-membership modifier, whereby values from the interim land use fuzzy membership are adjusted according to their proximity to drains. We use a distance decay function (Fig. 6). In this way, we increase likelihood in areas that have surface drains, but do not change the likelihood for areas were drain data are missing.

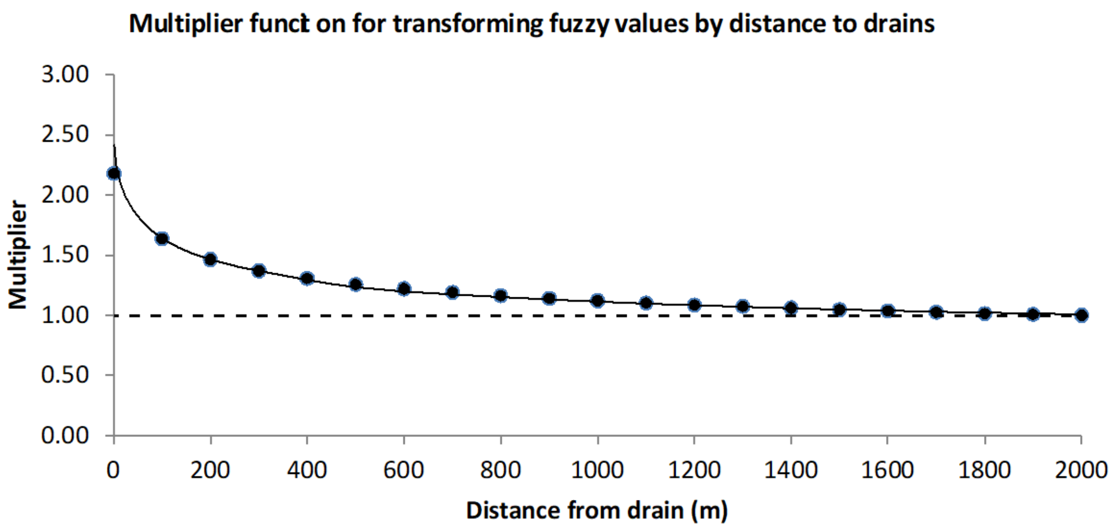


Figure 6 Multiplier function for adjusting fuzzy memberships according to distance from surface drains (fuzzy values >2000m are multiplied by 1 and do not change. A conditional rule is applied to ensure no adjustment results in a fuzzy value >0.99).

2.2.4 Land use intensity fuzzy membership

The likelihood of artificial drainage being present is influenced by land use type and intensity. Artificial drainage of poorly drained soils under high stocking-rate dairy is almost a necessity (i.e. difficult if not impossible to maintain high dairy stocking rates on poorly drained soils without artificial drainage), while at the other extreme, artificial drainage for a low stocked sheep farm with poorly drained soils may only return marginal benefits (i.e. is not as necessary).

The land use intensity fuzzy membership is developed using the Agribase farm dataset. Farm parcels with missing or indeterminate farm type classifications (e.g. NEW, UNS[7]) were manually classed if records contained sufficient enterprise data (e.g. analysis indicates

[7] Agribase farm type codes are listed in Appendix 3

>50% of farm area is in stonefruit so farm type is classed as FRUIT), and through inference of trading or station name (e.g. named as an orchard, deer farm, etc.). Records that could not be classed were deleted and became part of 'missing parcel' infilling.

Missing parcels are selected from LINZ's primary parcels, and neighbouring clusters of parcels are merged to form farms (typically a farm comprises many ownership parcels). Not Otherwise Farmed (NOF) land use parcels such as roads, hydro, and railways, are classed directly from primary parcel attributes. Parcels were then further classed as NOF if parcel area is predominantly protected by legislation (Section 3.2.1) or urban land use dominates (using settlements from the soon to be published 2016 LUCAS Land Use Map). Remaining unclassed parcels are then subjected to infilling by land use signatures and/or dominant neighbouring land use.

Signatures are created by overlaying Agribase farm types with land cover, slope, potential stock carrying capacity, and land use capability, to identify the dominant or average characteristic of a given land use. A similar overlay is performed with the unclassed parcels, and the two layers are matched to class the closest land use. This is complemented with an analysis to identify the dominant neighbouring land use by area (i.e./e.g. if a parcel is surrounded by dairy farms, and the signature suggests dairy land use, then the parcel is most likely to be a dairy farm).

The fuzzy membership is developed by a direct classification of non-pastoral land uses (Table 6), and a sub-classification of dominant pastoral farming types according to stocking rate intensity. Land uses that involve cropping are difficult to rate, as most cropping takes place during the drier months (and is, thus, non-limiting) but this needs to be balanced against wetness interfering or delaying cultivation and harvesting.

Table 6 Initial classification of Agribase farm types by intensity[7]

Likelihood class	Agribase farm types	Fuzzy value
Low	API, DOG, DPL, FIS, FOR, MPL, MTW, NAT, NOF, OPL, OTH, PIG, PKH, POU, RAB, RET, SHW, SLY, TOU, URB, ZOO	0.3
Moderate	ALA, ARA, EMU, FLO, GOA, GRA, HOR, LIF, OAN, OST	0.5
High	FRU, NUR, VEG, VIT	0.7
(pastoral farm types)	BEF, DAI, DEE, DRY, SHP, SNB	-

Key pastoral farms are classified by intensity according to stocking rate levels within each farm type (Fig. 7). Stocking rate is calculated using effective farm area from the enterprise data (if available and logical) or from effective area calculated via overlay with LCDB agricultural covers. Agribase **total stock** numbers are disaggregated to **stock class** numbers using percent make-up of Territorial Authority (TA) livestock numbers reported in the Agricultural Production Survey. Missing and outlier values were then replaced using stocking rate averages by TA and farm type. Intensity classifications are assigned by quartile (Fig. 7), and used as a basis for fuzzy membership values. Initial testing indicated these values were too high and needed to be weighted (Table 7).

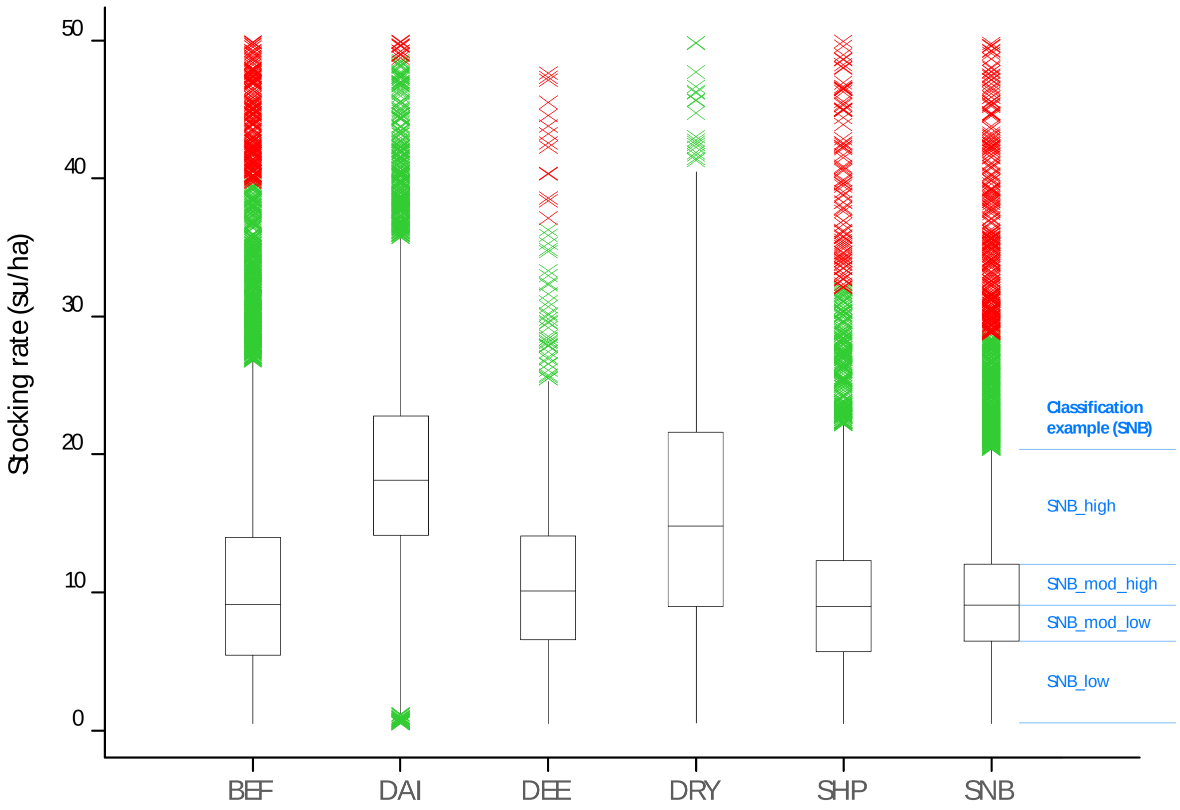


Figure 7 Boxplot of stocking rates calculated for key Agribase pastoral farming types, and example showing the method of classifying stocking rate by intensity (SNB).

Table 7 Likelihood of artificial drainage matrix for pastoral farm types by stocking intensity

	Low	Mod low	Mod high	High
DAI	0.61	0.65	0.7	0.75
DRY	0.50	0.5	0.61	0.65
BEF	0.50	0.5	0.61	0.65
SNB	0.45	0.50	0.56	0.63
DEE	0.40	0.5	0.50	0.61
SHP	0.40	0.5	0.50	0.61

2.3 Validation dataset development

A validation dataset of locations where artificial drainage has been observed has been constructed. This is the first attempt to compile such a dataset, and could be considerably improved if more resource had been available. The primary sources of observations include:

- National Soil Database points where artificial drainage is recorded in the site description. Records range in date from 1967 to 1990. Site coordinates would have been taken from topographical maps and thus may have a degree of location error. An initial 57 possible sites were cleaned down to 50.
- Two study site locations used in the Transfer Pathways Programme (TPP).
- Resource consent locations for earthworks relating to drainage by 'humping and hollowing' or flipping supplied by West Coast Regional Council. Dataset includes 240 points. It is not clear if point locations are for the main farm entrance (common for resource consents) or the actual areas of drained land.
- Resource consent data for tile drains supplied by Environment Southland. Drain lines are converted to points (one point per line, located at the line midpoint). 3,365 points.
- A new dataset of sub-farm surface drain networks for the Hauraki Plains digitised off hill shaded lidar is converted from lines to points to provide 3,909 points.
- A small dataset of observations made by the author mostly in the Wellington Region as part of soil mapping field work (55 observations).
- Tile and mole plan for Aorangi Research Farm (AgResearch) in the Manawatu. Lines converted to points to provide 386 sites.

A total of 8,000 observation points has been assembled. However, the points are highly clustered according to dataset source, and will thus produce location bias in the validation results. A map of the validation point locations is provided in Appendix 2.

2.3.1 Final modelling

All fuzzy memberships are converted to raster formats at 15m pixel resolution for final modelling. The model itself (Fig. 8) applies adjustments to land use and soil suitability input layers, which are then combined via ESRI's gamma fuzzy overlay. The model is applied to both methods of suitability aggregation (maximum value and scale weighted-area average) to produce two outputs. The model generates fuzzy probability values ranging from 0 (definitely not a member) to 1 (definitely a member).

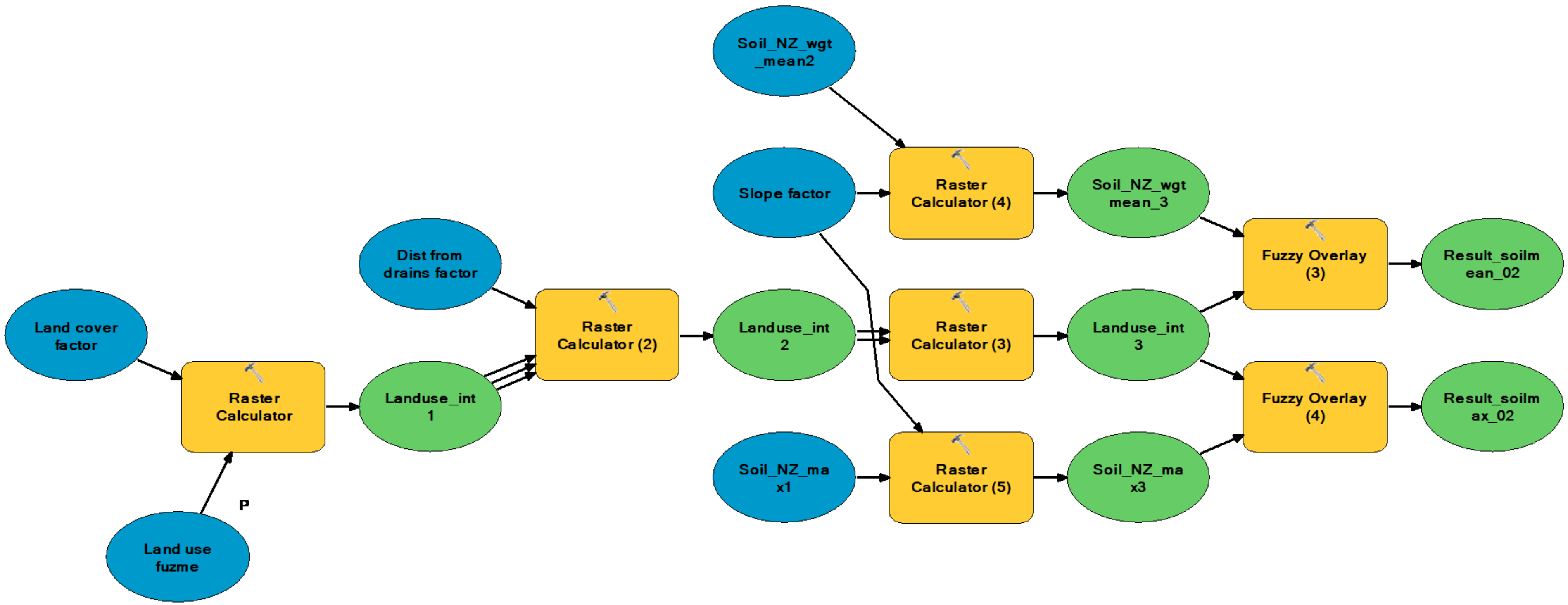


Figure 8 Overview of the fuzzy inference model.

3 Results

Results for the two outputs are tabulated and mapped by fuzzy probability increments that cluster around the midpoint (Table 8; Fig. 9). In the interests of simplification, values <0.5 are proposed as the least likely to be drained, while values >0.6 are rated as having a high likelihood. The midpoint of 0.55 is set as the absolute threshold for generating a single classification for discussion purposes. However, the full range of fuzzy probabilities (Fig. 9) should be used for modelling and analyses that require an estimate of uncertainty.

Results predicted using the “scaled weighted-area average method for suitability” provide a more realistic estimate of total areas, as the maximum value method reports total areas of polygons but in many cases only a proportion of the polygon will qualify as having high suitability for artificial drainage. The “maximum value” method has more relevance for validation purposes, as validation points will likely represent only small confine areas within the larger polygons used to determine the suitability layers.

An estimate of 2.5 million ha of artificially drained land at the 0.55 threshold level is higher than an earlier estimate of 2.0 million ha for cultivatable land (Bowler 1973), but compares well to the 2.7 million ha estimated for ‘arable land’[8] made using NZ’s first spatial land resources database in the early 1980s (Fletcher 1982).

Table 8 Hectares and percent of NZ land area under artificial drainage estimated at different levels of fuzzy probability (fuzziness), for the two methods of aggregating soil suitability (maximum value and scaled weighted-area average)

	Soil max			**Soil mean**	
fProb	**Hectares**	**%NZ**	**fProb**	**Hectares**	**%NZ**
45	3479068	13%	45	3149552	12%
50	**3192703**	**12%**	**50**	**2834619**	**11%**
55	**2914241**	**11%**	**55**	**2528708**	**9%**
60	**2574397**	**10%**	**60**	**2220358**	**8%**
65	2276090	9%	65	1966285	7%
70	1997898	7%	70	1673349	6%

[8] ‘Arable land’ is defined as Land Use Capability class 4 or better, with slopes <=15°. This is similar to the slope factor decay function used in this analysis (Section 3.2.2).

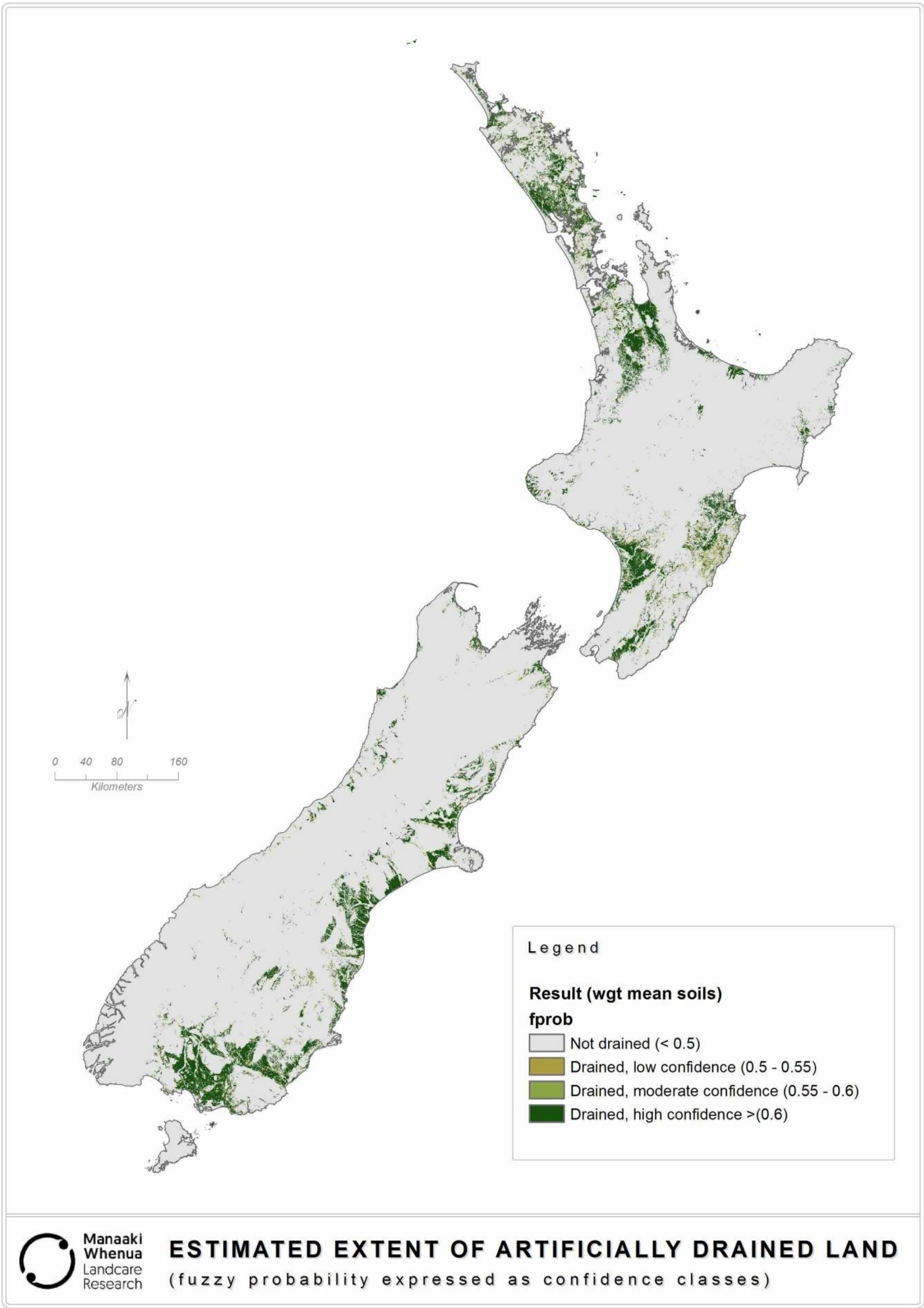


Figure 9 The estimated extent of artificially drained land in NZ.

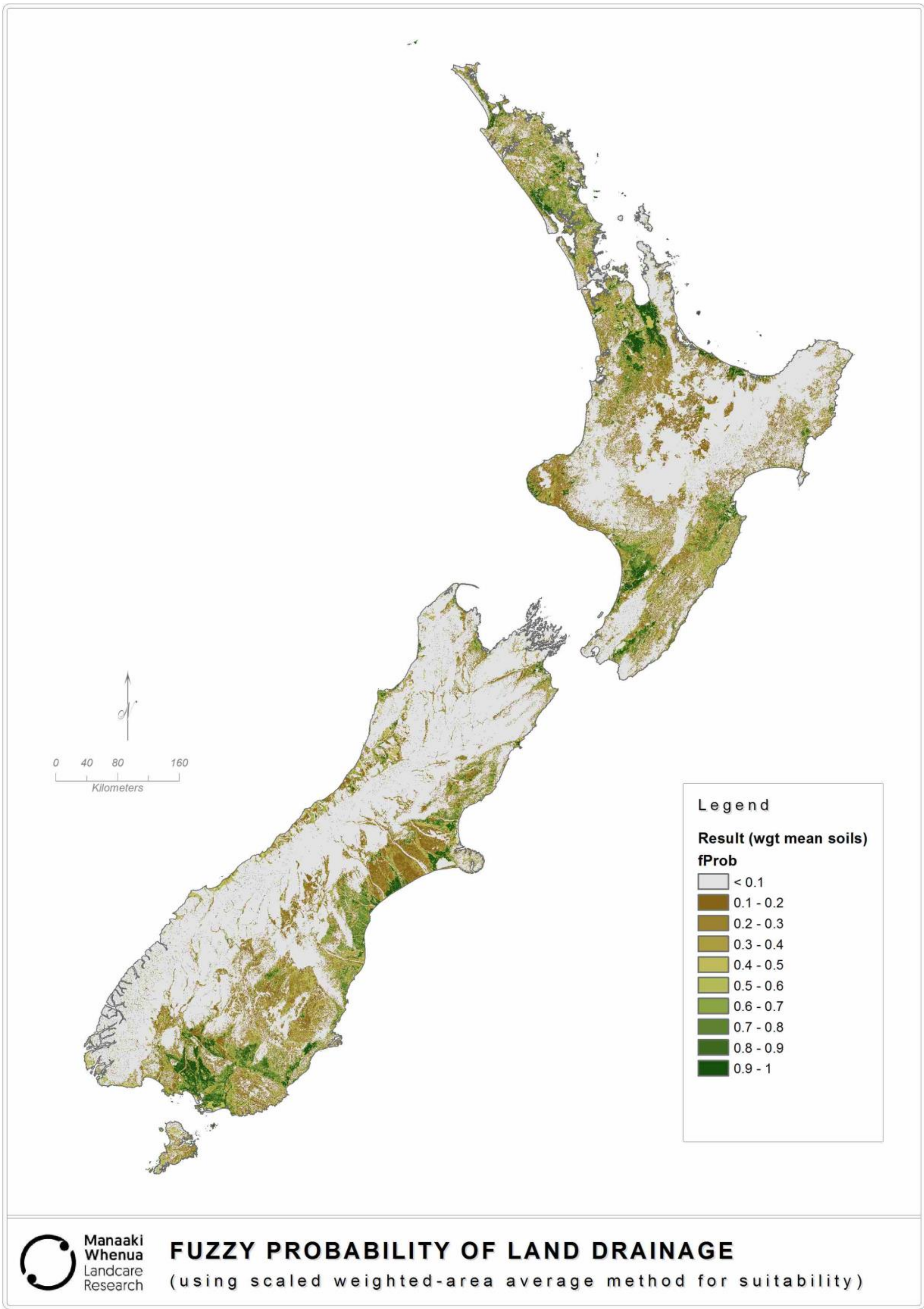


Figure 10 Fuzzy probability of land being artificially drained.

3.1 Validation

The model validates favourably. The total percentage of observed versus predicted drainage is 90% for the estimated based on the maximum value suitability method (Table 9), and 87% for the scaled weighted-area average method (Table 10). A small but notable percentage in the 0.4–0.5 probability range falls just outside our proposed confidence thresholds, suggesting our threshold values could be refined. For individual datasets, both the NSD and West Coast consents achieve the lowest rates of agreement.

Table 9 Percent of validation points that coincide with land predicted to be artificially drained (maximum value method)

		Confidence class				
Dataset	**# points**	**High (>0.6)**	**Moderate (0.55-0.6)**	**Low (0.5-0.55)**	**Very low (undrained) (0.4-0.5)**	**Undrained (<0.4)**
Aorangi	386	96%	2%			2%
ES tile drains	3365	86%	3%	1%	3%	7%
Hauraki	3909	87%	2%		5%	5%
NSD	50	76%	4%		16%	4%
TPP	2	100%				
Wellington	55	96%			2%	2%
West Coast consents	240	56%	5%		20%	19%
All datasets (total)	**8007**	**87%**	**2%**	**1%**	5%	6%

Table 10 Percent of validation points that coincide with land predicted to be artificially drained (scaled weighted-area average method)

		Confidence class				
Dataset	**#points**	**High (>0.6)**	**Moderate (0.55-0.6)**	**Low (0.5-0.55)**	**Very low (undrained) (0.4-0.5)**	**Undrained (<0.4)**
Aorangi	386	63%	21%	13%	1%	2%
ES tile drains	3365	82%	3%	3%	4%	8%
Hauraki	3909	87%	2%	1%	5%	5%
NSD	50	80%	0%		16%	4%
TPP	2	100%	0%			
Wellington	55	91%	5%			4%
West Coast consents	240	52%	7%	1%	17%	23%
All datasets (total)	**8007**	**83%**	**4%**	**2%**	5%	7%

Reasons for lower agreement were investigated by extracting values for each input layer to individual validation points, and then averaged by confidence class to determine which inputs were 'dragging down' the final fuzzy probability estimates (Table 11).

Table 11 Average values for model input layers by confidence class

Confidence class	# points	Average values for model input layers					
		Slope	Distance	Land use	Land cover	Wgt soil	Max soil
High (>0.6)	6610	1.00	1.55	0.50	1.00	0.91	0.94
Moderate (0.55–0.6)	291	0.99	1.36	0.44	0.98	0.50	0.64
Low (0.5–0.55)	187	0.97	1.23	0.51	0.97	0.40	0.54
Very low (undrained) (0.4–0.5)	537	0.98	1.47	0.47	**0.57**	0.56	0.61
Undrained (<0.4)	382	0.96	1.35	0.47	**0.86**	**0.14**	**0.19**

Two key discrepancies are apparent for the undrained classes:

- Low values for land cover suggest that a significant proportion of points coincide with non-agricultural land covers (agricultural covers should have a value near 0.99). This is especially true with the NSD and West Coast datasets. The former contains dated records (back to 1967) and land cover has changed since this time. The West Coast is more difficult to untangle because of uncertainty regarding point placement (farm entrance vs drained area) but a large number of points coincide with native vegetation.
- Very low soil-suitability values in the Undrained (<0.4) class suggest the majority of these points coincide with soils that have been mapped as having minor or nil wetness/drainage impediments. This was confirmed with a visual inspection of a sample, with a high proportion of the errant validation points coinciding with soils mapped as well drained, and to a lesser extent, moderately well drained. Some of this is likely to be mapping error, but we believe that the major cause is scale-related, in that the scale of the soils data (1:50,000) is far more coarse than the paddock-scale detail of the validation points.

From this we deduce that validation accuracy could be improved by a more rigorous cleaning of the validation dataset, especially with respect to dated records. Further, inaccuracies with soils data had a strong influence on the result, suggesting that the model itself will perform even better with more accurate data. However, this should be balanced with recognition that the validation dataset is opportunist in character, as it is based on what could be sourced within the project frame.

4 Conclusions

A fuzzy inference method was developed for estimating both the suitability of land for artificial drainage, and the likelihood of artificial drainage in New Zealand.

- New input datasets were developed with value beyond this project, including a land intensity layer, a line network of artificial surface drains maintained by councils, a new dataset of land under statutory protections, and a validation point dataset of observed artificial drainage.
- Two national 'suitability for artificial drainage' estimates were produced, based on the *mean-area weighted* suitability of multiple soil types occurring within single soil polygons, and the *maximum value* suitability of compound soil units. The former is best used for reporting areas, while the latter is more appropriate for validation purposes. At the 0.5 fuzzy probability threshold, we estimate the area of land suitable for drainage in NZ at 5.4 million ha or 20% of NZ's total land area (mean-area weighted method).
- Two national 'likelihood of being artificially drained' estimates were produced. Likelihood modifies suitability by taking into account current land cover, land use type and intensity, slope, and the proximity of receiving drain networks (to transfer water away from drained areas).
- We estimate that 2.5 million hectares of land is currently artificially drained at the 0.55 fuzzy probability level (moderate confidence). Confidence classes are provided.
- The estimate achieves a 90% validation accuracy using 8,000 observation points. Accuracy could be improved with more rigorous cleaning of the observations, and a proportion of the inaccuracy arises from scale limitations associated with the soil input data.

5 Recommendations

We have developed a national GIS layer and map that predicts the current extent of artificially drained land in NZ at different levels of confidence. Our estimate achieves a prediction accuracy of 90%.

We recommend that the artificial drainage layer be made available for use in national and regional modelling applications; in particular, for those applications involving contaminant source and flow modelling that have not yet been able to fully account for the contribution from landscapes with artificial drainage.

We also recommend the continued growth and enhancement of artificial drainage datasets; namely the surface drain dataset, and the point observation (validation) data. Understanding of artificial drainage and its importance to surface water quality will continue to grow, and improvements in these datasets will contribute to better estimates and certainties.

6 Acknowledgements

We would like to acknowledge Lincoln Agritech who funded this work through their MBIE Transfer Pathways programme. We are also grateful to councils and staff who provided drainage scheme data (Hauraki DC, Bay of Plenty RC, Environment Canterbury, Environment Southland, Hawke's Bay RC, Otago RC, Horizons RC, Taranaki RC, Waikato RC, and Wellington RC), along with Drs Dave Horne and Ross Monaghan who provided expert discussion on different aspects of artificial drainage in NZ. We also acknowledge Matt Irwin and Massey University who made available historical drainage plans of the Drainage Extension Service.

7 References

Bowler DG 1973. The drainage problem and research needs. DSIR Information Series 96. Pp. 102–106.

Fletcher JR 1982. Land in New Zealand which could benefit from drainage. In: McAuliffe KW ed. Proceedings of the First National Land Drainage Seminar. Occasional Report No. 4. Palmerston North, New Zealand, Department of Soil Science, Massey University. Pp. 5–12.

Hewitt AE 2010. New Zealand Soil Classification. 3rd edition. Landcare Research Science Series No.1. Lincoln, New Zealand, Manaaki Whenua Press.

Hudson AW, Hopewell HG, Bowler DG, Cross MW. 1962. The draining of farm lands. 2nd edn. Palmerston North, Massey University. 319 p.

Leamy ML 1974. Soils of Stewart Island (Rakiura), New Zealand. Soil Survey Report 22. Lower Hutt, New Zealand, Soil Bureau DSIR.

Manderson A, Belliss S 2016. Proposal for estimating artificial drainage extent in New Zealand. Report prepared for Lincoln Agritech. Maanaki Whenua – Landcare Research Contract Report LC2690.

Milne JDG, Clayden B, Singleton PL, Wilson AD 1995. Soil description handbook. Lincoln, New Zealand, Manaaki Whenua Press. 157 p.

Monaghan RM, Smith LC, Muirhead RW 2016. Pathways of contaminant transfers to water from an artificially-drained soil under intensive grazing by dairy cows. Agriculture, Ecosystems & Environment 220: 76–88.

Pearson L 2015. Artificial subsurface drainage in Southland. Technical report. Invercargill, Environment Southland.

Stenger R, Wilson SR, Barkle GF, Close ME, Woodward SJR, Burbery LF, Pang L, Rekker J, Wöhling Th, Clague JC, McDowell R, Thomas S, Clothier B, Lilburne L, Miller B 2016. Transfer Pathways Programme (TPP) – new research to determine pathway-specific contaminant transfers from the land to water bodies. In: Currie LD, Singh R eds Integrated nutrient and water management for sustainable farming. Occasional Report No. 29. Palmerston North, New Zealand, Fertilizer and Lime Research Centre, Massey University. 6 p.

Appendix 1 – Expert drainage suitability ratings by NZSC code

NZSC	Suitability
SAT	0.09
AFST	0.09
AM	0.05
ATT	0.05
ATX	0.05
BAM	0.6
BAMP	0.6
BAO	0.75
BAP	0.08
BAT	0.08
BFA	0.05
BFAL	0.14
BFC	0.06
BFL	0.05
BFM	0.6
BFMA	0.6
BFMP	0.6
BFP	0.08
BFT	0.09
BFW	0.15
BLA	0.08
BLAD	0.05
BLAM	0.5
BLD	0.05
BLF	0.1
BLM	0.6
BLT	0.06
BLX	0.05
BMA	0.07
BMG	0.05
BMM	0.6
BMT	0.1
BOA	0.06
BOC	0.05
BOH	0.05
BOI	0.06
BOM	0.6
BOMA	0.6
BOP	0.11
BOT	0.08
BOW	0.15
BSA	0.05
BSM	0.6
BSMP	0.6
BSP	0.05
BST	0.08
BXT	0.14
EMG	0.05
EMM	0.6
EMT	0.09
EOC	0.07

NZSC	Suitability
EODC	0.05
EOJ	0.07
EOJC	0.08
EOM	0.6
EOMC	0.6
EOMJ	0.6
EOT	0.06
EPT	0.75
ERO	0.999
ERT	0.13
ERW	0.27
EVM	0.6
EVMC	0.6
EVT	0.13
GAH	0.999
GAO	0.999
GAP	0.999
GAT	0.999
GAY	0.999
GOA	0.99
GOC	0.999
GOE	0.98
GOI	0.999
GOJ	0.999
GOM	0.999
GOO	0.999
GOP	0.999
GOQ	0.999
GOT	0.999
GRA	0.999
GRF	0.999
GRO	0.999
GRQ	0.999
GRT	0.999
GSA	0.999
GSC	0.999
GSO	0.999
GSQ	0.999
GST	0.999
GTO	0.999
GUF	0.999
GUFQ	0.999
GUT	0.999
LGT	0.999
LIM	0.75
LIT	0.32
LOA	0.05
LOM	0.6
LOT	0.06
LOV	0.05
LOVA	0.08

NZSC	Suitability
LPT	0.75
MIM	0.6
MIT	0.05
MIW	0.05
MOBA	0.05
MOBL	0.08
MOBT	0.05
MOI	0.05
MOL	0.05
MOM	0.6
MOT	0.05
MOZ	0.05
MPT	0.87
NOA	0.21
NOM	0.6
NOT	0.14
NPA	0.75
NPT	0.999
NXA	0.05
NXM	0.6
NXT	0.1
OFA	0.999
OFS	0.999
OHA	0.999
OHM	0.99
OMA	0.999
OMM	0.999
PIC	0.15
PID	0.16
PIM	0.6
PIMD	0.6
PIT	0.06
PJA	0.2
PJC	0.15
PJM	0.6
PJMU	0.6
PJMW	0.6
PJN	0.27
PJT	0.13
PJU	0.15
PJW	0.15
PLM	0.6
PLT	0.11
PPC	0.999
PPF	0.999
PPJ	0.999
PPJX	0.999
PPT	0.999
PPU	0.91
PPX	0.999
PUJ	0.88

NZSC	Suitability
PUM	0.6
PUT	0.05
PXJ	0.5
PXJC	0.5
PXJM	0.6
PXJN	0.5
PXM	0.6
PXMC	0.6
PXMJ	0.6
PXT	0.15
RF?	0.05
RFA	0.05
RFAW	0.06
RFM	0.6
RFMA	0.6
RFMQ	0.6
RFMW	0.6
RFT	0.13
RFW	0.08
RHI	0.05
ROA	0.09
ROAW	0.05
ROM	0.6
ROT	0.09
ROW	0.1
RSA	0.05
RSK	0.05
RSM	0.6
RST	0.06
RTB	0.31
RTBA	0.05
RTBL	0.41
RTBP	0.05
RTM	0.6
RTT	0.06
RXT	0.06
SAH	0.05
SAM	0.6
SAW	0.05
SIM	0.6
SIT	0.06
SJK	0.14
SJL	0.05
SJM	0.6
SJQ	0.1
SJT	0.13
SZQ	0.05
UDM	0.71
UDP	0.75
UEM	0.6
UEP	0.75

NZSC	Suitability
UEY	0.5
UPS	0.75
UPT	0.75
UST	0.05
UYM	0.6
UYT	0.22
UYZ	0.15
WF	0.06
WGF	0.999
WGFQ	0.999
WGFU	0.999
WGQ	0.999
WGT	0.999
WH	0.15
WHA	0.05
WO	0.06
WS	0.05
WST	0.05
WT	0.05
WW	0.07
WX	0.15
XNT	0.15
XOT	0.06
XPT	0.6
ZDH	0.99
ZGT	0.999
ZOH	0.06
ZOT	0.21
ZPH	0.79
ZPHP	0.999
ZPOZ	0.999
ZPP	0.87
ZPQ	0.999
ZPT	0.24
ZPU	0.999
ZPZ	0.999
ZXF	0.34
ZXH	0.2
ZXP	0.5
ZXQ	0.15
ZXU	0.54
BOMW	0.6
EPJ	0.75
LGO	0.999
RXA	0.06
ZDYH	0.999
NOL	0.06
ZDQ	0.999
BSA*	0.05
OLO	0.999
ZX	0.5

Appendix 2 – Maps of key inputs

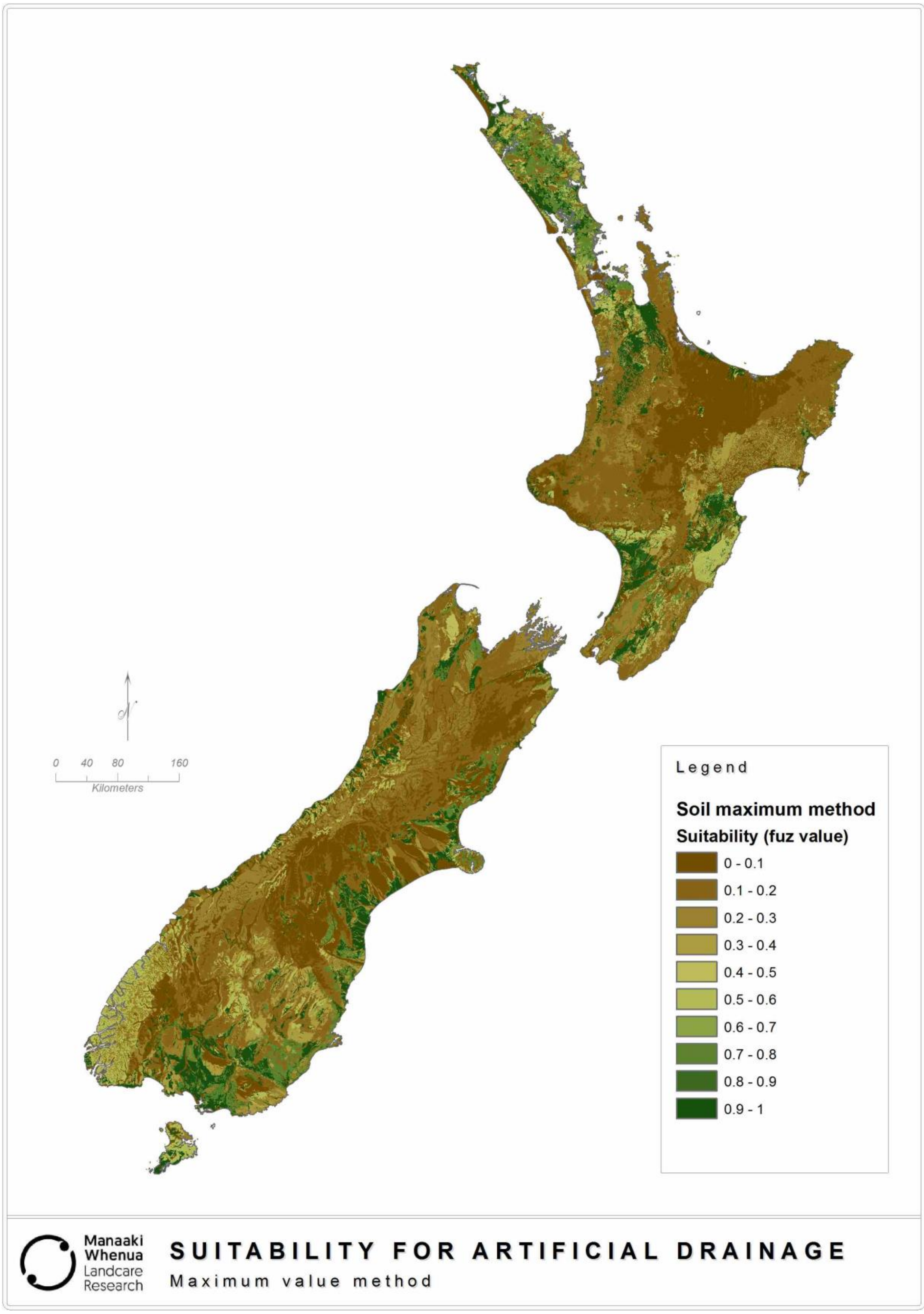
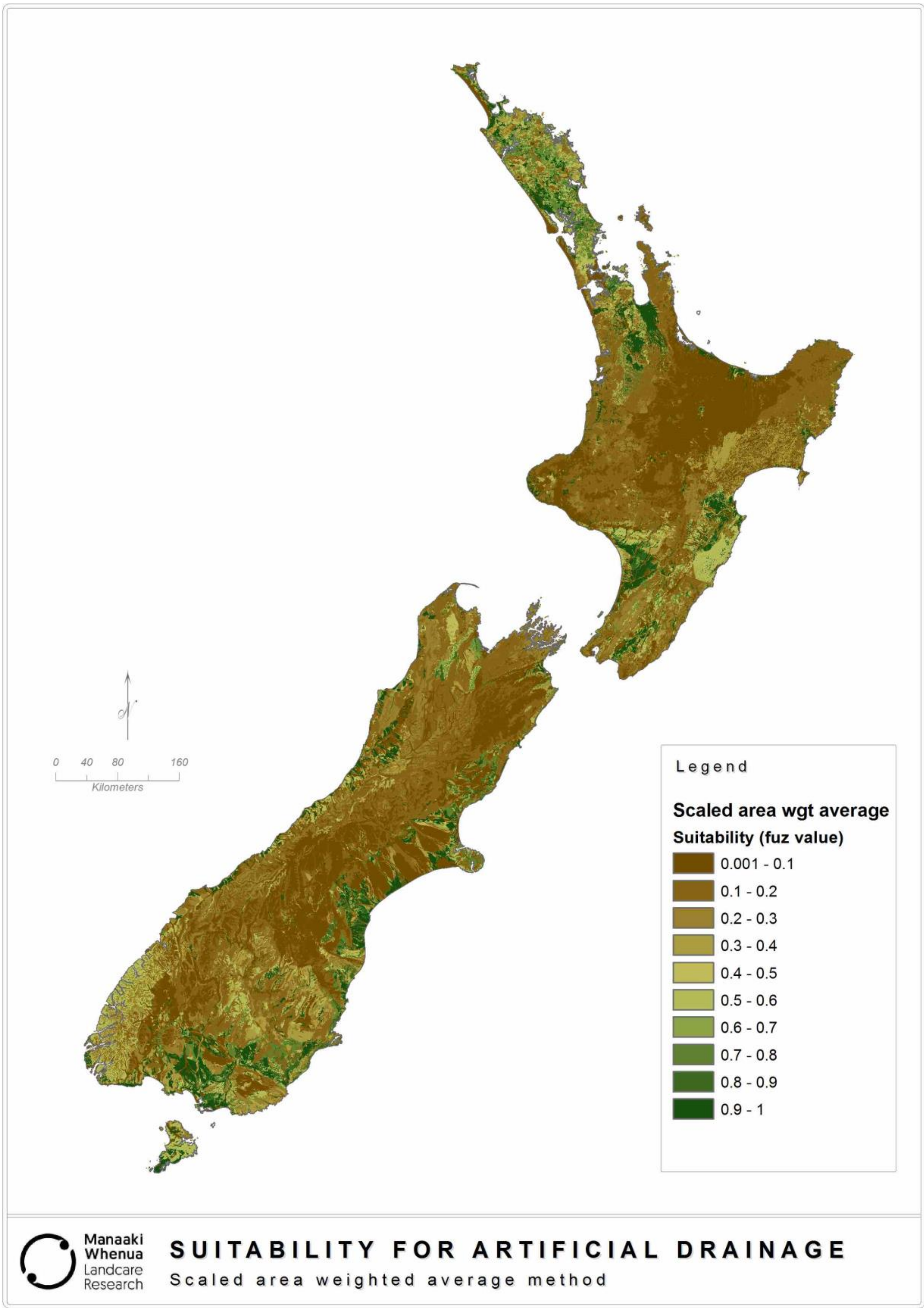


Figure 11 Fuzzy membership layer for artificial drainage suitability (maximum value method).

Figure 12 Fuzzy membership layer for artificial drainage suitability (scaled area-weighted average method).



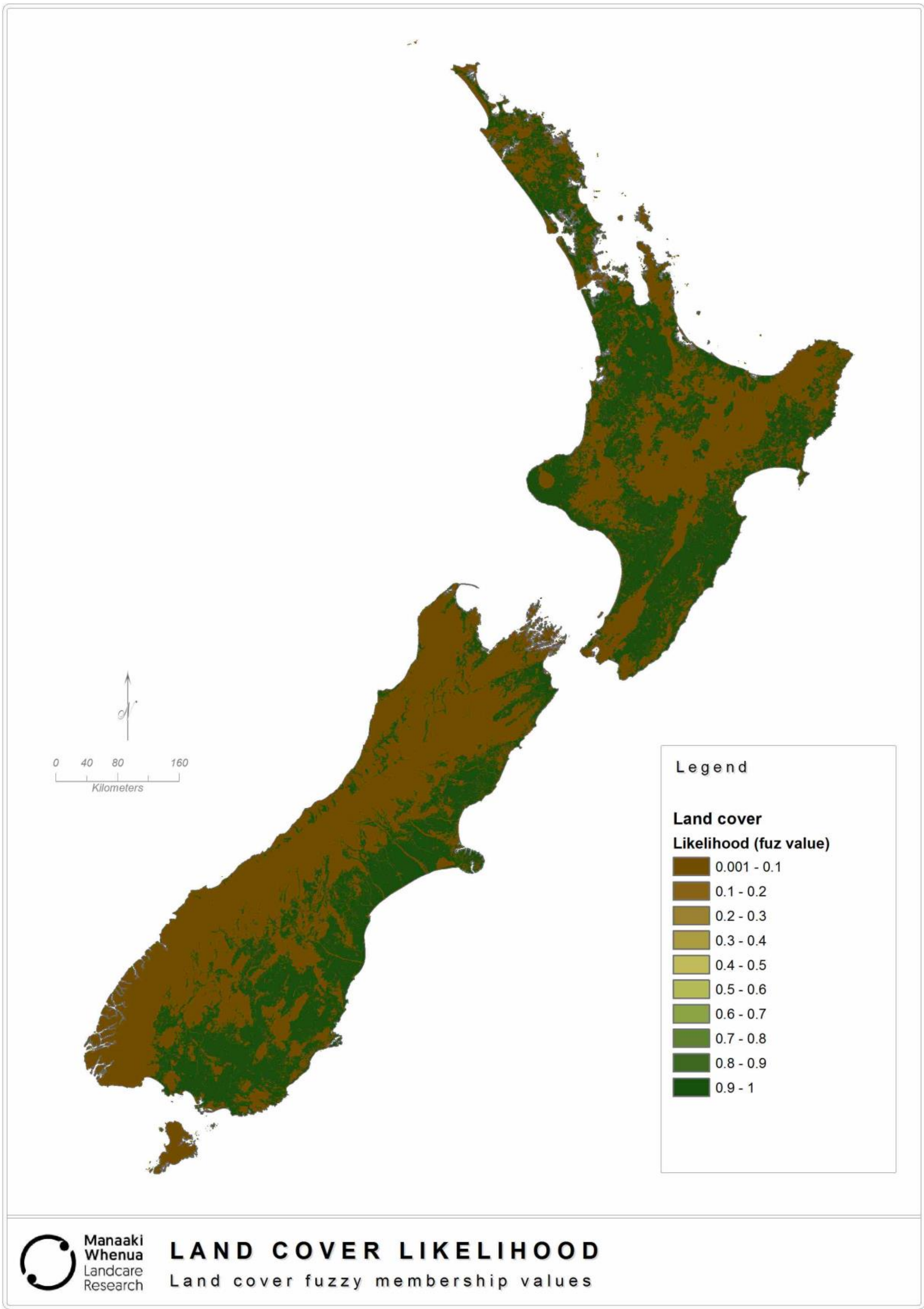


Figure 13 Fuzzy membership layer for land cover.

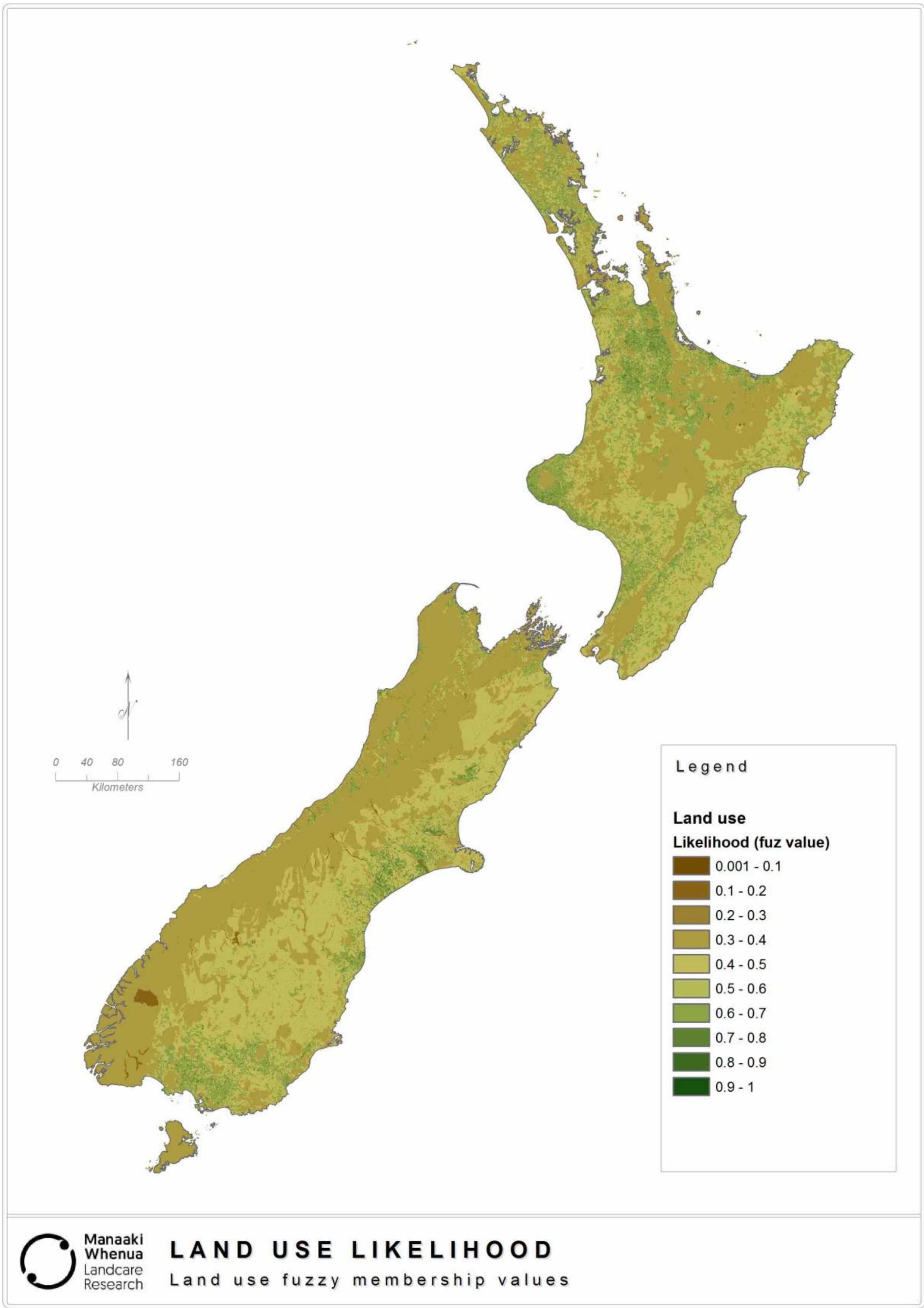


Figure 14 Fuzzy membership layer for land use (weighted).

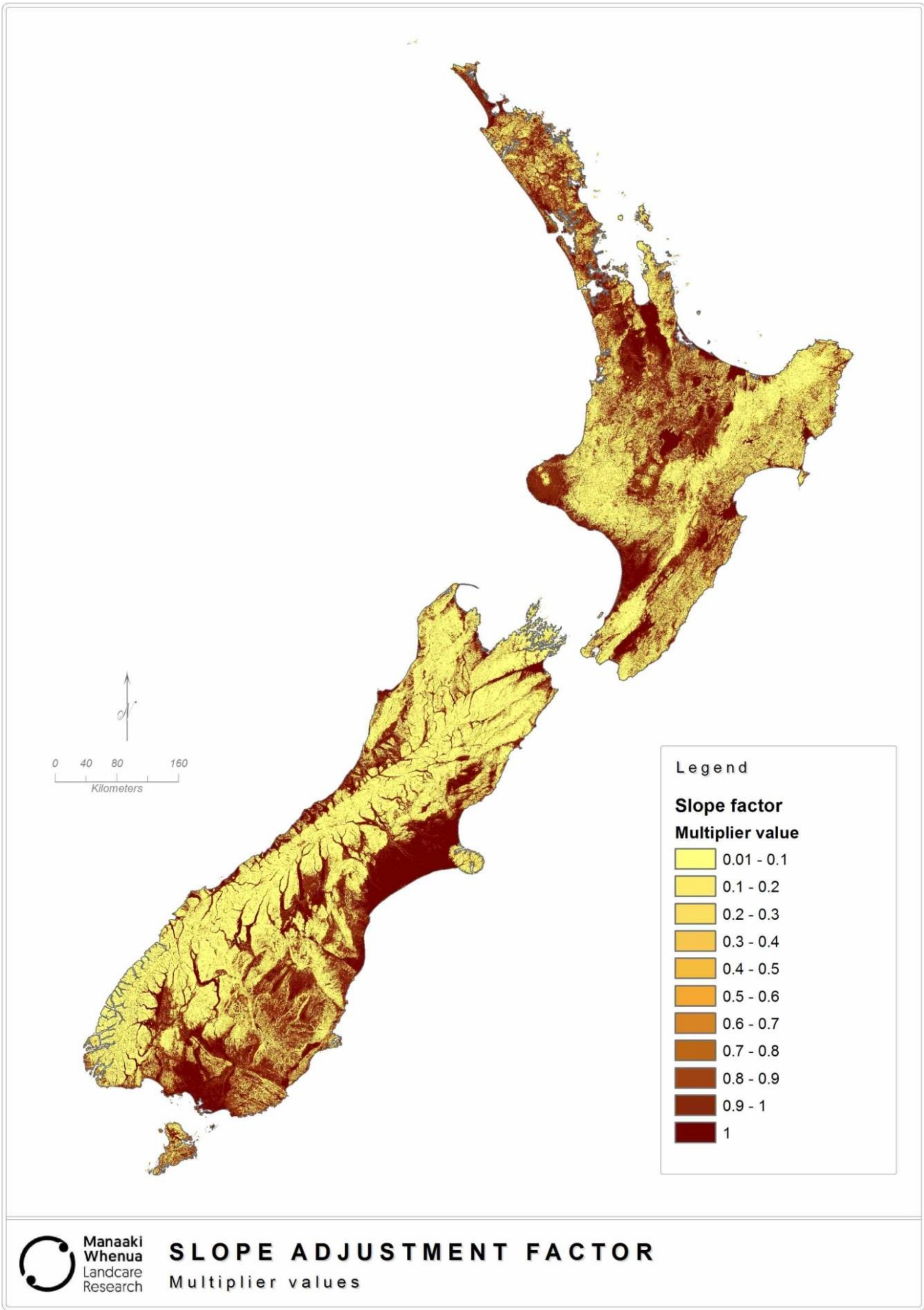


Figure 15 Slope adjustment factor.

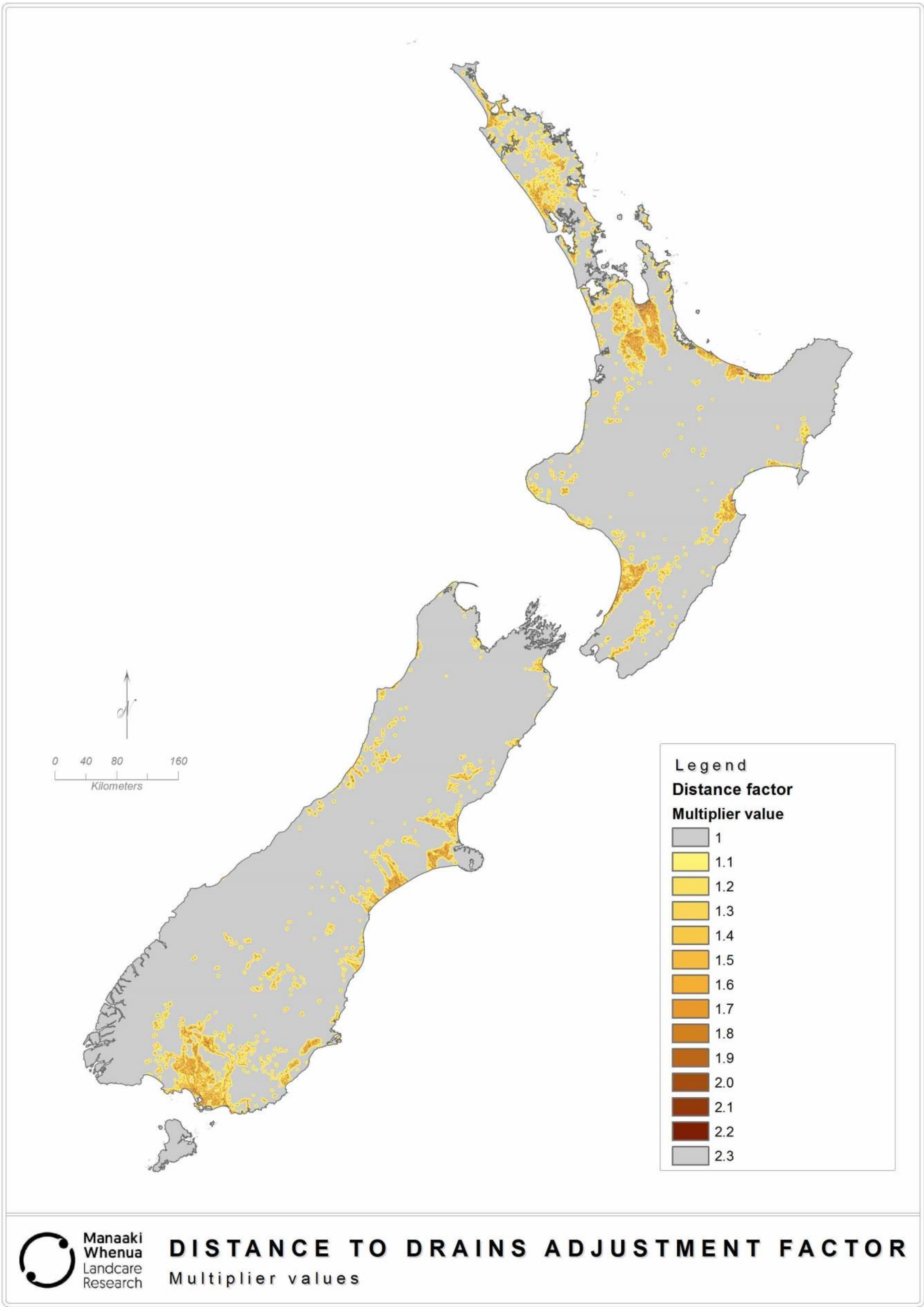


Figure 16 Distance to drains adjustment factor.

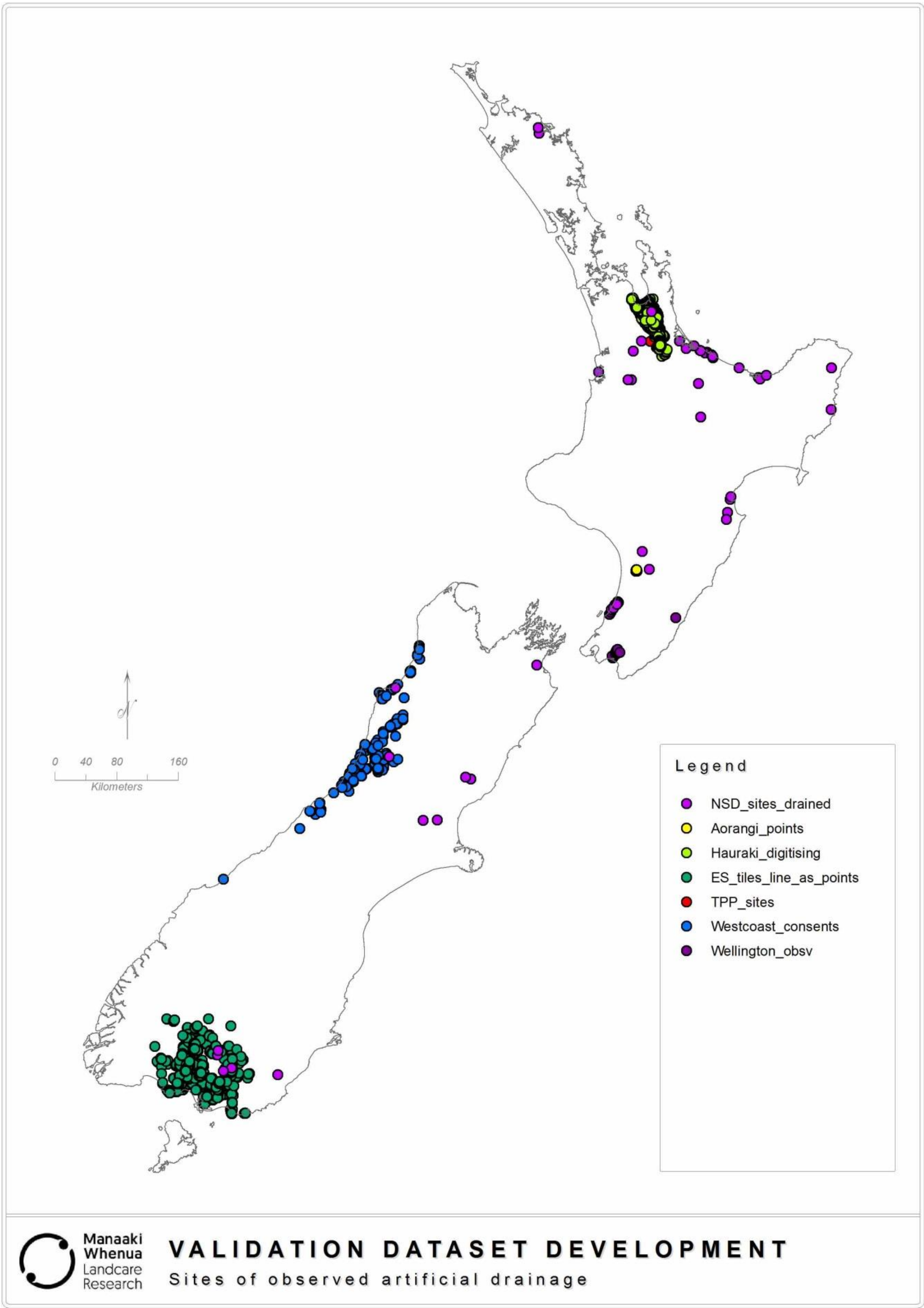


Figure 17 Validation set locations (point dataset).

Appendix 3 – Agribase farm type descriptions

Code	Description	Modifications
ALA	Alpaca and/or Llama Breeding	
API	Beekeeping and hives	
ARA	Arable cropping or seed production	
BEF	Beef cattle farming	
DAI	Dairy cattle farming	
DEE	Deer farming	
DOG	Dogs	
DRY	Dairy dry stock	
EMU	Emu bird farming	
FIS	Fish, Marine fish farming, hatcheries	
FLO	Flowers	
FOR	Forestry	
FRU	Fruit growing	
GOA	Goat farming	
GRA	Grazing other people's stock	
HOR	Horse farming and breeding	
LIF	Lifestyle block	
MTW	Meat slaughter premises	
NAT	Native Bush	
NEW	New Record – Unconfirmed Farm Type	Redefined by enterprise attributes or parcel infilling
NOF	Not farmed (i.e. idle land or non-farm use)	Expanded to include roads, hydro, and protected land
NUR	Plant Nurseries	
OAN	Other livestock (not covered by other types)	
OPL	Other planted types (not covered by other types)	
OST	Ostrich bird farming	
OTH	Enterprises not covered by other classifications	
PIG	Pig farming	
POU	Poultry farming	
SHP	Sheep farming	
SNB	Mixed Sheep and Beef farming	
TOU	Tourism (i.e. camping ground, motel)	
UNS	Unspecified (i.e. farmer did not give indication)	Redefined by enterprise attributes or parcel infilling
VEG	Vegetable growing	
VIT	Viticulture, grape growing and wine	
ZOO	Zoological gardens	