Manaaki Whenua
Landcare Research

Mapping of native Myrtaceae species in New Zealand

Prepared for: Department of Conservation

December 2017

Mapping of native Myrtaceae species in New Zealand

Contract Report: LC3065

Susan K. Wiser, Jerry A. Cooper, Elise A. Arnst, Sarah J. Richardson

Manaaki Whenua – Landcare Research

Reviewed by:	*Approved for release by:*
Peter Bellingham Senior Researcher Manaaki Whenua – Landcare Research	Gary Houliston Portfolio Leader – Enhancing Biodiversity Manaaki Whenua – Landcare Research

Contents

Summary v
1 Introduction 1
2 Background 1
3 Objectives 1
4 Methods 2
4.1 Overview of the native Myrtaceae occurrence data set 2
4.2 The data processing pipeline 3
4.3 Data sources and processing 3
4.4 Estimating sampling effort 10
4.5 Defining abundance classes 10
5 Results 14
6 Conclusions 74
7 Recommendations 75
8 Acknowledgements 75
9 References 75

Summary

Project and client

- In September 2017 the Department of Conservation (DOC) contracted Manaaki Whenua – Landcare Research (MWLR) to produce improved distribution and abundance maps of native Myrtaceae. A first draft of such maps was produced in August 2017 by Manaaki Whenua.

Objectives

- MWLR have already produced PDFs for DOC of 'polygonised' species maps for 19 native Myrtaceae taxa using National Vegetation Survey (NVS) databank plot data, and New Zealand Virtual Herbarium (NZVH) and i-Naturalist Naturewatch occurrence records. We categorised species occurrences (percentage of plots, collections and observations recording the target species) using ecological districts as polygons. Ecological districts are environmentally and biogeographically homogeneous areas, and using them to cautiously extrapolate point location data is robust. The objective of this project is to enhance and improve these maps.

Methods

- First we ensured that occurrence records at the variety or subspecies level were incorporated into the data set at the species level. Note that *Kunzea ericoides* is treated *sensu lato* and does not include recent taxonomically split species. Together with the exclusion of *Metrosideros kermadecensis* (with a natural range restricted to the Kermadec Islands), this reduces the number of mapped taxa to 18 species. Outer islands were excluded from this analysis.
- Records were incorporated for georeferences within 1 km of the coast and assigned to the appropriate ecological district (under the assumption that uncertainty/truncation in the georeferences had inadvertently located these records offshore).
- Numerous records were incorporated from NZVH currently without Lat/Lon data but with NZTM/NZMG[1] coordinates.
- For those taxa restricted to a subset of Land Cover Database (LCDB) classes, the ecological district maps were intersected with the extent of these classes to increase the spatial resolution of display.
- We checked the veracity of the 100 records sourced from the NVS databank, representing the most pronounced outliers based on current knowledge of distributional limits given in the literature resources such as the New Zealand flora manuals, Figure 5.3 of *Vegetation of New Zealand* (Wardle 1991) and the biological floras (Allan 1982; Breitwieser et al. 2010). For examples, see the previously produced

[1] New Zealand Transverse Mercator projection (NZTM) and New Zealand Map grid (NZMG)

maps of *Metrosideros excelsa* (which show presence in the lower North Island) and *Metrosideros robusta* (which show presence in Southland).

- We tested whether the thresholds and/or definitions of the six levels of occurrence could be adjusted. For example, should a minimum number of plots be required in an ecological district to define a taxon as 'common'?
- We considered whether the NZVH and NatureWatch records could be incorporated into the measure of sampling effort.
- BIOWEB was included as an additional data source (threatened species records). Duplicate records with data from other sources were identified and removed as far as possible within the time constraints.
- Known natural species ranges were quantified from the published literature.

Results

- We have produced distribution maps for each taxon, which include a measure of abundance. Distributions are masked by known species ranges and LCDB classes judged suitable for the species.
- We have provided a layer depicting the number of Myrtaceae species per ecological district.
- We have provided additional maps to aid interpretation and to make the steps in the mapping process transparent:
    - i presence/absence maps for each taxon at the ecological district scale, based on all records
    - ii range maps for each taxon at the ecological district scale, based on the literature
    - iii maps of the coverage of the LCDB classes used to mask the primary maps
    - iv a map of the overall sampling effort, showing the number of NVS plots per ecological district.
- We also discuss the limitations of these maps.

Conclusions

- The most narrowly distributed Myrtaceae species on the New Zealand mainland are *Kunzea sinclairii* and *Metrosideros bartlettii*. The most broadly distributed is *Leptospermum scoparium*.
- Nine species are designated as common in at least one ecological district.
- Of the 18 Myrtaceae species examined, the most species that co-occur in a single ecological district is 14 (two ecological districts). Six ecological districts support 13 different species.
- Seventeen ecological districts appear to support no Myrtaceae species.

Recommendations

- The Ministry for Primary Industries (MPI) has funded a project to map Myrtaceae distributions by 31 May 2019, using more sophisticated species distribution modelling approaches than could be applied in the current project. The data and methods developed in this project will support the MPI project and will result in a major improvement.
- However, the underlying database could be markedly improved by:
    - targeted digitisation of herbarium records from areas of known distribution but without extant data (most New Zealand herbaria are only partially digitised)
    - targeted digitisation of vegetation plot records from areas of known distribution but no readily available vegetation plot data
    - improving data quality from existing sources (Bioweb, NZVH herbarium network data, especially from the herbaria AK and NZFRI[2]) by, for example, improved flagging of cultivated records, improved geospatial resolution and validation, and verifying suspect identifications
    - improving data quality in the NVS databank by further examination of spatial outliers
    - gathering distribution data on the recently described new taxa of *Kunzea* and mapping these new taxa rather than the former broad concept of *Kunzea ericoides*
    - incorporating records of occurrences in the pre-human record to improve current estimates of natural species ranges – these are currently qualitative and subjective.

[2] AK – Auckland Museum, NZFRI – Scion herbarium (Index Herbarium codes)

1 Introduction

In response to the recent incursion of myrtle rust into New Zealand, the Department of Conservation (DOC) contracted Manaaki Whenua – Landcare Research (MWLR) to produce improved Myrtaceae distribution and abundance maps.

2 Background

Myrtle rust (*Austropuccinia psidii*) is a non-native pathogen that can infect many species of Myrtaceae. It was first documented in Australia in 2010 and is now resident across native Australian Myrtaceae. Myrtle rust was found on Raoul Island, in the Kermadec Islands, in late March 2017. A Ministry for Primary Industries (MPI) response was initiated in early April 2017. The disease is now considered to be established on Raoul Island, and DOC has taken responsibility for its ongoing management there.

Myrtle rust was found in Northland, in the North Island, on 2 May 2017. There are currently a number of areas in North Island where it has been positively identified: Kerikeri and Kaikohe, Te Kūiti Taranaki, Te Puke and West Auckland. The disease is, at November 2017, still being actively managed under the MPI-led response. It has not yet been found in the South Island or associated with natural areas of vegetation.

All native Myrtaceae species are considered to be at risk from myrtle rust (Ministry of Agriculture and Forestry 2011). Some of these species are currently abundant and widespread, while some are classed as threatened (e.g. *Metrosideros bartlettii* is nationally critical).

It is unclear how myrtle rust will respond in New Zealand climatic conditions. For example, there may be temperature gradients that could limit the range and rate of infection, and its impact on species populations may differ (e.g. the vulnerability of seedlings versus mature stems may differ among species).

Identification of the distribution of native myrtle species is fundamental to a national approach to monitoring the extent and impact of myrtle rust on native plant communities. Species distribution maps will guide the design of monitoring programmes to monitor the occurrence of myrtle rust and its long-term impacts on forest structure and composition, and to identify differences in the resistance of species over their range.

3 Objectives

MWLR have already produced PDFs for DOC of 'polygonised' species maps for 19 native Myrtaceae species using plot data archived in the National Vegetation Survey Databank (NVS; https://nvs.landcareresearch.co.nz/), the New Zealand Virtual Herbarium (NZVH; http://www.virtualherbarium.org.nz/) and i-Naturalist Naturewatch

(http://naturewatch.org.nz/) with categorised species occurrences (based on the fraction of NVS plots with records), using ecological districts as recording units.

The data harvesting, data validation and cleaning process (data pipeline) were based on those developed for a project funded by the Ministry of Business, Innovation and Employment (2015–2019) called Innovative Data Analysis (IDA), which was intended to support improved methods for national environmental reporting. We chose ecological districts as a final mapping unit because they are environmentally and biogeographically homogeneous areas, and using them to cautiously extrapolate point location data is robust. The extrapolation from point occurrence to polygon data also allows the inclusion of 'sensitive' records (threat-listed species and records on private land), which cannot otherwise be mapped. The objective of this project is to enhance and improve these maps.

4 Methods

4.1 Overview of the native Myrtaceae occurrence data set

Occurrence records for the listed species of the native Myrtaceae recorded from New Zealand were assembled from five sources:

- NVS public domain data, as supplied to the Global Biodiversity Information Facility (GBIF) in June 2017
- NVS level 2 data for tree species (i.e. data where permission for use must be obtained from data owners, as assembled for the IDA programme in April 2017)
- data provided to the NZVH network as of 1 October 2017
- research-grade records from i-Naturalist/NatureWatchNZ
- the Bioweb database as at 12 September 2017 – Bioweb is administered by DOC and includes the national distribution and abundance records of threatened plant species.

Although data have already been assembled for the earlier analyses, we used our data assembly pipeline to repeat the process, incorporating improvements learned from the previous exercise. This provided over 80,000 occurrence records for 639 taxa (including lower-level taxa at the variety and subspecies level and all species in the Myrtaceae family, exotic or native).

We have selected the subset of 18 Myrtaceae species accepted as indigenous to the New Zealand mainland for mapping. We have done all mapping at the species level, and records of infra-specific taxa are included within the parent species. *Kunzea ericoides* is treated in the broad sense, because the determinations for most records precede the revision of de Lange (2014). There are few records of the newly defined taxonomic entities, and the task of re-interpreting older records, especially for observation-only records, is not straightforward. *Metrosideros kermadecensis* was excluded because its natural range does not include the New Zealand mainland.

4.2 The data processing pipeline

The IDA data delivery pipeline uses the tools provided by rOpenSci[3] to harvest data from GBIF and other Integrated Publishing Toolkit (IPT)[4] data providers into Darwin Core Archive (DwC-A) data format, as a common standard for all biodiversity occurrence data. Specifically, we use the R package *spocc* (Chamberlain et al. 2017). An extension of these tools was intended to allow access to New Zealand-relevant data sources and data-cleaning services. This meant, for example, that DwC-A file dumps of level 2 NVS data could feed into the same data processing pipeline.

In addition, taxon scrubbing services built on the New Zealand Organisms Register (NZOR; (http://www.nzor.org.nz/) provided New Zealand-specific taxon name resolution in a code-consistent manner with the R-OpenSci taxon scrubbing library. NZOR now has a re-developed taxon-matching service (created as part of the IDA and MWLR internal reinvestment work programmes), but we do not have R-OpenSci code-compliant libraries to access the service. As a consequence, IDA has components of a data pipeline that are relevant only to some data sources. For the current project it was necessary to go beyond the usual data sources and NZOR, and so while the data assembly and analysis are repeatable and employ tools and experience developed within the IDA project, it is not a documented IDA pipeline.

4.3 Data sources and processing

All non-NVS data were extracted and integrated from the sources. Given the restricted nature and use of the final data set (native Myrtaceae for mapping), only key fields were assembled in a simple format (i.e. data source, identifier, preferred taxonomic name, collection date, latitude and longitude). It should be noted that the final data set may contain multiple observations of the same individual. It was not possible within the timeframe to identify equivalent records from different data sources due to the high variability in geospatial precision between data sets, with some providing only one or two decimal digits of precision (Bioweb) and some none at all (NZFRI herbarium at Scion).

4.3.1 Taxonomic updating and selection

In this exercise, names were linked and corrected to standard form of the currently accepted name in the MWLR Plant Names Database (PND[5]). Only names of indigenous (non-endemic)/endemic species were propagated. Further, records of hybrids were excluded from the analyses. Names at infra-specific rank were treated at the species rank (e.g. *Leptospermum scoparium* var. *incanum*). Records at genus level or higher were excluded.

---

[3] rOpenSci is a project to create R packages that allow access to data repositories (see https://ropensci.org/).

[4] A free, open-source software tool that enables biodiversity databases to be published and shared through the GBIF network.

[5] *Ngā Tipu o Aotearoa – New Zealand Plants* (http://nzflora.landcareresearch.co.nz/).

Records of the recently recognised splits (10 taxa) from *Kunzea ericoides* were treated as *K. ericoides sensu lato*.

4.3.2 National Vegetation Survey Databank

We checked the veracity of the 100 records sourced from the NVS representing the most pronounced outliers based on current knowledge of distributional limits given in literature resources such as the New Zealand flora manuals (Allan 1982), Figure 5.3 of *Vegetation of New Zealand* (Wardle 1991) and the biological floras (Breitwieser et al. 2010). These were identified from previous analyses where there were only single records in an individual ecological district. This process resulted in 41 corrections.

4.3.3 NVS level 1 public data

These data were harvested directly from the MWLR IPT[6] installation, which feeds data into the GBIF network. The GBIF IPT data for the NVS represent a single observation of each species per plot, per survey event. The subset corresponding to the family Myrtaceae was extracted. Previously identified singleton records from the level 1 data were either removed as errors or re-assigned to another species. A total of 59,298 level 1 records of indigenous species are available for mapping; 150 records with coordinates in the sea were excluded.

4.3.4 NVS level 2 restricted data

As part of the IDA project we had permission to assemble a data set of tree species from the restricted level 2 data within the NVS. The subset corresponding to the family Myrtaceae was extracted and transformed into DwC format. It should be noted that these are data for a subset of indigenous Myrtaceae species that are trees[7]. Previously identified singleton records from the level 2 data were either removed as errors or re-assigned to another species. A total of 10,501 records were reduced to 7,413 unique records (the original data contain multiple records of the same species per plot due to multiple sampling protocols); 340 records had coordinates in the sea greater than 1 km from land and were excluded.

4.3.5 New Zealand Virtual Herbarium (NZVH)

Data were downloaded directly from the NZVH IPT provider database hosted on behalf of the NZVH by MWLR. This database includes threat-listed taxa (not publically available) and records with publically obscured coordinates (97 records). The data set has 14,756 total records for the Myrtaceae, of which 13,349 are records of species native to New Zealand. Of these, 12,000 have Lat/Lon georeferences and a further 140 georeferences in other coordinate systems (e.g. NZTM or NZMG). Of the other coordinate references, 80 were able to be converted to Lat/Lon. Data from the NZFRI (1,615 records from Scion) were only provided to the NZVH with integer values for Lat/Lon. One degree of Lat/Lon introduces too

[6] Internet Publishing Toolkit, which supports data harvesting by the GBIF network.

[7] 'Trees' as defined within the IDA project.

much uncertainty in location for these data to be included in the downstream analysis, and they were excluded.

After final data checking, and exclusion of records flagged as plants in cultivation, non-native taxa, and observations at genus level or above, 5,924 records were available for mapping; 1,136 NZVH records were found to have coordinates located in the sea and were excluded. Note that correct georeferencing appears to be a significant problem with herbarium records, and the quality of the remaining records remains unquantified. Within the NZVH the breakdown of original provider records is shown in Table 1.

Table 1 The number of Myrtaceae records utilised from the NZVH, by data provider

Institute code	*No. of records*
AK	3,123
CANU	335
CHR	2,339
LINC	139
MPN	163
NZFRI	1,616
UNITEC	180
WELT	1,595

4.3.6 BioWeb

The data were provided as a shapefile, which was extracted into a database for processing. The data set consists of 1,181 records, of which all but two are from the North Island and with a substantial focus on the west coast and Bay of Plenty. The Lat/Lon values were provided with only a single decimal digit precision. One hundred and sixteen records had coordinates located in the sea and were excluded; this is most likely to be a consequence of the lack of precision in the Lat/Lon fields in the BioWeb data provided.

4.3.7 i-Naturalist/Naturewatch

Records of Myrtaceae were downloaded directly from the NatureWatchNZ data repository by Jon Sullivan (Lincoln University). This provided access to records for threat-listed species and enabled observations of plants in cultivation to be flagged. A total of 3,803 records were provided, of which 1,755 were observations of native taxa not flagged as 'planted' and of 'research grade'. Fifty-one records of non-New Zealand observations were removed from the analysis.

4.3.8 Post-integration combined processing

The combined data set of 75,568 point records was intersected with the LCDB (version 4.1[8]). The LCDB was used to assign land/ocean status to the records and, if on land, the 2015 LCDB cover class to each record. As indicated above, records located in the sea were identified, and those within 1 km of land were assigned to the nearest ecological district (Table 2) and LCDB class. Where the LCDB layer was larger than the ecological district layer, we excluded these LCDB polygons to ensure analyses were carried out across a consistent area.

For the spatial processing done within QGIS (2/18.13)[9], the intersection of points with polygons was carried out using SAGA[10] 'add polygon attributes to points', and the proximity distances of points to polygons were evaluated using NNJoin plugin. Other calculations were carried out using Esri ArcMap 10.2.1[11]. The intersection of points with polygons was carried out using an 'intersect' from the Overlay subset of the Analysis Tools, and the proximity of points to polygons was evaluated using the 'near' function from the Proximity subset of the Analysis Tools.

Table 2 Data sources and off-shore georeferences

Data source	*Number of records having offshore Lat/Lon*	*Number of records having offshore Lat/Lon within 1 km of the coastline*
Bioweb	116	36
NVS Level 1	150	112
NVS Level 2	340	300
NatureWatch	51	17
NZVH	1,136	463
Total	1,793	928

Records in LCDB Class 1 – Built-up Area, Class 2 – Urban Parkland/Open Space, Class 5 – Transport Infrastructure, and Class 14 – Snow & Ice were excluded (Table 3) because these may not represent occurrences of the taxon in its native range. It should be noted, however, that some of the occurrence records do not have sufficient geospatial resolution to accurately position them relative to LCDB class polygons and are likely to be misclassified.

[8] LCDB: https://lris.scinfo.org.nz/layer/48423-lcdb-v41-land-cover-database-version-41-mainland-new-zealand/

[9] QGIS: http://www.qgis.org/en/site/

[10] SAGA: http://www.saga-gis.org/en/index.html

[11] Esri: https://www.esri.com/en-us/home

Table 3 Exclusion of occurrence records based on their LCDB class

Data source	*Number of occurrence records in LCDB Class 1, 2, 5 or 14*
Bioweb	73
NVS level 1	7
NVS level 2	67
NatureWatch	153
NZVH	311

One hundred and ninety-two NVS data records representing the only occurrence of a species in an ecological district (singletons) were excluded. We adopted a conservative approach, and such singletons require verification. Note that only 100 records of the originally identified singletons were re-assessed prior to this exercise, so residual singletons remained in the data.

A data matrix of the natural range of each taxon at the ecological district scale was prepared, based on known and confirmed published range data (Table 4). Range data from NZPCN[12] were sourced from the text associated with descriptions on that species' web page. The NZPCN website maps were not consulted as a source of information because they include NVS data. This matrix was used to exclude 895 records from all data sources assumed to be out of the natural range. Figures 37 to 54 show the range maps.

[12] New Zealand Plant Conservation http://www.nzpcn.org.nz/

Table 4 Literature sources used to define species ranges at the ecological district scale

Species	*NZ Flora (Breitwieser et al. 2010)*	*NZPCN*	*Wardle 1991*	*Simpson 2005*	*de Lange 2014*	*Wellington Botanical Society 2008*	*Simpson 2011*
Kunzea ericoides	y	y	y		y		
Kunzea sinclairii	y	y			y		
Leptospermum scoparium	y	y	y				
Lophomyrtus bullata	y	y					
Lophomyrtus obcordata	y	y					
Metrosideros albiflora	y	y					
Metrosideros bartlettii	y	y		y			
Metrosideros carminea	y	y					
Metrosideros colensoi	y	y	y				
Metrosideros diffusa	y	y					
Metrosideros excelsa	y	y	y	y			
Metrosideros fulgens	y	y	y				
Metrosideros parkinsonii	y	y					
Metrosideros perforata	y	y					
Metrosideros robusta	y	y	y	y			
Metrosideros umbellata	y	y		y			y
Neomyrtus pedunculata	y	y				y	
Syzygium maire	y	y					

A 'y' indicates the range of the taxon is described in that literature source.

As a result of these exclusions, the final data set consisted of 72,957 records. The final breakdown of records per species is provided in Table 5.

Table 5 Number of records per species

Taxon	*Number of occurrence records*
Kunzea ericoides	5,409
Kunzea sinclairii	143
Leptospermum scoparium	6,684
Lophomyrtus bullata	802
Lophomyrtus obcordata	1,059
Metrosideros albiflora	346
Metrosideros bartlettii	172
Metrosideros carminea	544
Metrosideros colensoi	280
Metrosideros diffusa	12,990
Metrosideros excelsa	716
Metrosideros fulgens	6,050
Metrosideros parkinsonii	193
Metrosideros perforata	5,909
Metrosideros robusta	5,404
Metrosideros umbellata	14,413
Neomyrtus pedunculata	11,370
Syzygium maire	473

The final breakdown of occurrence records by data source is provided in Table 6.

Table 6 Number of occurrence records per data source

Data source	*Number of occurrence records*
Bioweb	1,021
NVS level 1	58,305
NVS level 2	7,208
NatureWatch	1,525
NZVH	4,898
Total	72,957

Note that some ecological districts have known occurrences (published or from personal knowledge) for some of the species we list as absent from that ecological district. This is

because none of our data sources contain digitised records for those species in those areas. Our goal was to produce maps from clearly documented primary digital data using a repeatable process. These known deficiencies illustrate the need to digitise exemplar records held in herbaria as documented primary occurrence data.

4.3.9 Total land class areas per ecological district

Specific land-cover classes within ecological districts are likely to support particular species of Myrtaceae. Two groups of species with different predominant land cover class affiliations were identified by examining the frequency histograms of species occurrences in each land class nationally. Group 1 for *Kunzea ericoides sensu lato* and *Leptospermum scoparium* includes all land classes, except those initially excluded (1, 2, 5, 14), and also excluding Cropland (30), High Producing Grassland (40) and Orchards and Vineyards (33). Group 2 for all indigenous species of Myrtaceae except *Kunzea ericoides sensu lato* and *Leptospermum scoparium* includes the land classes Indigenous Forest (69), and Mānuka & Kānuka (52). The total land area per ecological district was calculated for these two groups. These are referred to as the two taxon groups below.

4.4 Estimating sampling effort

From the entire NVS databank a data set was extracted of all plot locations ever sampled. These data were used to create a 'sampling effort' data file containing the unique number of sampling events per ecological district (irrespective of date or data collection method). Plots associated with LCDB classes 1, 2, 5, 14, 30, 33 and 40 were excluded from further calculations. Records classified as offshore but within 1 km of land were assigned to the nearest ecological district and the nearest land class, as described previously.

Sampling effort per ecological district was calculated separately for the two groups of LCDB classes associated with the two species groups described in section 4.3.9. Thus our measure of sampling effort is based entirely on records from the NVS databank and the relevant LCDB class areas in which these records fall; the level of sampling based on NZVH and i-Naturalist/NatureWatchNZ is not incorporated. This is because whether a taxon is absent cannot be inferred from these latter types of data sources.

4.5 Defining abundance classes

The occurrence data set was used to generate a table containing a count of the number of occurrence records for each taxon in each ecological district. A matrix of presence/absence for each of the 18 species in each of 253 ecological districts (distant offshore islands excluded) was generated. An abundance category for each taxon in each ecological district was generated by combining these data according to the criteria summarised in Table 7, using the following method.

- We examined the sampling effort by ecological district (masked by the appropriate LCDB classes).
- We then calculated the plots per hectare in this area and constructed frequency histograms depicting all ecological districts nationally.

- Finally we considered an ecological district well sampled when the number of plots in that ecological district was in the upper tercile (1/3) of this distribution. This threshold contributed to the definition of the abundance classes.
- Thus sampling effort threshold was calculated separately for the two sets of LCDB classes associated with the two taxon groups.

Common: the taxon is common in a well-sampled ecological district. Specifically, the number of plot occurrences of that taxon in the ecological district is greater than 20% of the total number of plots in that ecological district and the sampling effort in that ecological district is within the upper tercile (1/3) of the sampling effort nationally. Note that designation of this abundance class is based on NVS data alone.

Present – medium/high sampling: the species is present in a low number of plots but the sampling effort is high, so the confidence that the taxon is not common is high. Specifically, the number of plot occurrences in an ecological district is less than or equal to 20% of the total number of plots in that ecological district and sampling effort in that ecological district is within the upper tercile of the sampling effort nationally. Note that designation of this abundance class is based on NVS data alone.

Present – low sampling: the species is present in plots but the sampling effort is low, so confidence in the designation as 'Present' is low. The sampling effort is not within the upper tercile. Note that designation of this abundance class is based on NVS data alone.

Rare or under-sampled: the species is absent from NVS plot data for the ecological district but present in other data sources (NZVH, i-Naturalist/Naturewatch).

Absent – moderate sampling: the species is absent with high confidence. Specifically, the taxon is absent from the ecological district (although present elsewhere in New Zealand) and the sampling effort is within the upper tercile. Designation as 'Absent' is based on absence from all data sources (NVS, NZVH, i-Naturalist/NatureWatchNZ).

Absent – low sampling: the species is absent from the ecological district (although present elsewhere in New Zealand) and the sampling effort is not within the upper tercile. Designation as 'Absent' is based on absence from all data sources (NVS, NZVH, i-Naturalist/NatureWatchNZ).

No data: There are no records from any of the data sources in the ecological district within the LCDB classes.

The distribution of the occurrence records for each taxon across the abundance classes is provided in Table 8.

Table 7 Summary of criteria for designating abundance classes

Occurs in > 20% of the plots in the ecological district	*Sampling effort within upper tercile (per taxon group)*	*Supported by NVS presence/ absence (for ecological districts with NVS plots)*	*Records from NVZH or Naturewatch within the ecological district*	*Abundance class*
Yes	Yes	Yes	N/A	Common
No	Yes	Yes	N/A	Present – medium/high sampling
Yes or No	No	Yes	N/A	Present – low sampling
No	No	No	Yes	Rare or under-sampled
No	Yes	No	No	Absent – moderate sampling
No	No	No	No	Absent – low sampling
No	No	No	No	No data

Table 8 The number of ecological districts (total n = 253) for each taxon in each of the abundance classes

Species	*Common*	*Present – medium/high sampling*	*Present – low sampling*	*Rare or under-sampled*	*Absent – medium/high sampling*	*Absent – low sampling*	*No data*
Kunzea ericoides	11	34	95	31	30	49	3
Kunzea sinclairii	0	0	0	1	82	165	5
Leptospermum scoparium	9	62	106	33	7	33	3
Lophomyrtus bullata	0	7	29	12	73	128	4
Lophomyrtus obcordata	1	20	52	29	55	91	5
Metrosideros albiflora	0	1	6	6	80	155	5
Metrosideros bartlettii	0	0	0	1	82	165	5
Metrosideros carminea	0	0	1	41	78	129	4
Metrosideros colensoi	0	2	8	47	76	116	4
Metrosideros diffusa	14	8	93	23	52	59	4
Metrosideros excelsa	0	3	17	18	78	132	5
Metrosideros fulgens	7	7	64	21	64	86	4
Metrosideros parkinsonii	0	0	9	8	81	150	5
Metrosideros perforata	7	11	68	28	60	77	2
Metrosideros robusta	2	5	65	18	71	88	4
Metrosideros umbellata	10	5	53	18	63	99	5
Neomyrtus pedunculata	10	11	87	18	56	66	5
Syzygium maire	0	6	9	26	74	133	5
Total	**71**	**182**	**762**	**379**	**1,162**	**1,921**	**77**

5 Results

We have produced six sets of maps and associated shapefiles, as follows.

- **The primary distribution maps for each taxon**: these are coded by abundance class at the ecological district scale, but the area displayed is masked to only display the LCDB classes believed to support that taxon. Taxa are assumed absent outside their known range, as depicted at the scale of the ecological district. Note that this can result in sharp boundaries between ecological districts that would be expected to be not so sharp in reality (Figures 1–18).
- **Presence/absence maps for each taxon at the ecological district scale** based on all records: no screening based on range limits or LCDB classes was applied (Figures 19–36).
- **Range maps for each taxon at the ecological district scale** based on the literature sources described in Table 4: these were used to restrict the display in the primary maps (Figures 37–54).
- **Maps of the coverage of the LCDB classes used to mask the primary maps**: Group 1 comprises *Leptospermum scoparium* and *Kunzea ericoides sensu lato*. Group 2 comprises the remaining 16 taxa (Figures 55–56).
- **A map of the richness (number of species) of native Myrtaceae, by ecological district** (Figure 57).
- **An overall sampling effort map**, showing the number of NVS plots per ecological district (Figure 58).

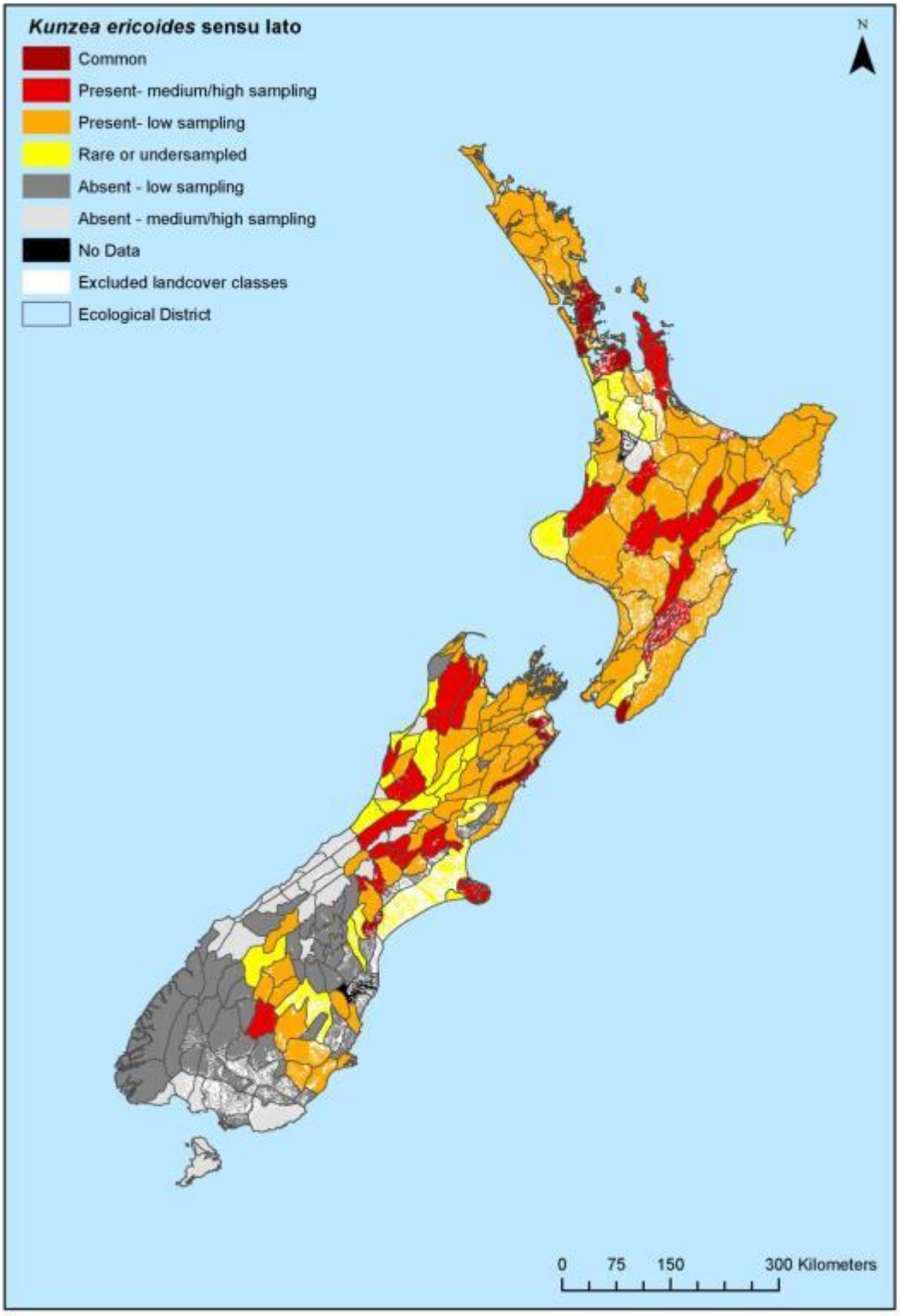

Figure 1 ***Kunzea ericoides sensu*** **lato – primary distribution**

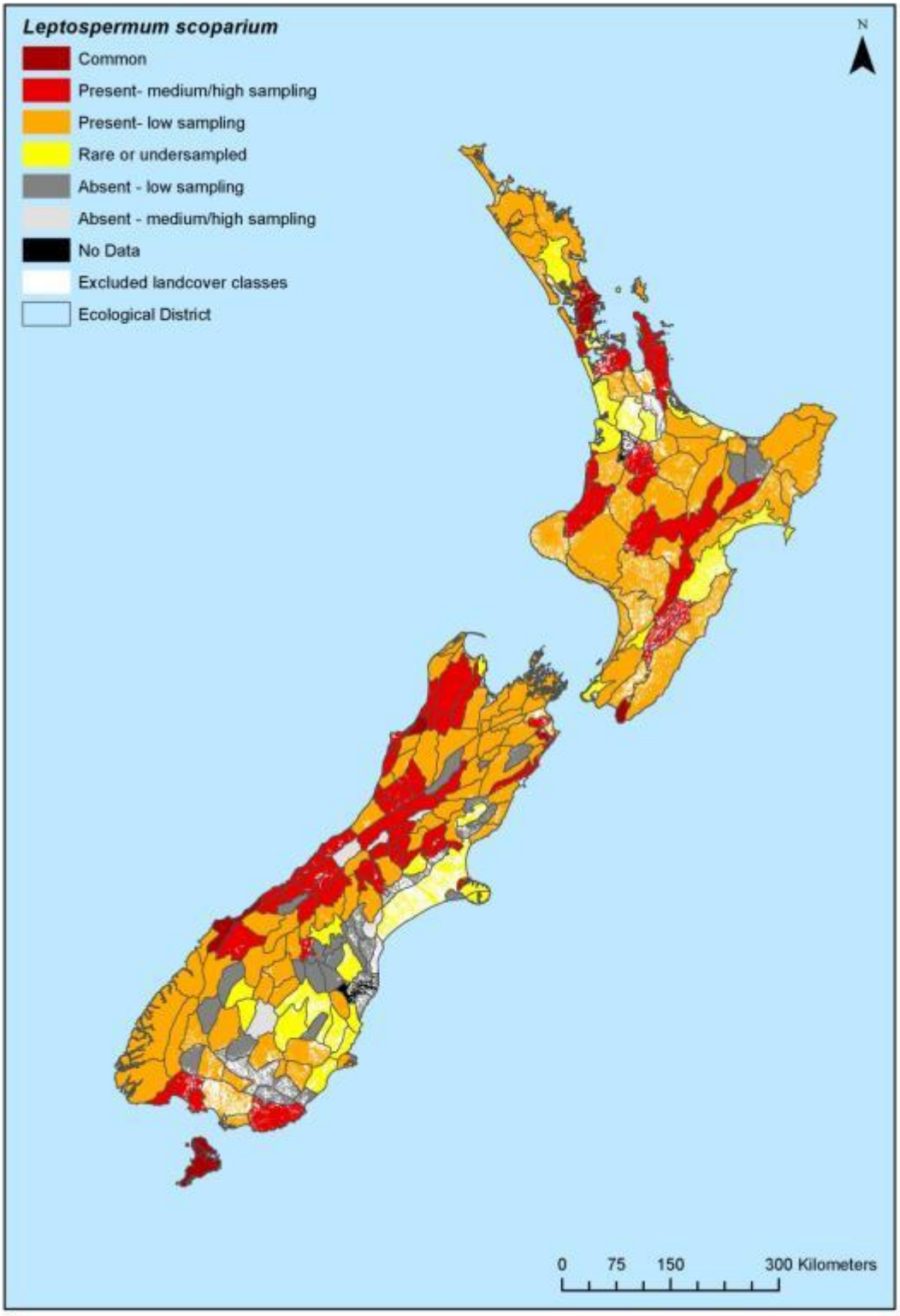

Figure 2 ***Leptospermum scoparium*** **– primary distribution**

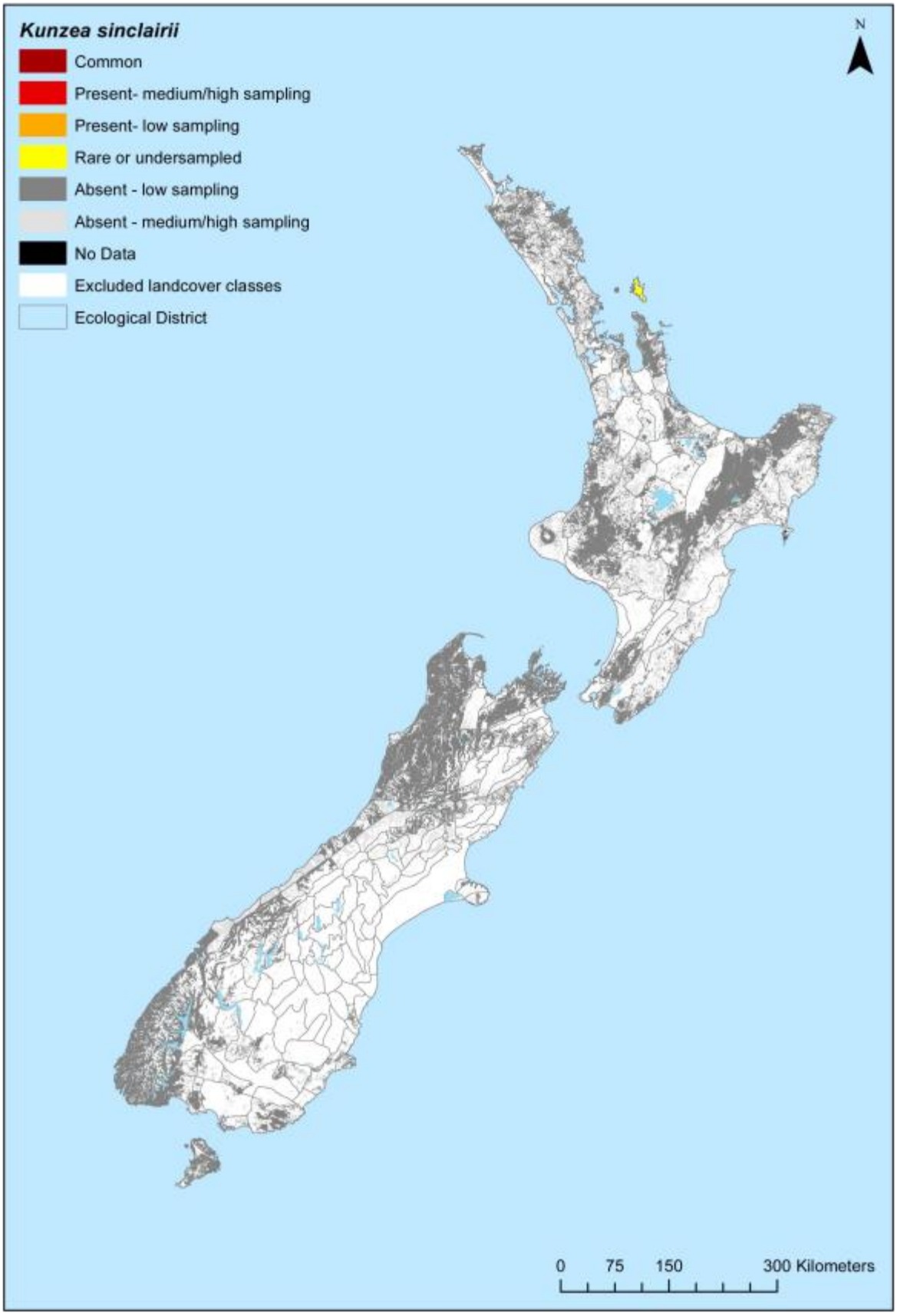

Figure 3 ***Kunzea sinclairii*** **– primary distribution**

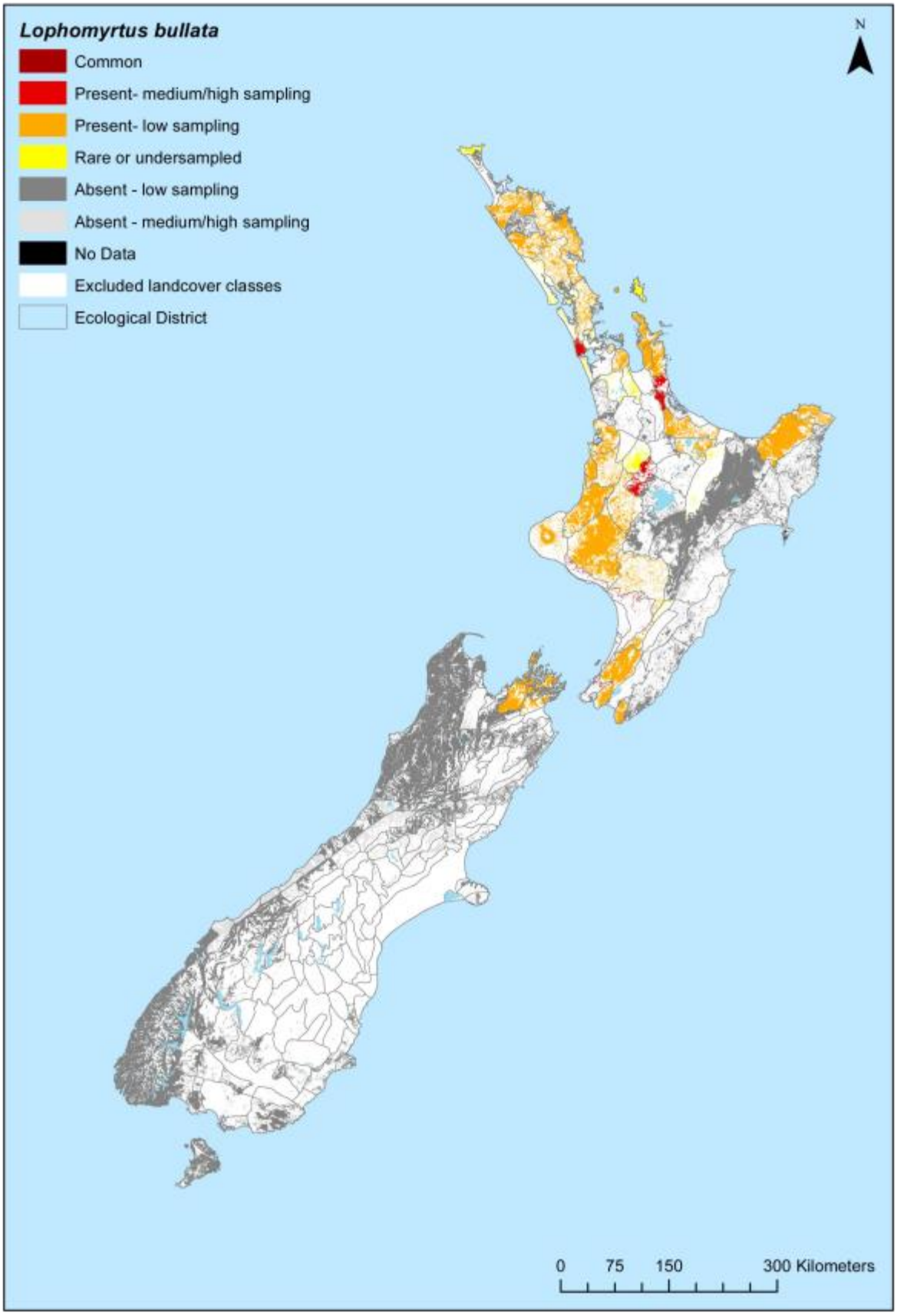

Figure 4 ***Lophomyrtus bullata*** **– primary distribution**

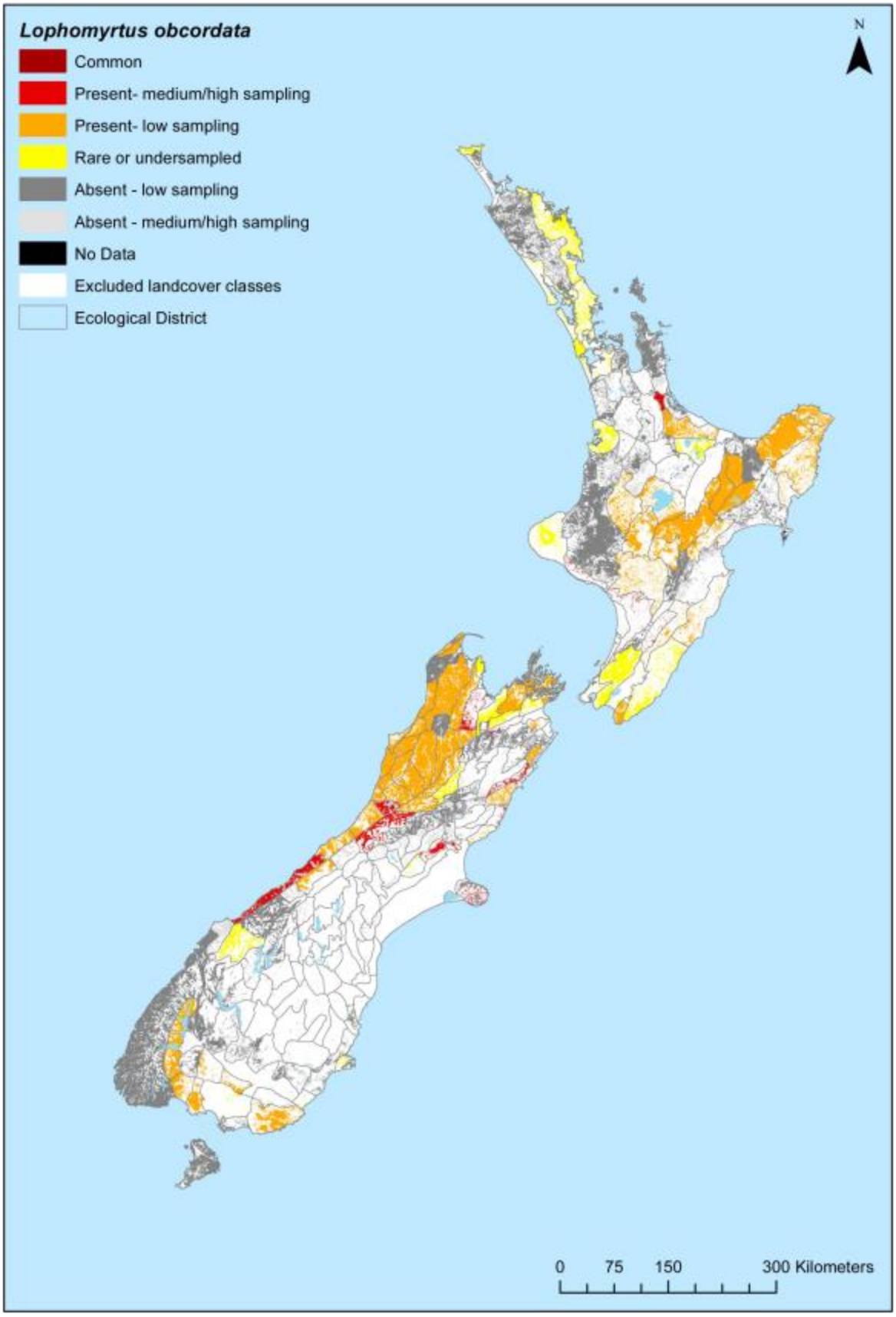

Figure 5 ***Lophomytrus obcordata*** **– primary distribution**

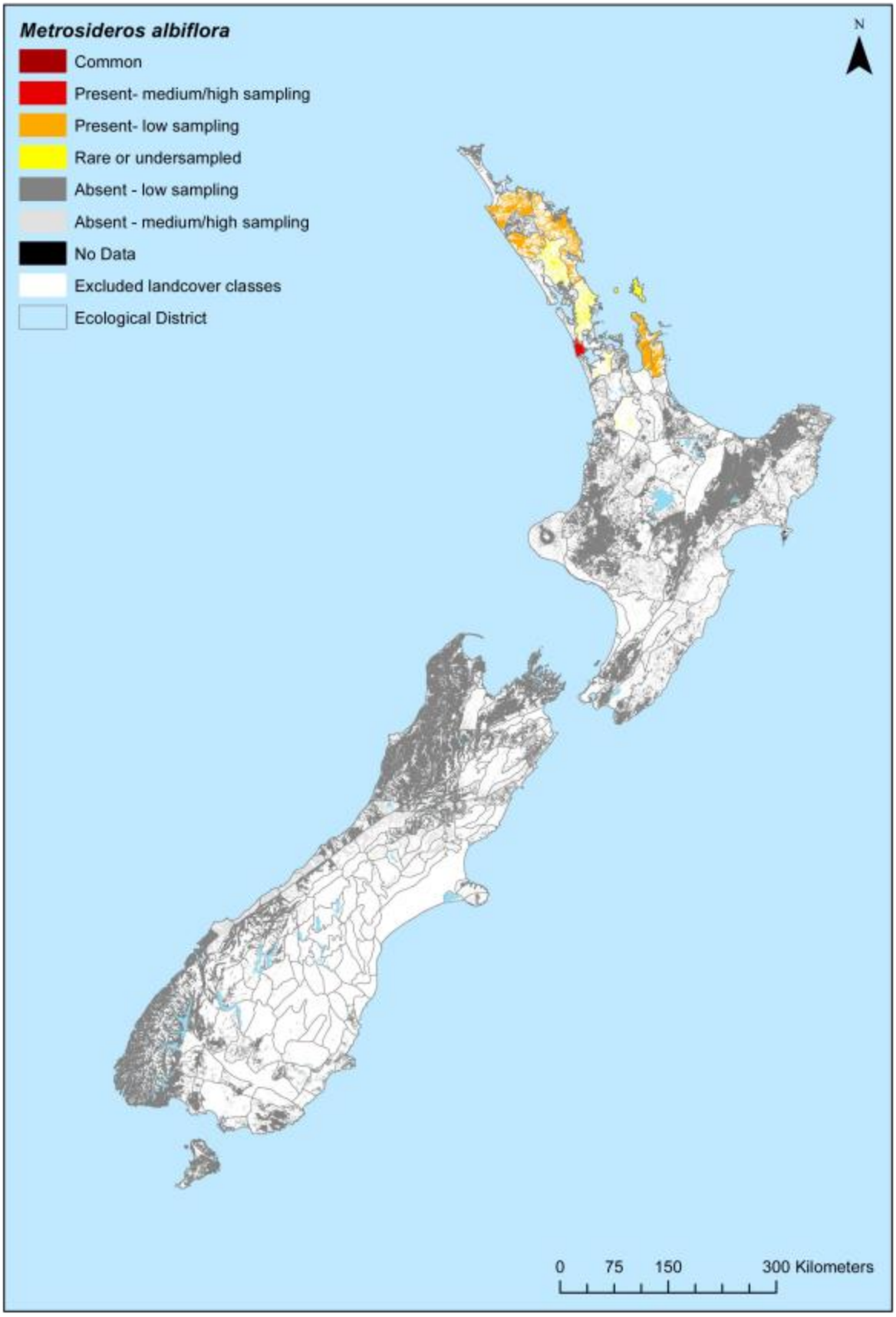

Figure 6 ***Metrosideros albiflora*** **– primary distribution**

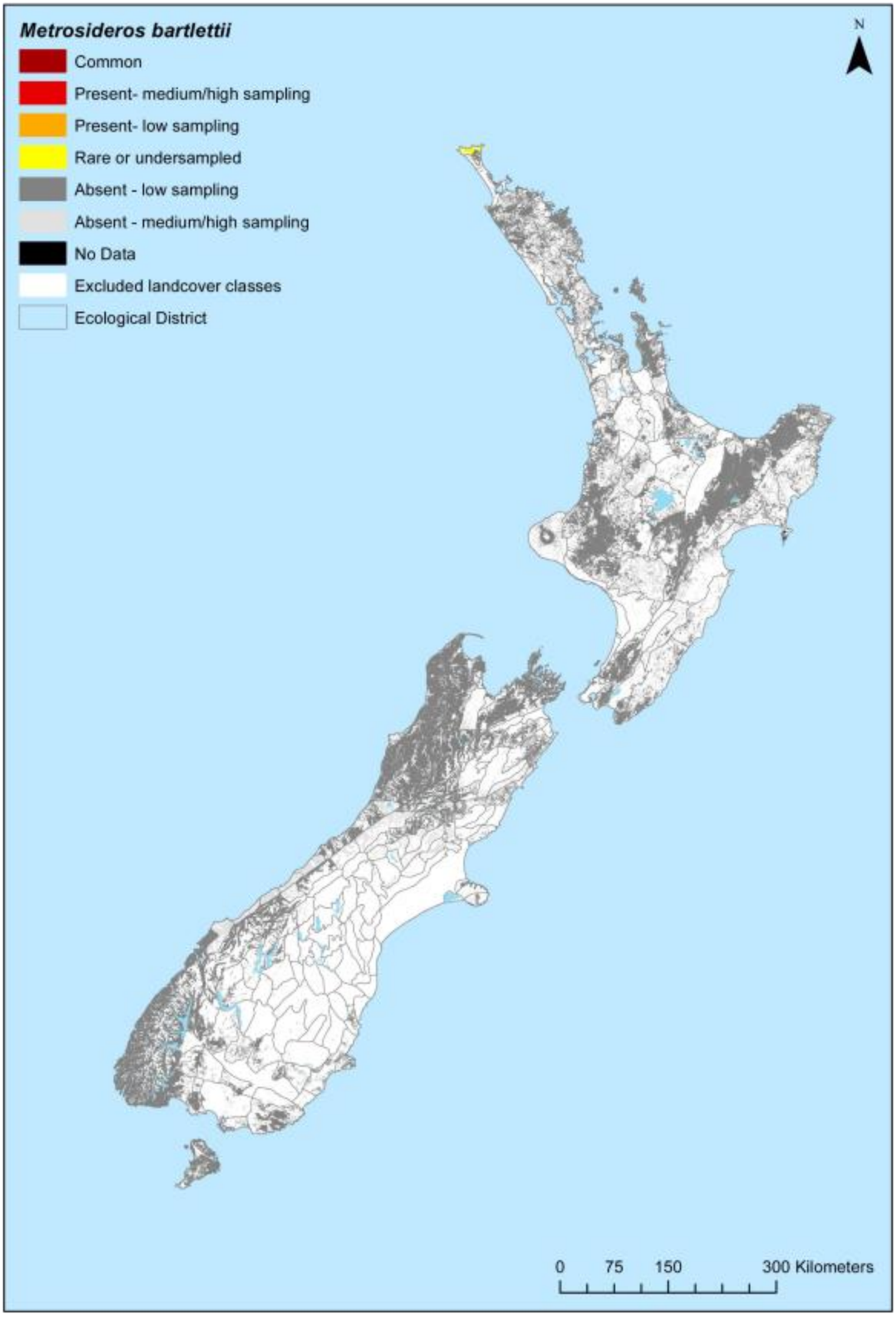

Figure 7 ***Metrosideros bartlettii*** **– primary distribution**

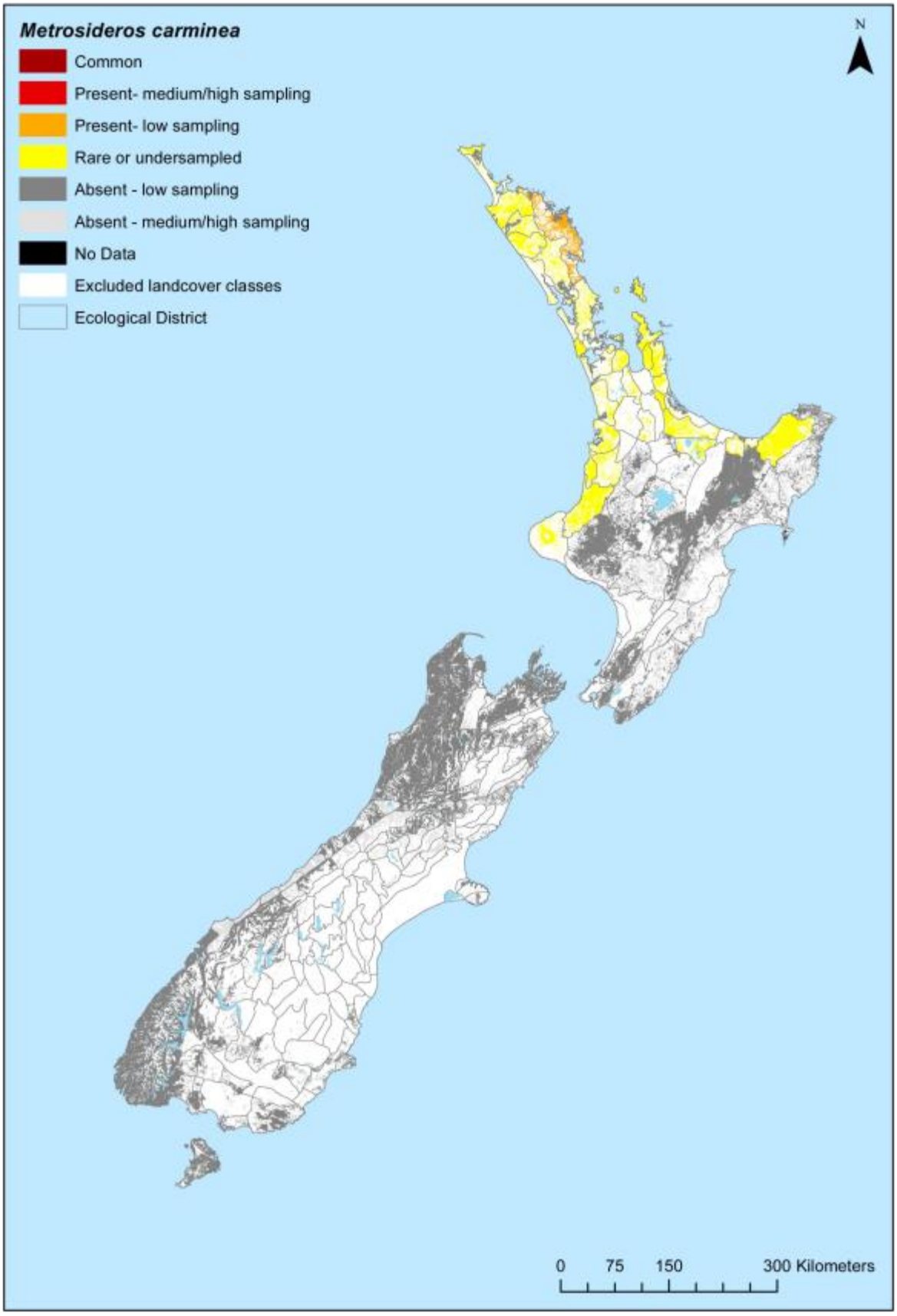

Figure 8 *Metrosideros carminea* – primary distribution

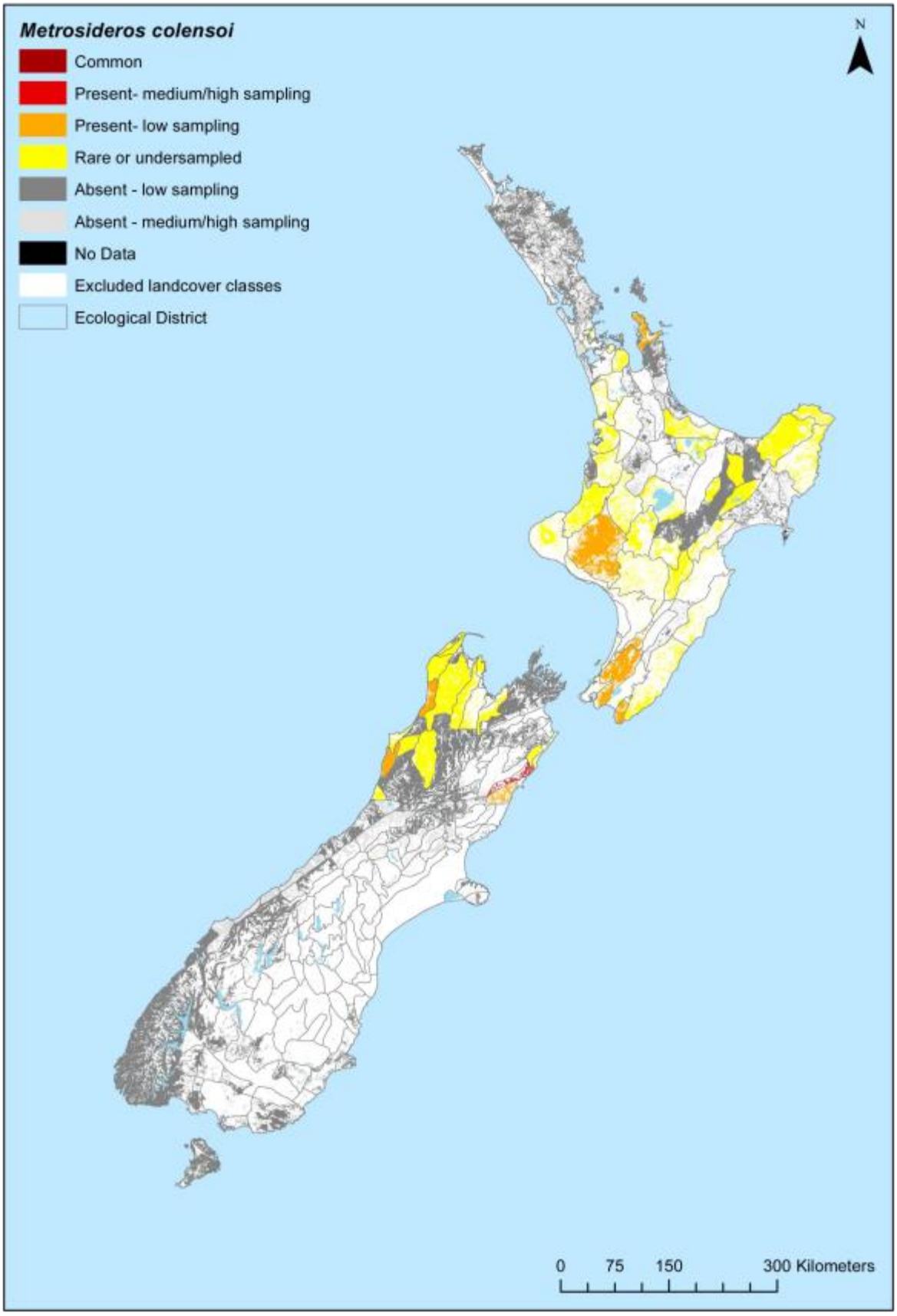

Figure 9 *Metrosideros colensoi* – primary distribution

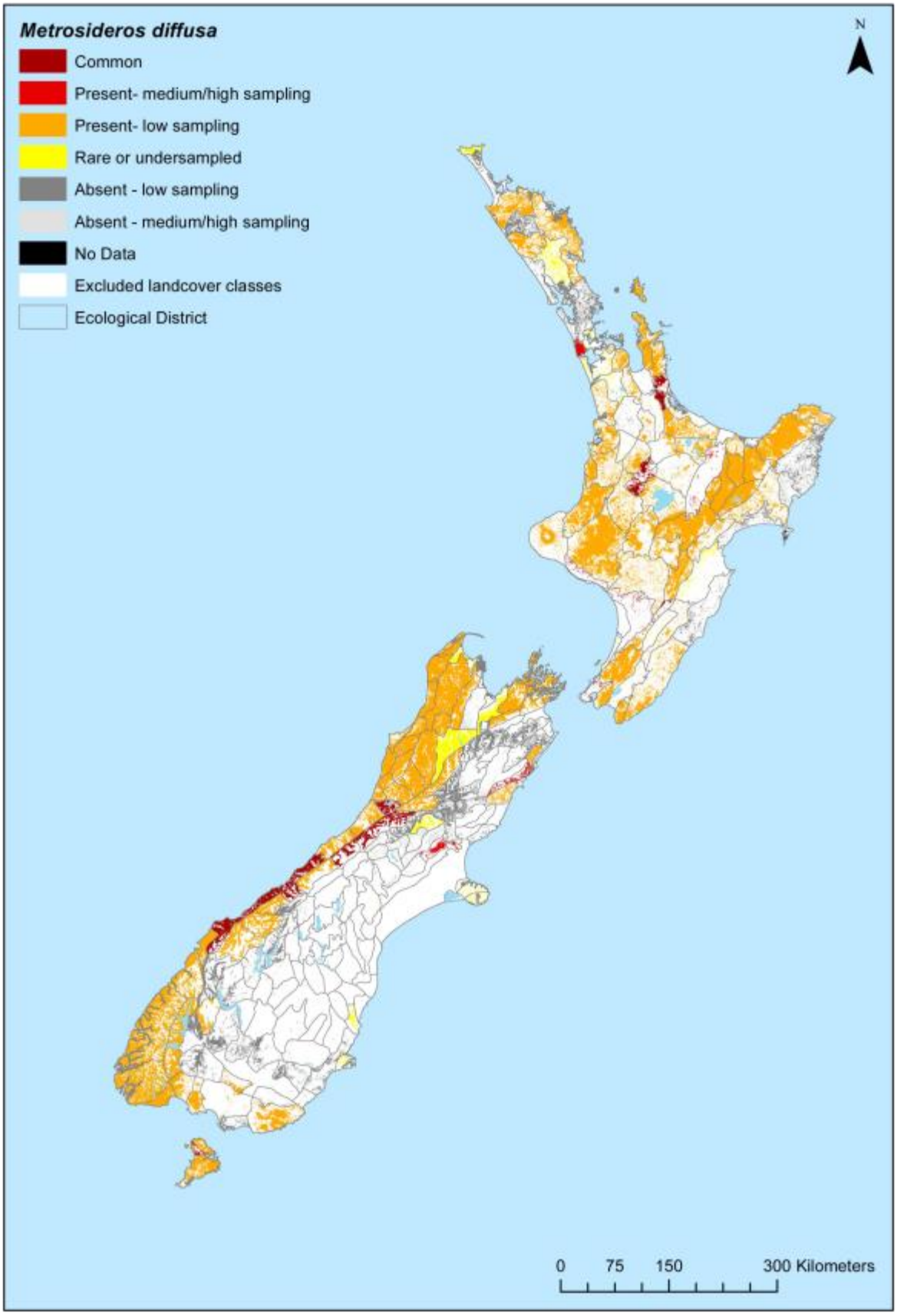

Figure 10 ***Metrosideros diffusa*** **– primary distribution**

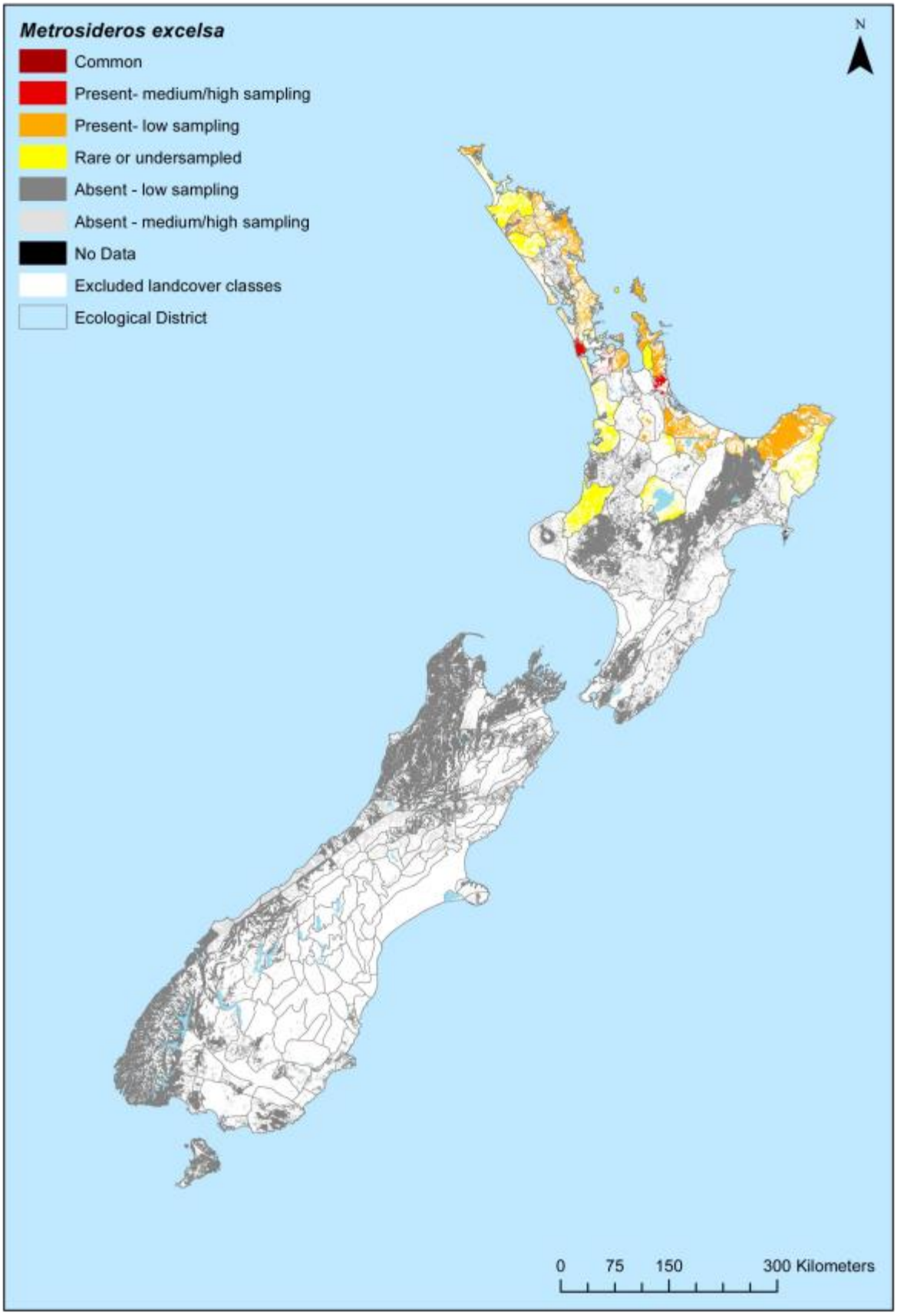

Figure 11 ***Metrosideros excelsa*** **– primary distribution**

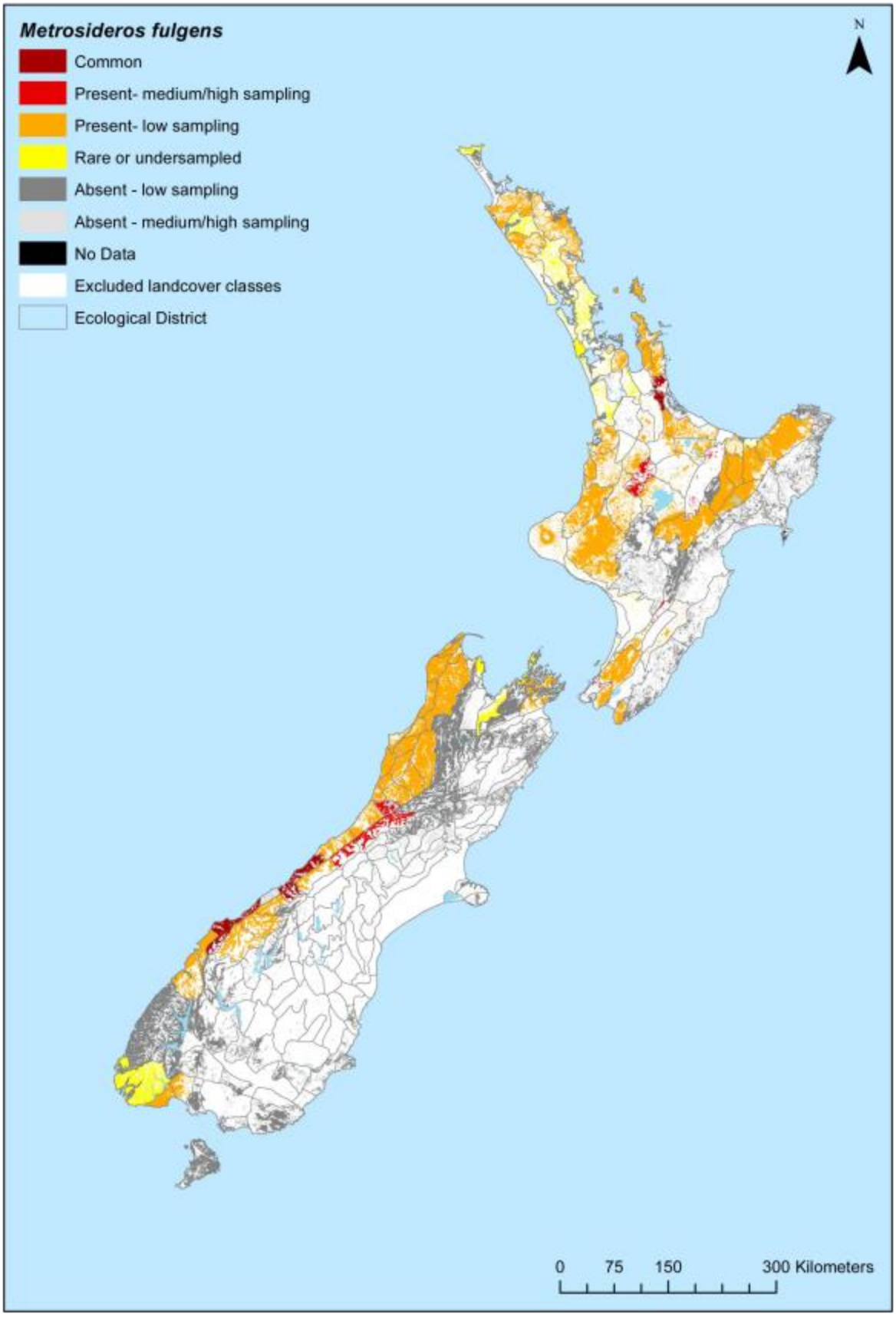

Figure 12 ***Metrosideros fulgens*** **– primary distribution**

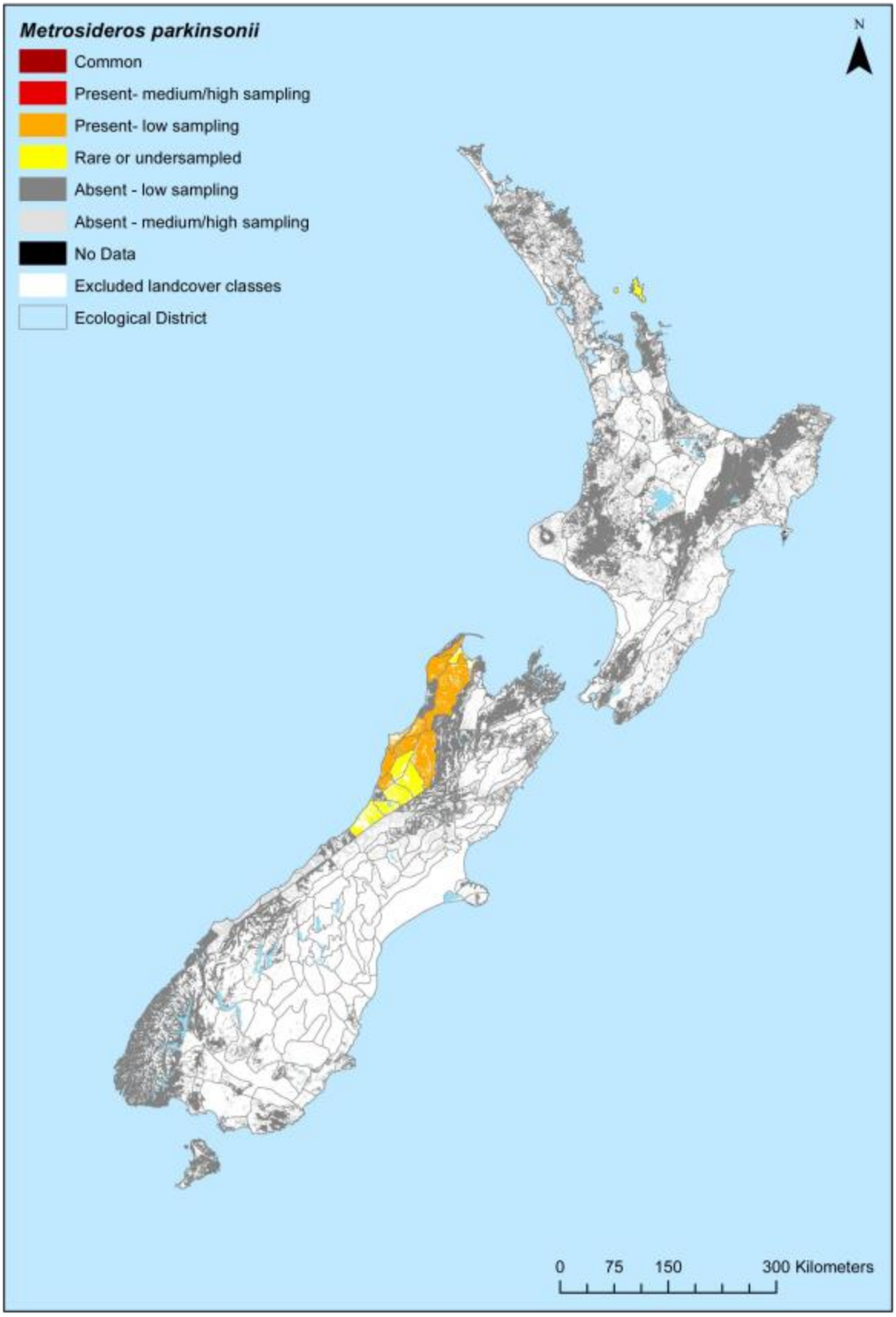

Figure 13 ***Metrosideros parkinsonii*** **– primary distribution**

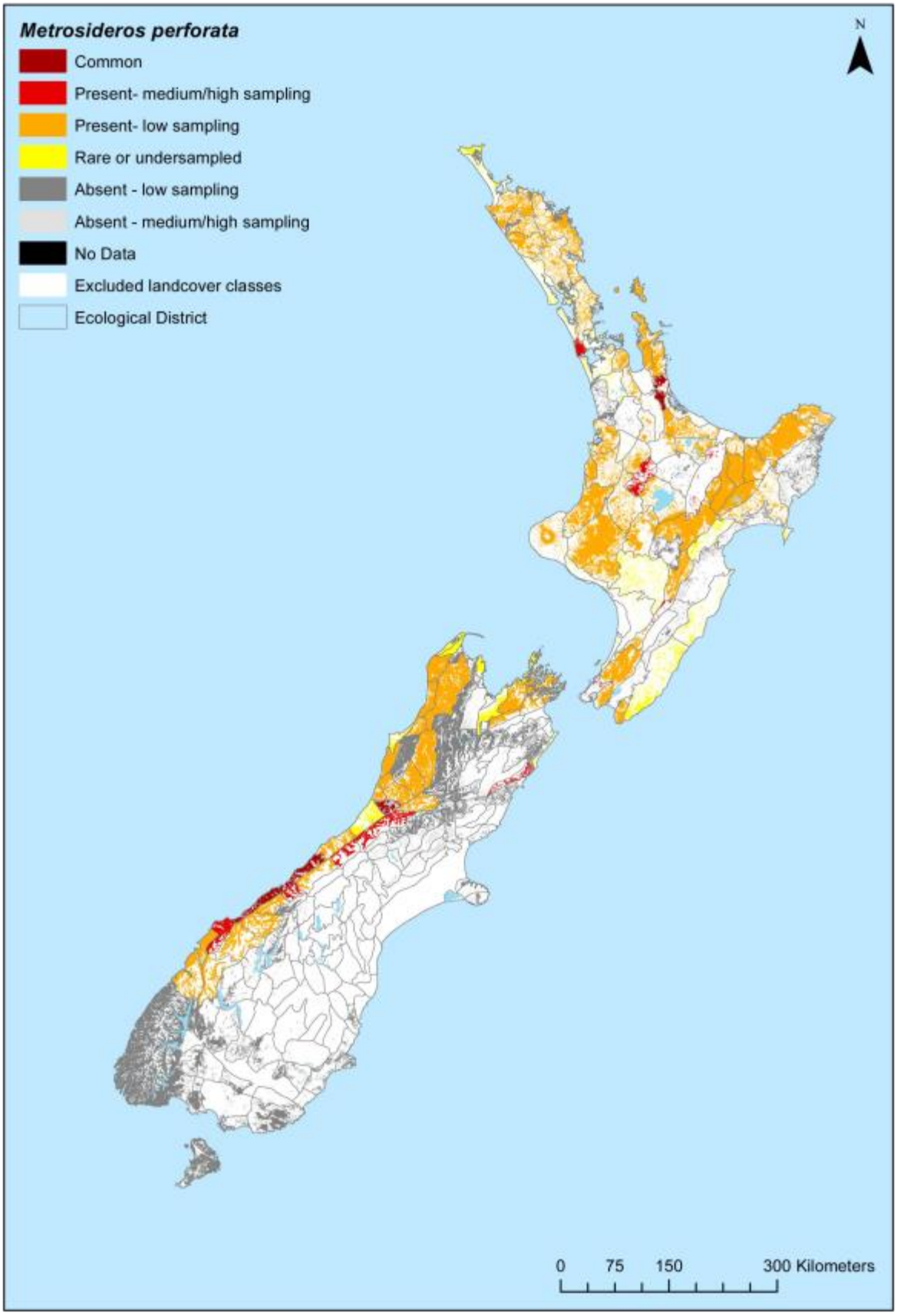

Figure 14 ***Metrosideros perforata*** **– primary distribution**

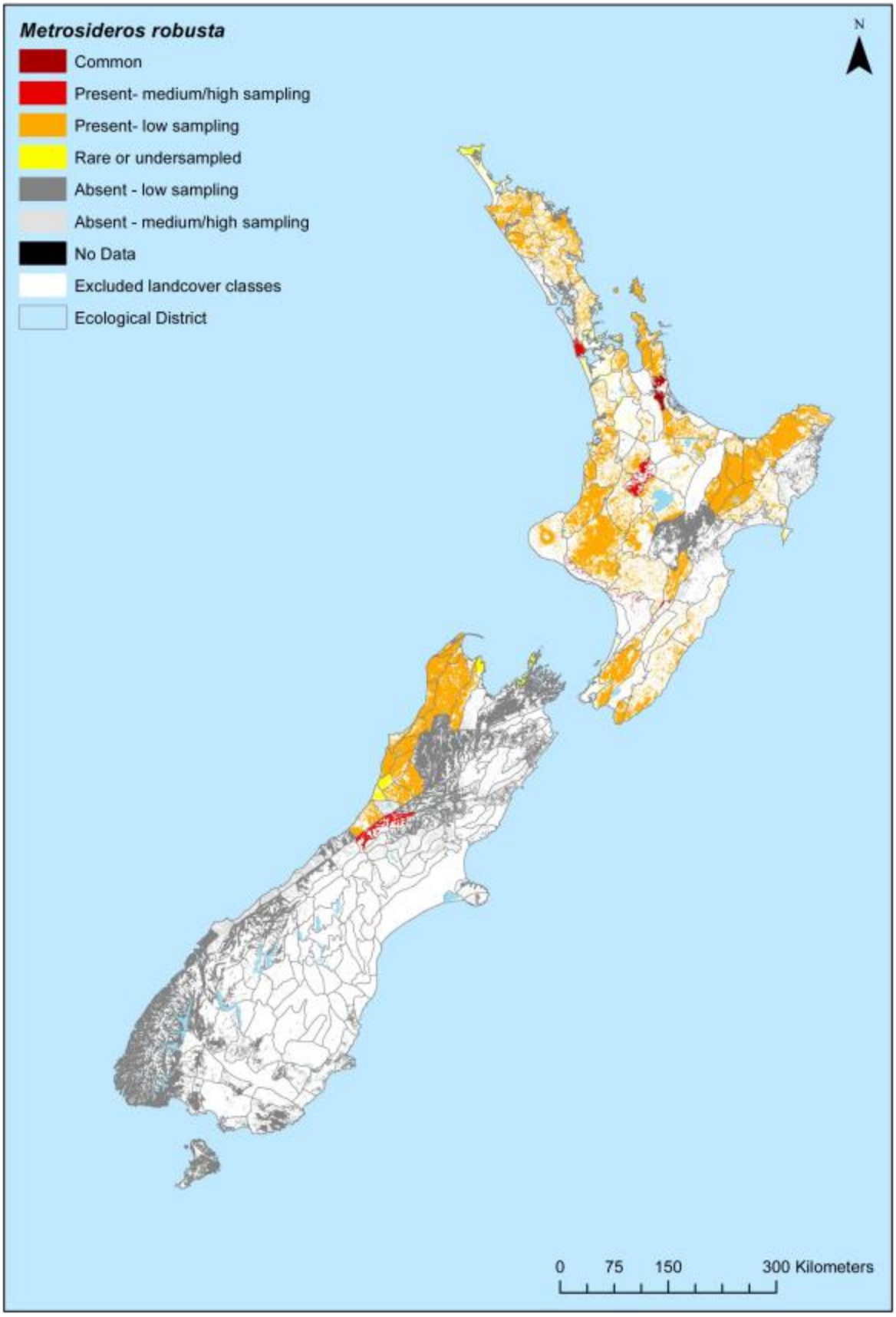

Figure 15 ***Metrosideros robusta*** **– primary distribution**

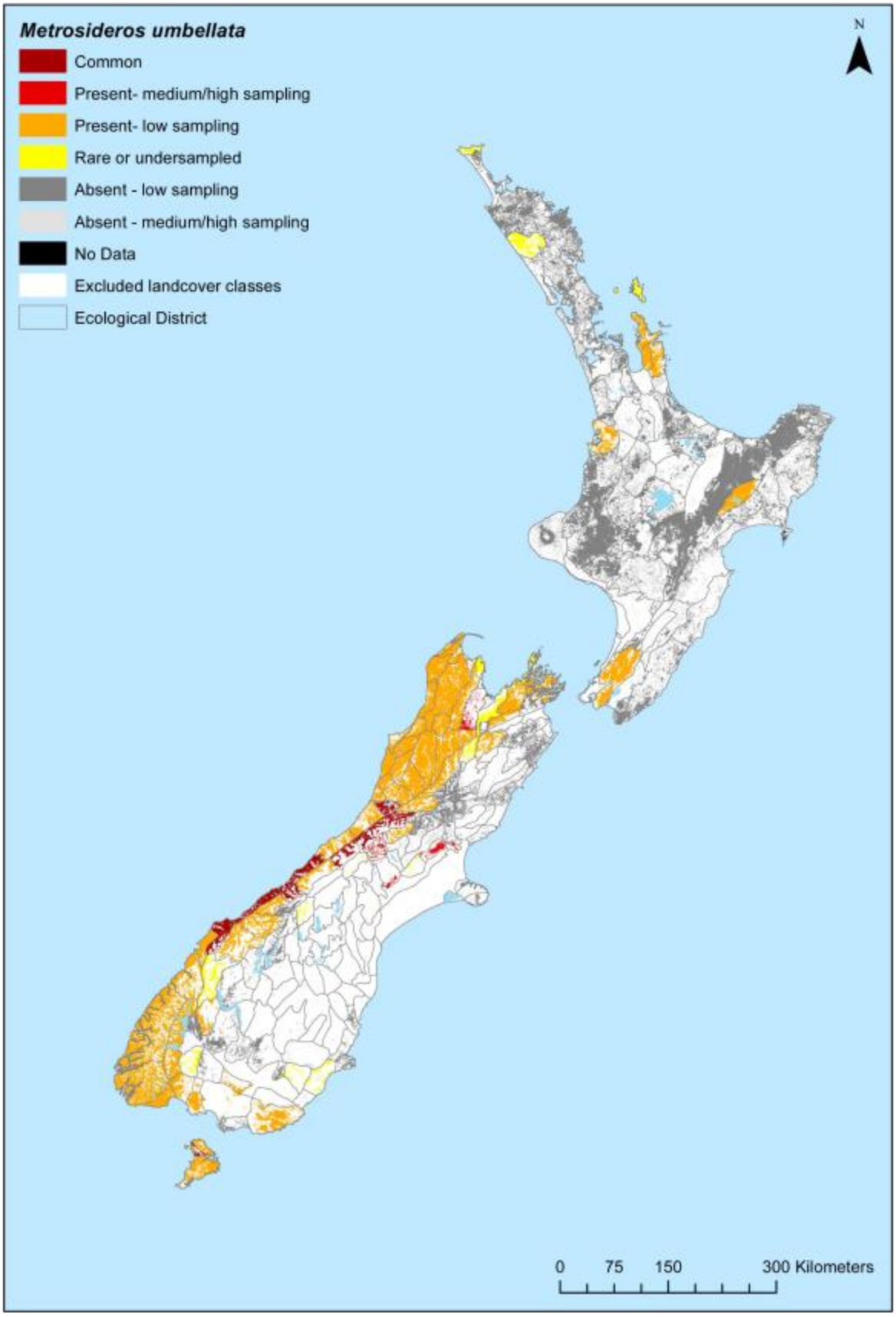

Figure 16 *Metrosideros umbellata* – primary distribution

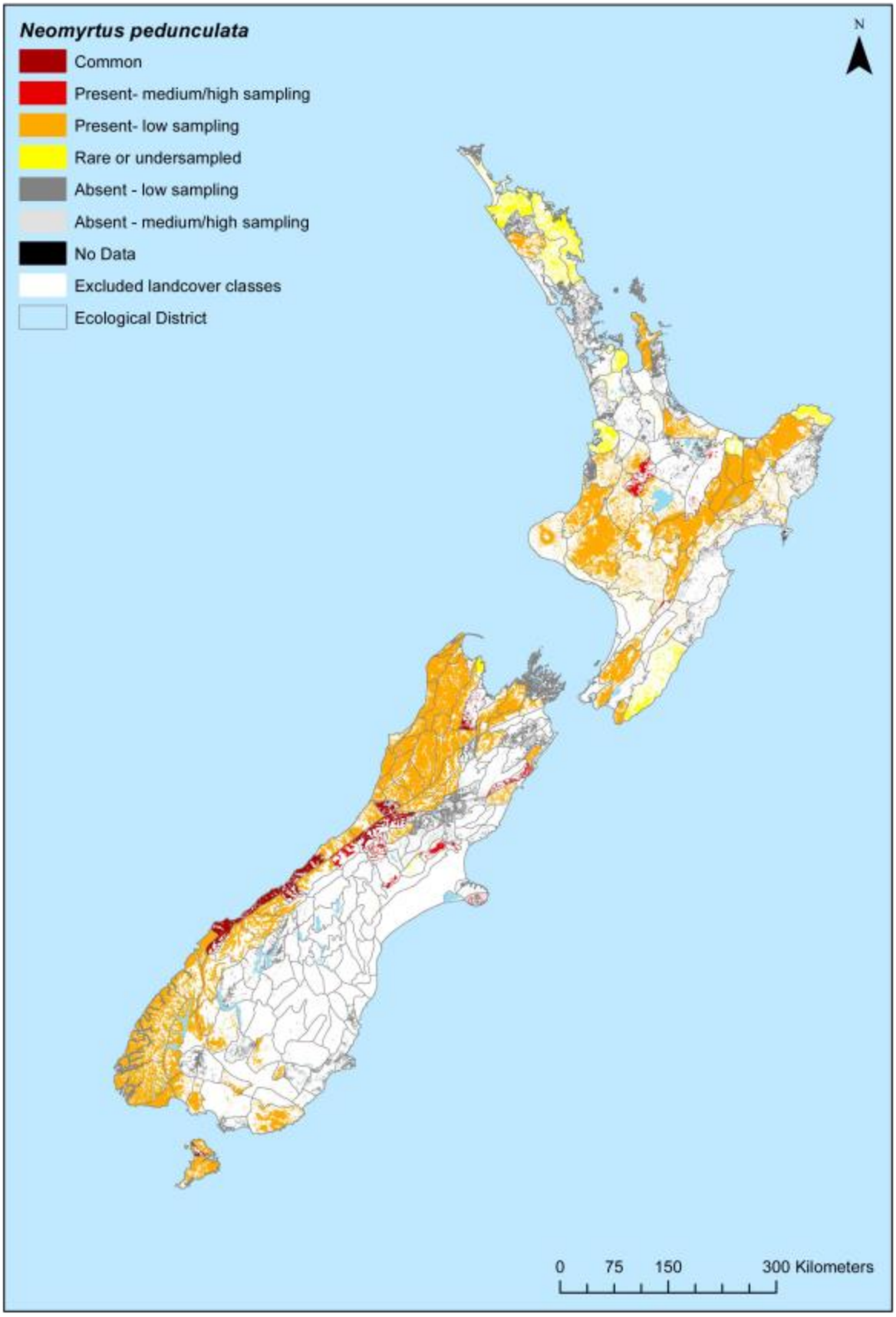

Figure 17 ***Neomyrtus pedunculata*** **– primary distribution**

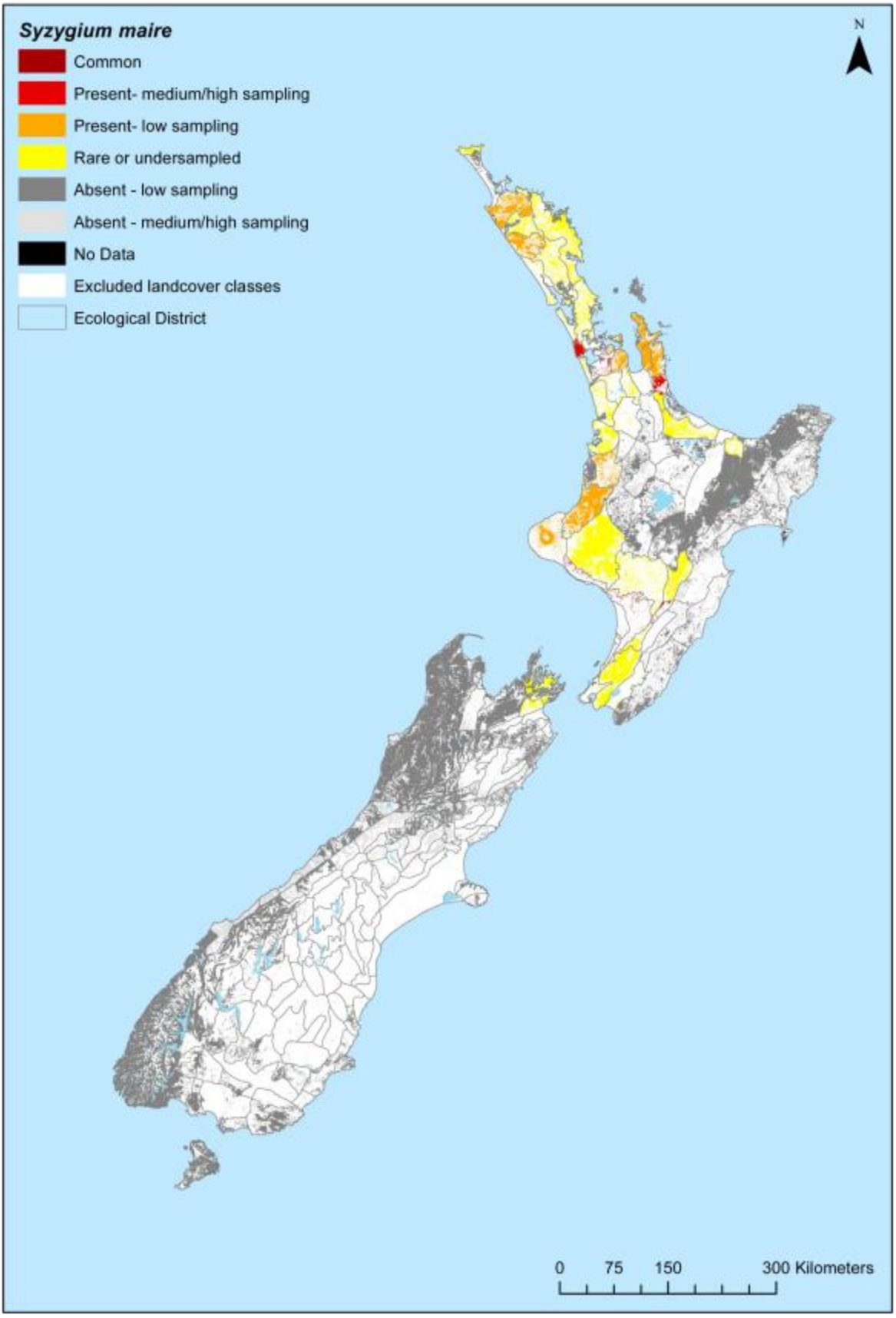

Figure 18 *Syzygium maire* – primary distribution

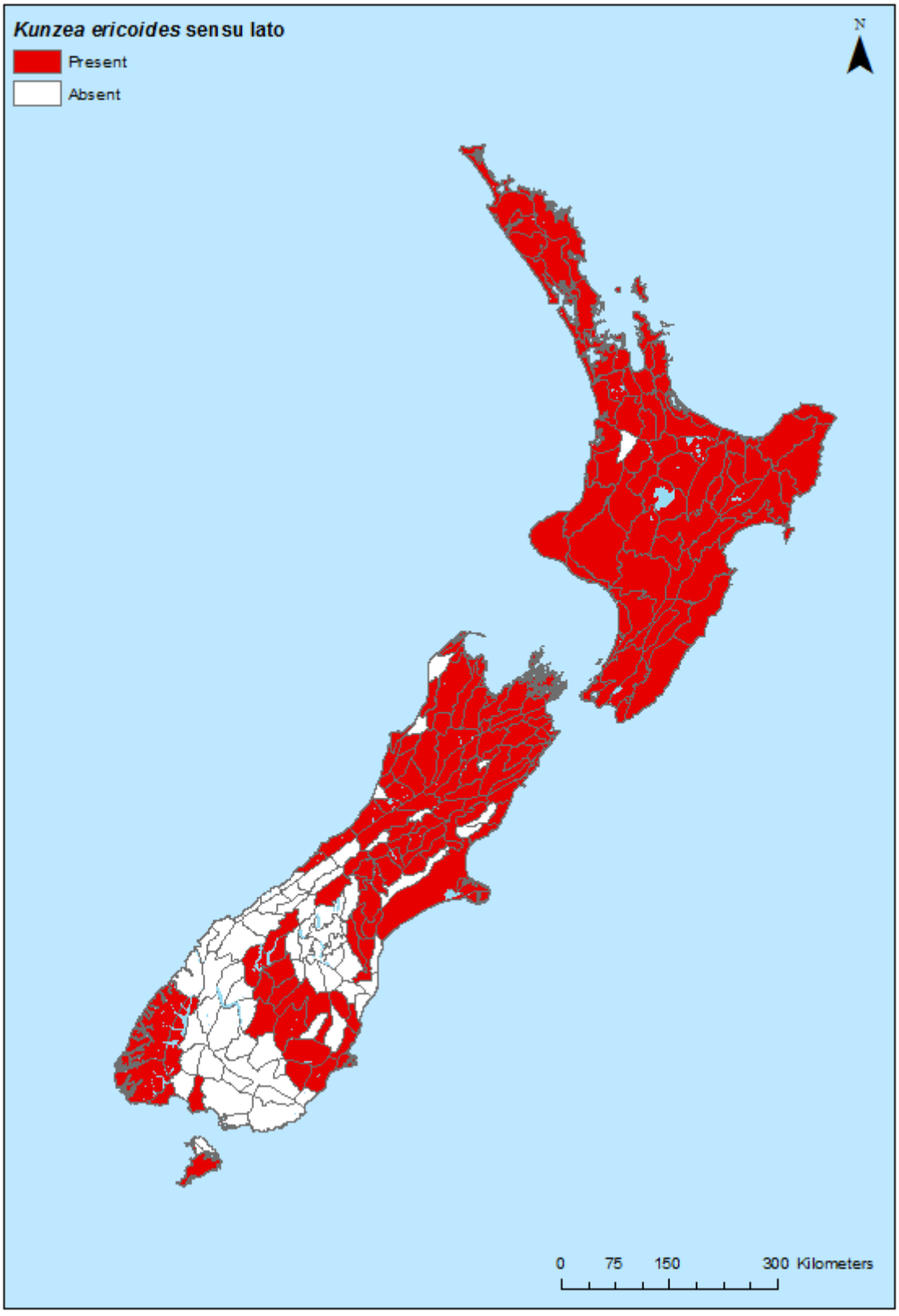

Figure 19 ***Kunzea ericoides sensu lato*** **– presence/absence**

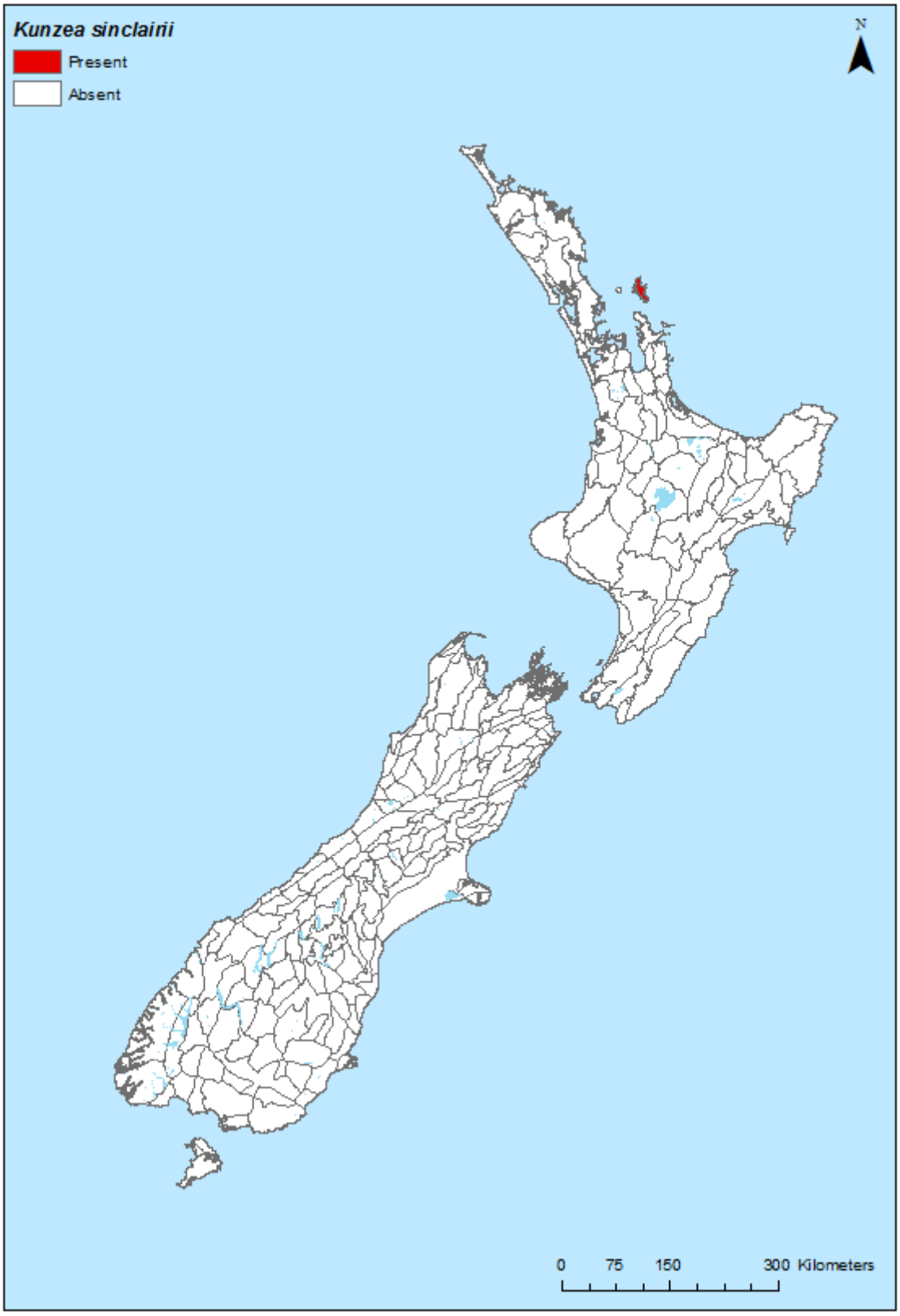

Figure 20 *Kunzea sinclairii* – presence/absence

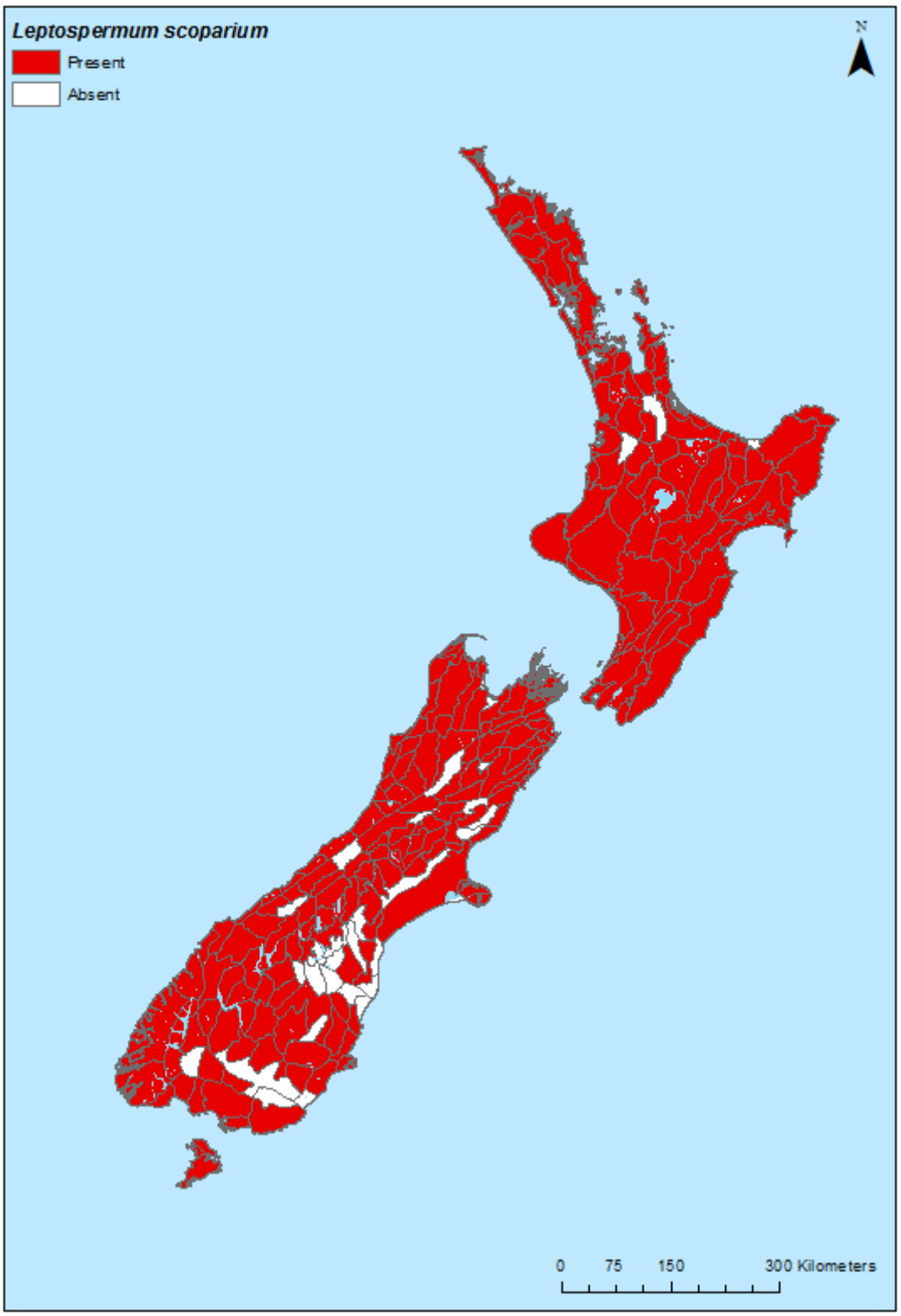

Figure 21 ***Leptospermum scoparium*** **– presence/absence**

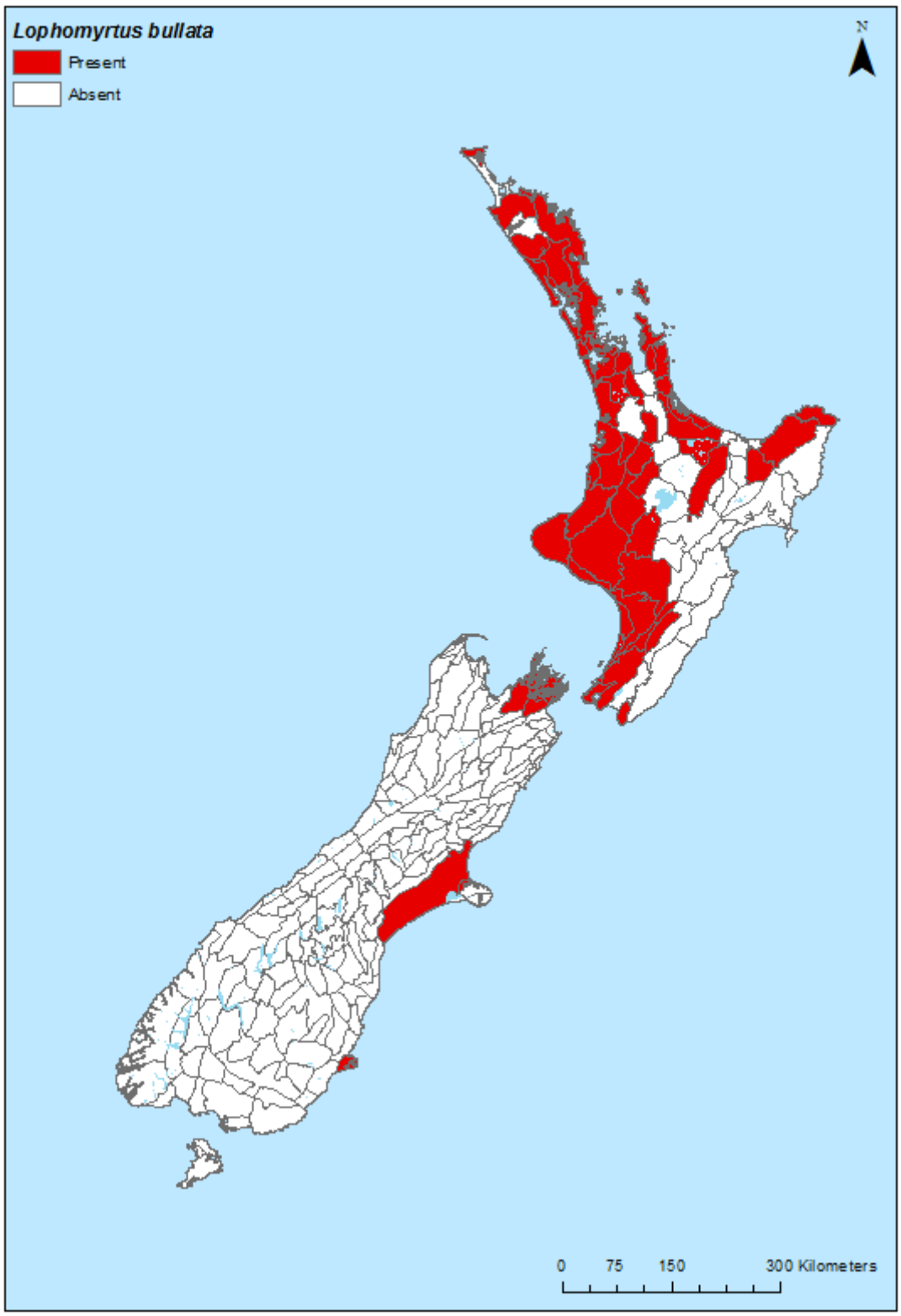

Figure 22 ***Lophomyrtus bullata*** **– presence/absence**

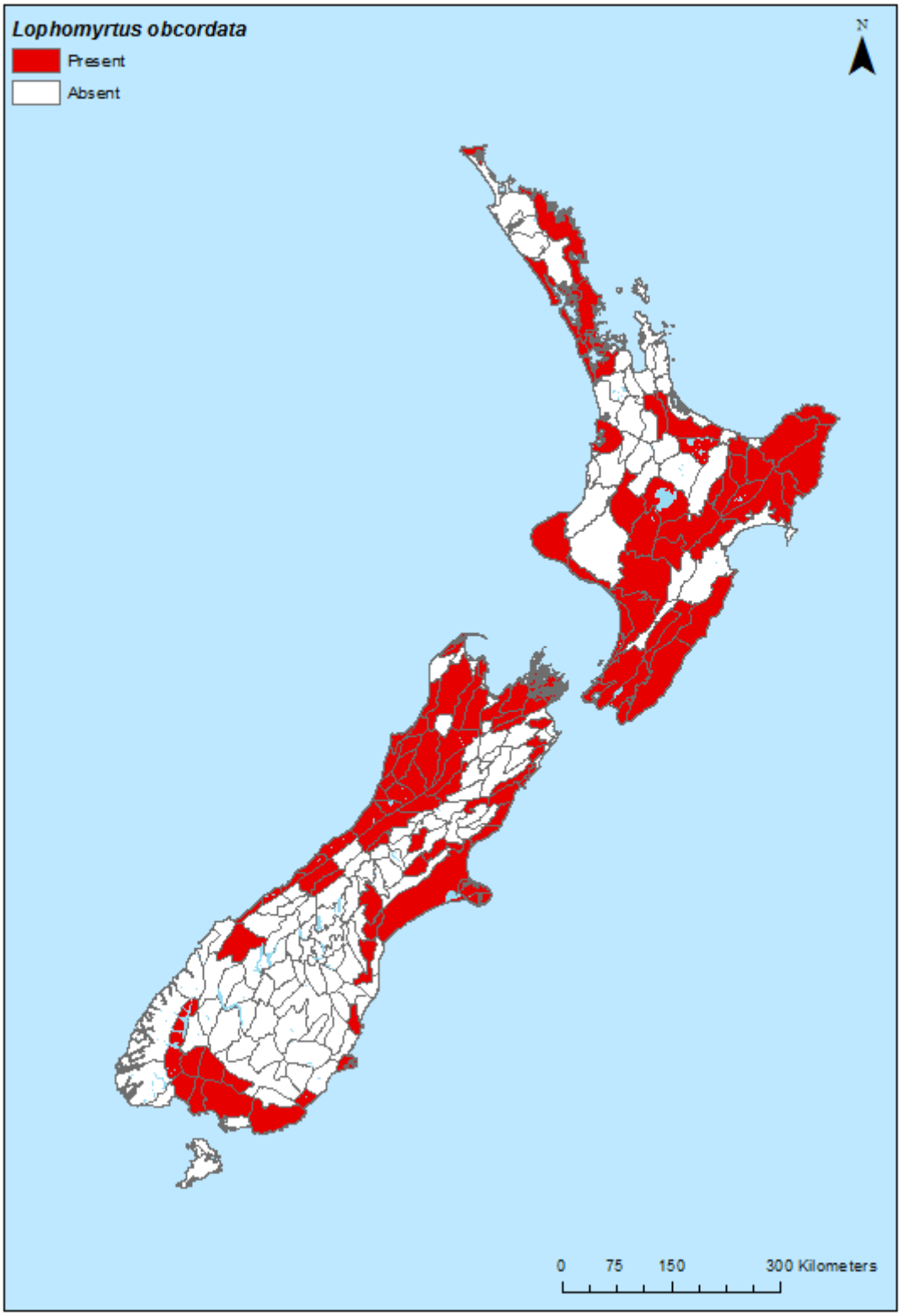

Figure 23 ***Lophomyrtus obcordata*** **– presence/absence**

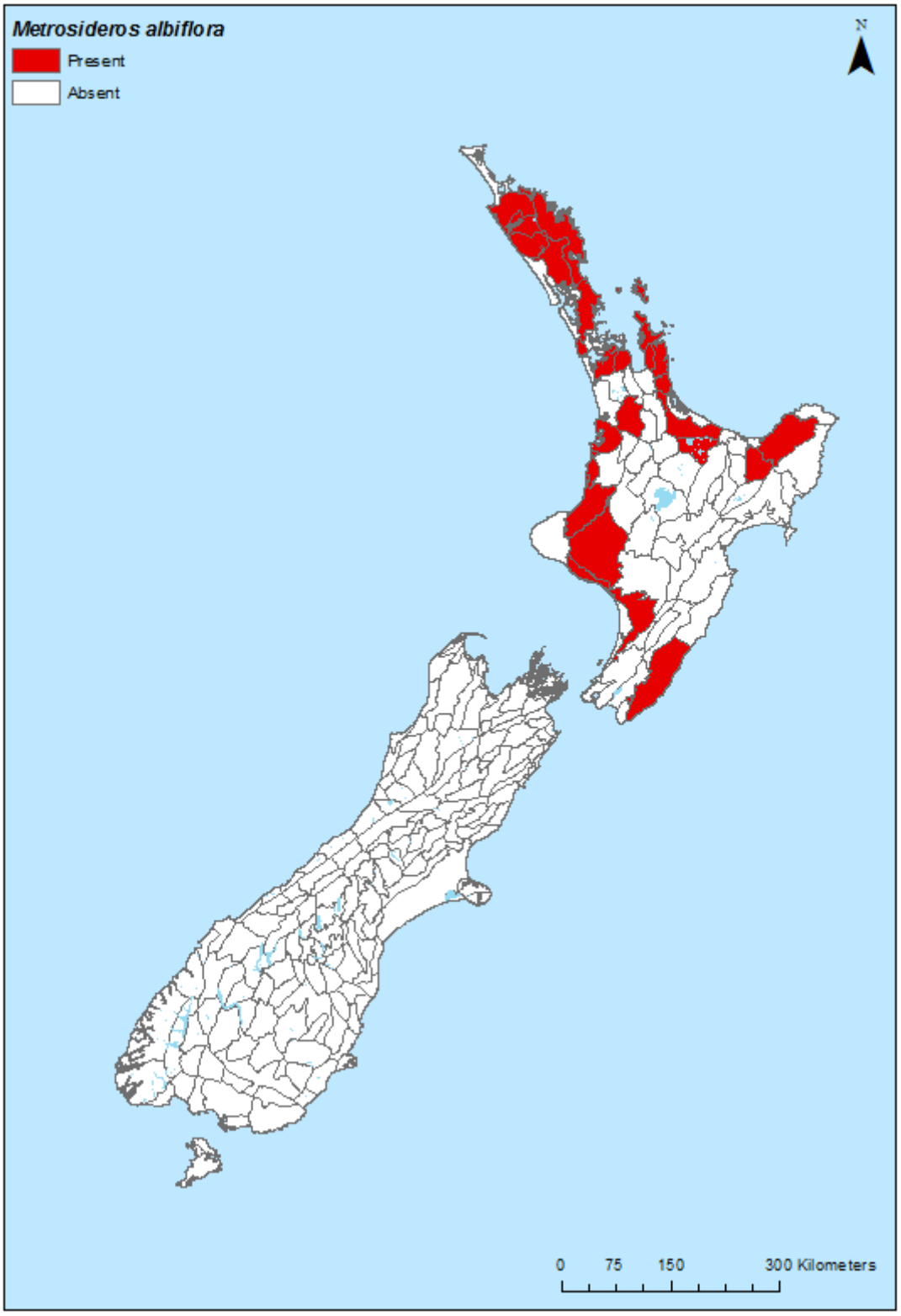

Figure 24 ***Metrosideros albiflora*** **– presence/absence**

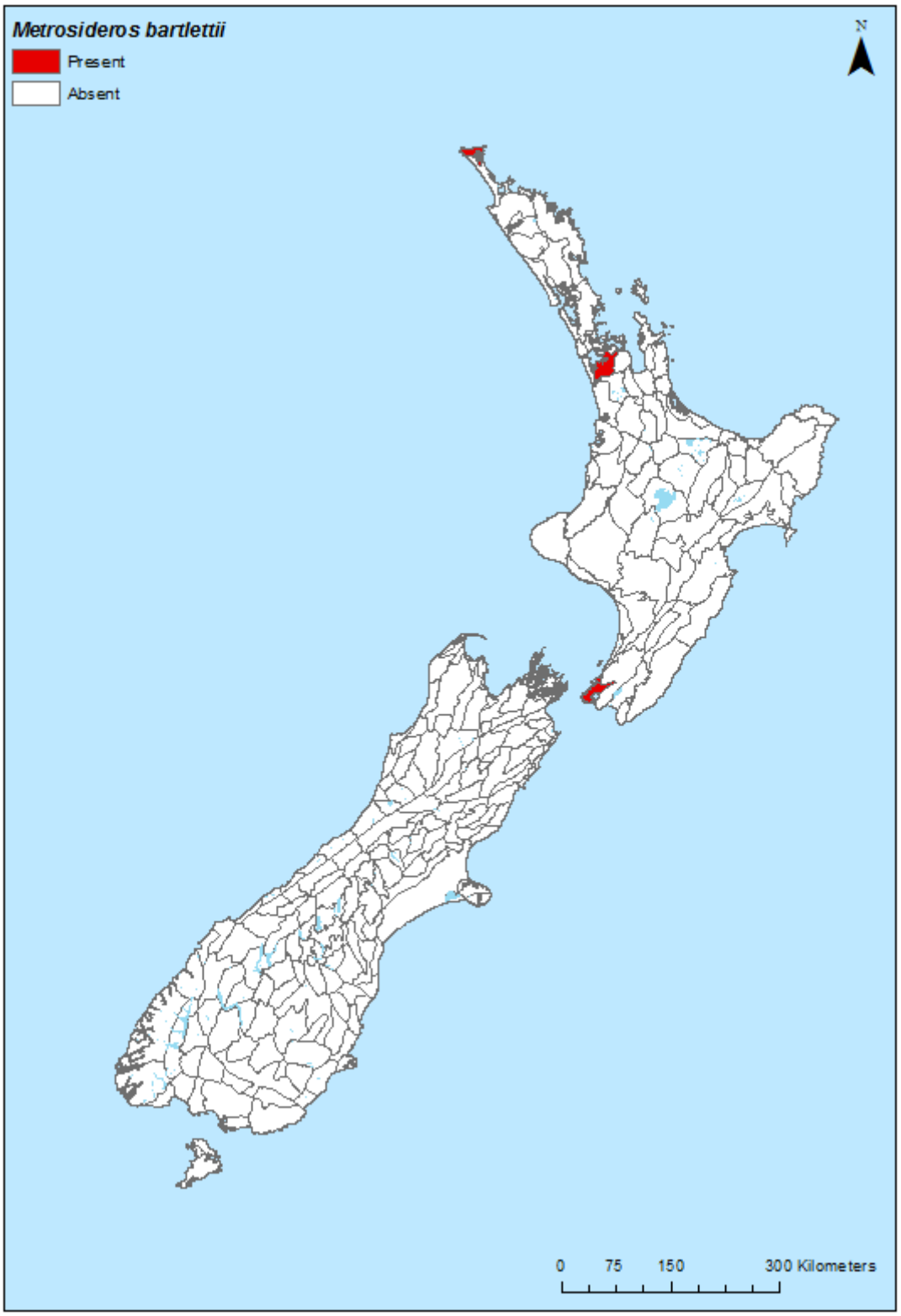

Figure 25 ***Metrosideros bartletii*** **– presence/absence**

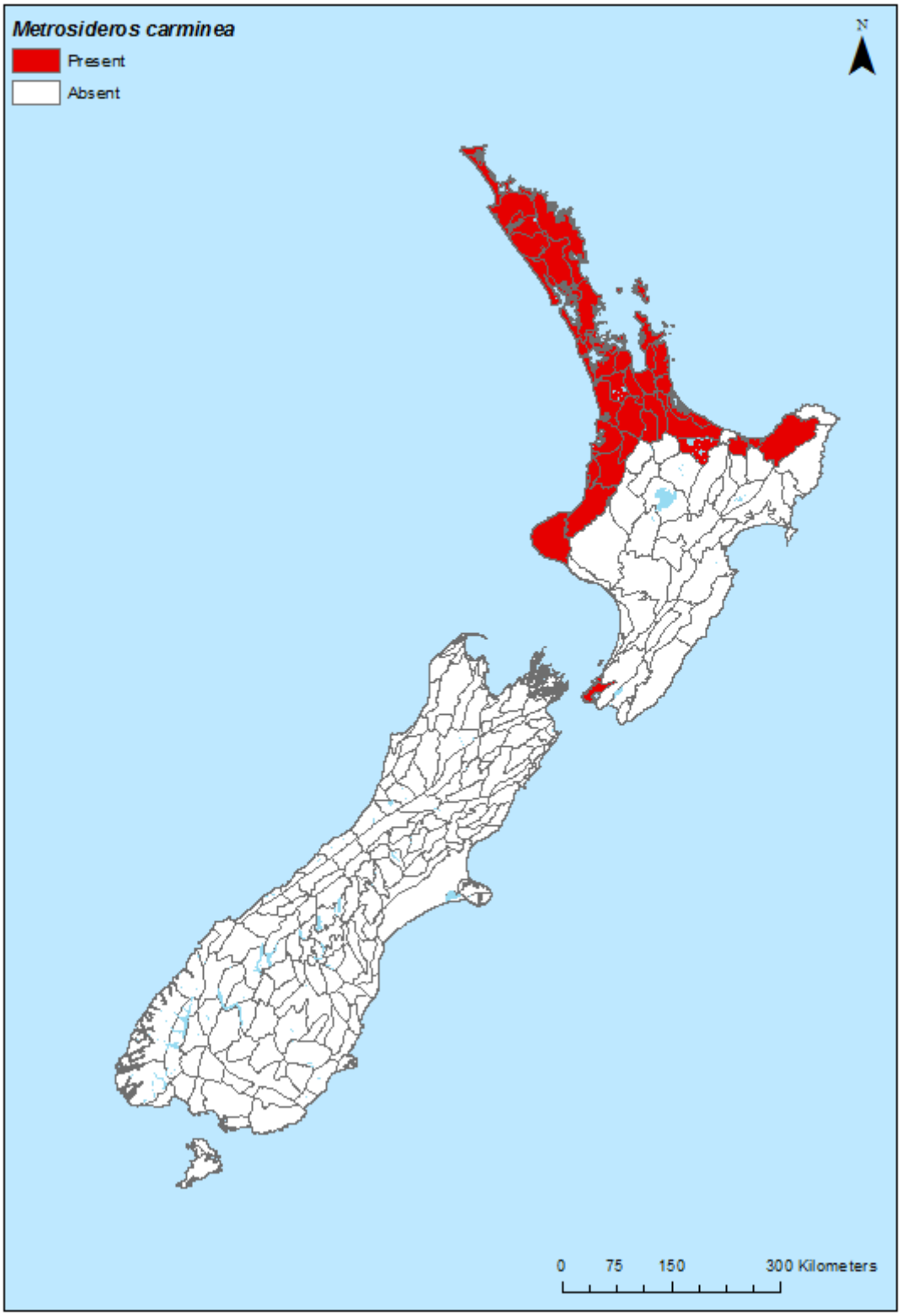

Figure 26 ***Metrosideros carminea*** **– presence/absence**

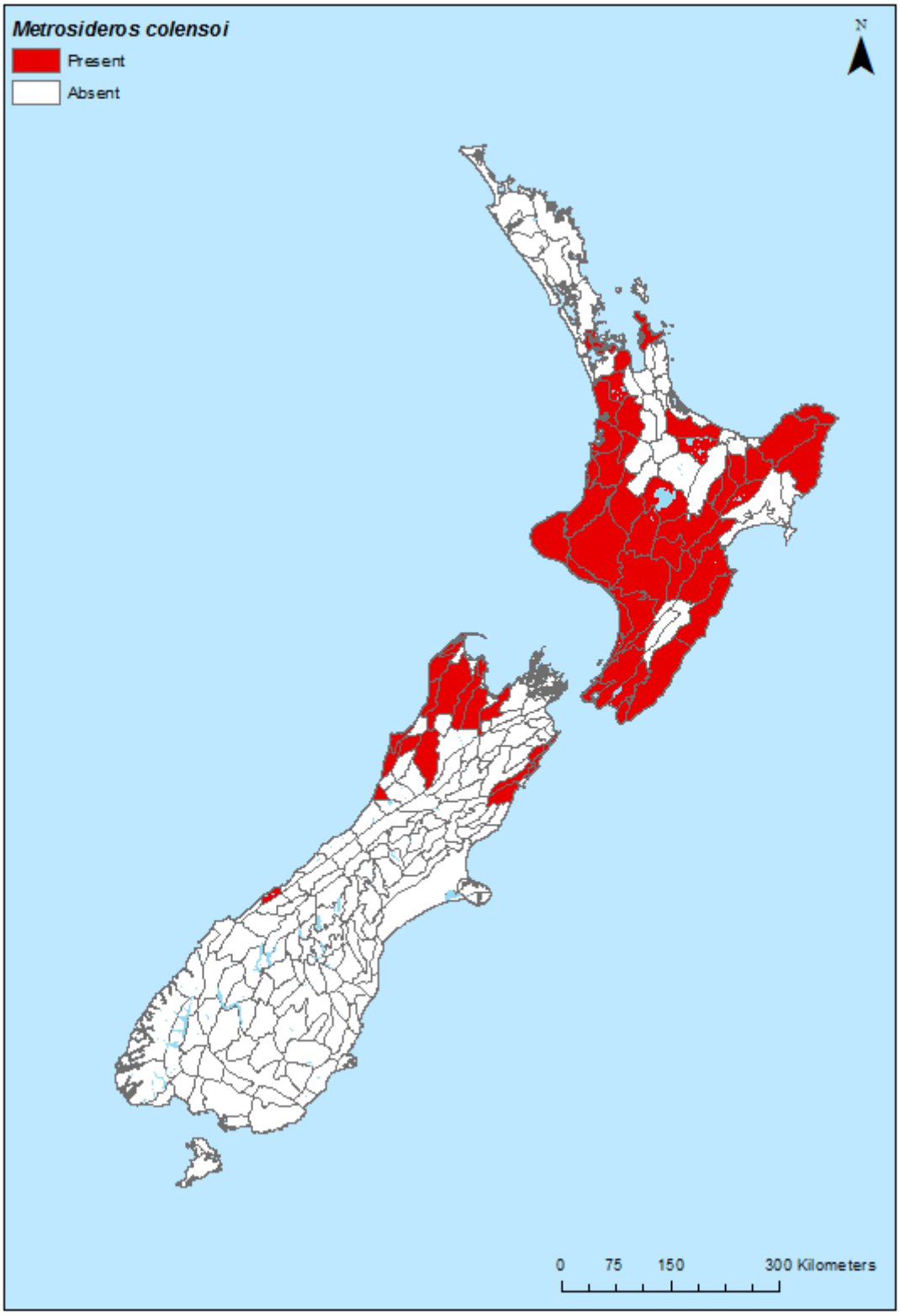

Figure 27 ***Metrosideros colensoi*** **– presence/absence**

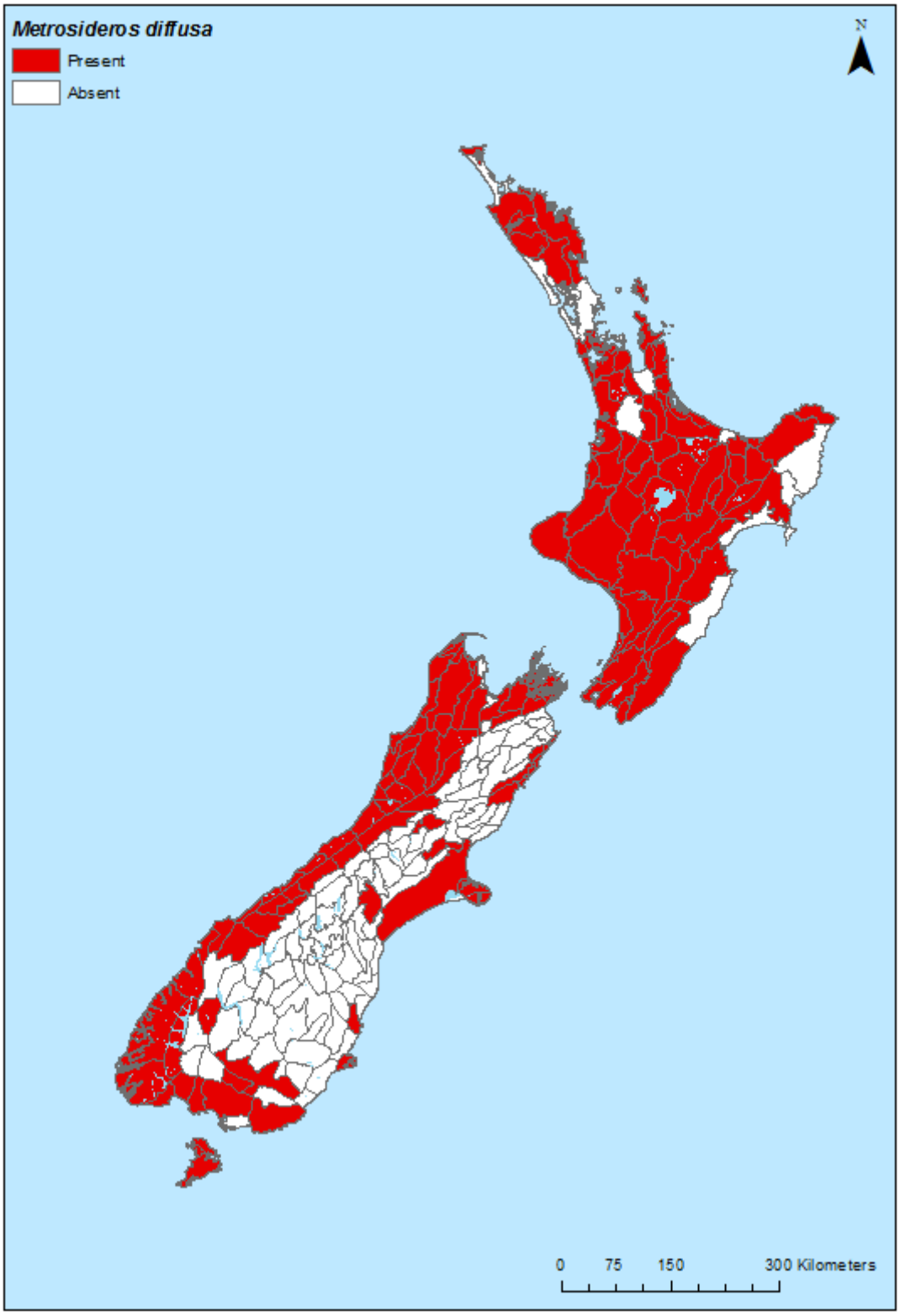

Figure 28 ***Metrosideros diffusa*** **– presence/absence**

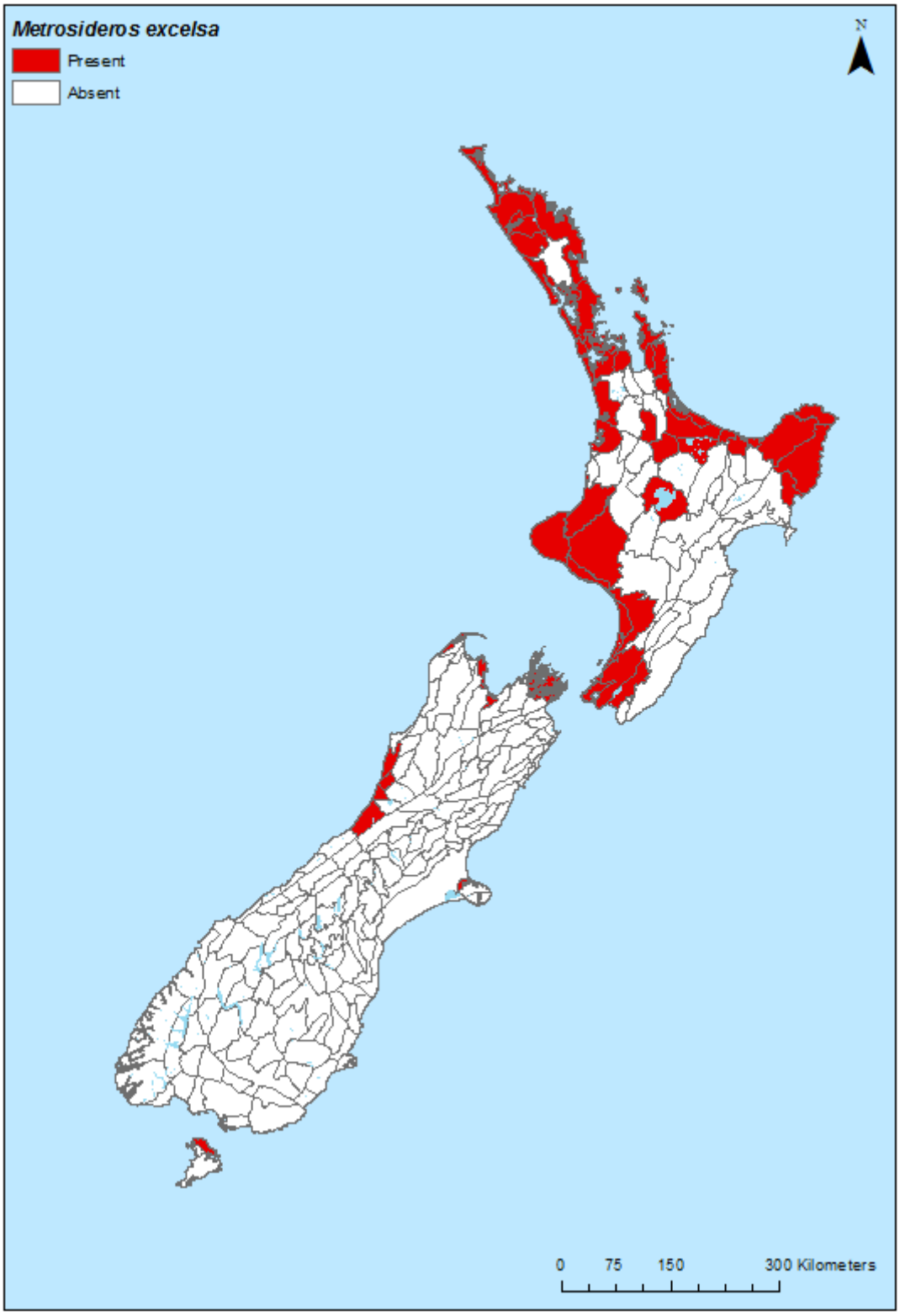

Figure 29 ***Metrosideros excelsa*** **– presence/absence**

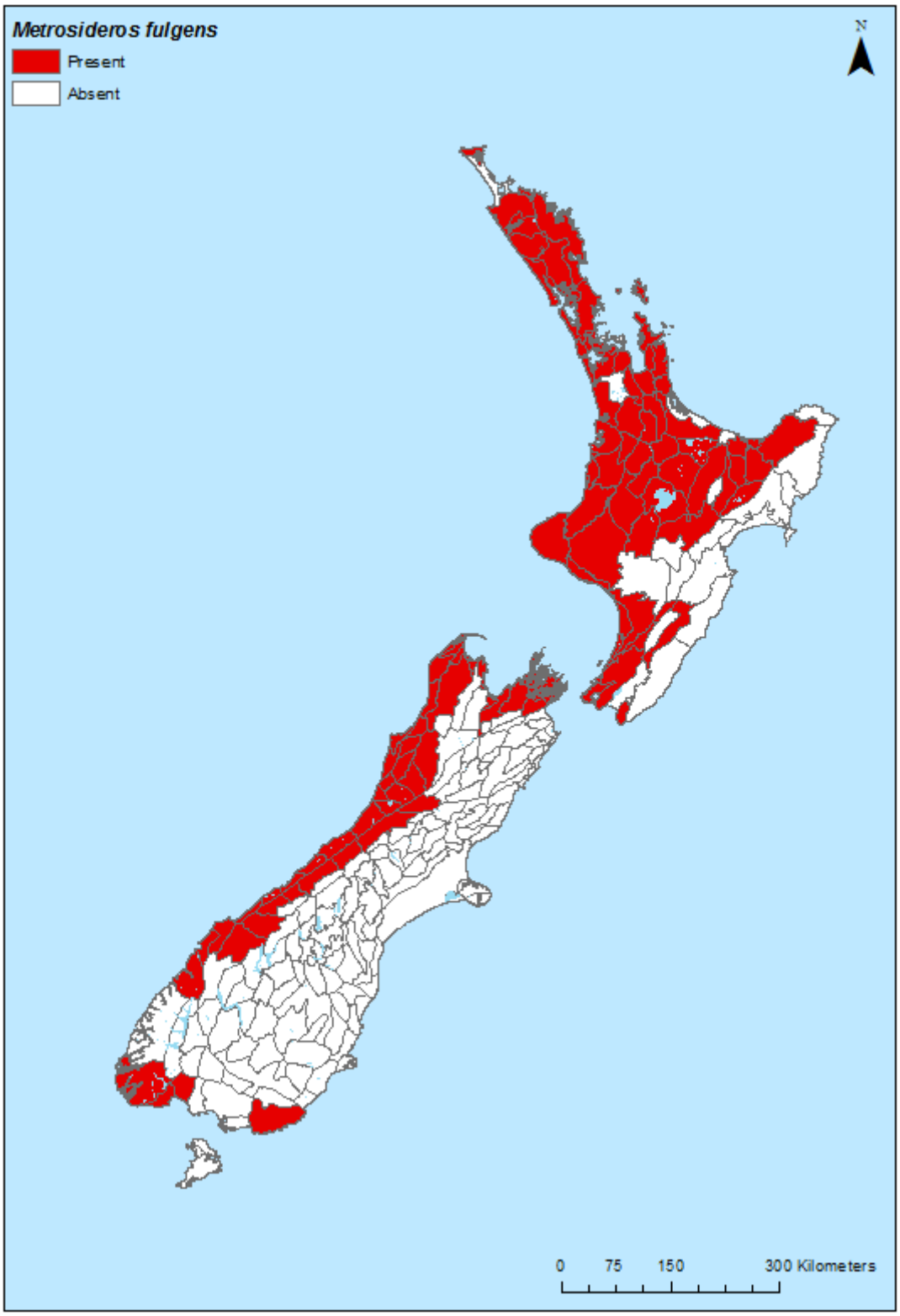

Figure 30 *Metrosideros fulgens* – presence/absence

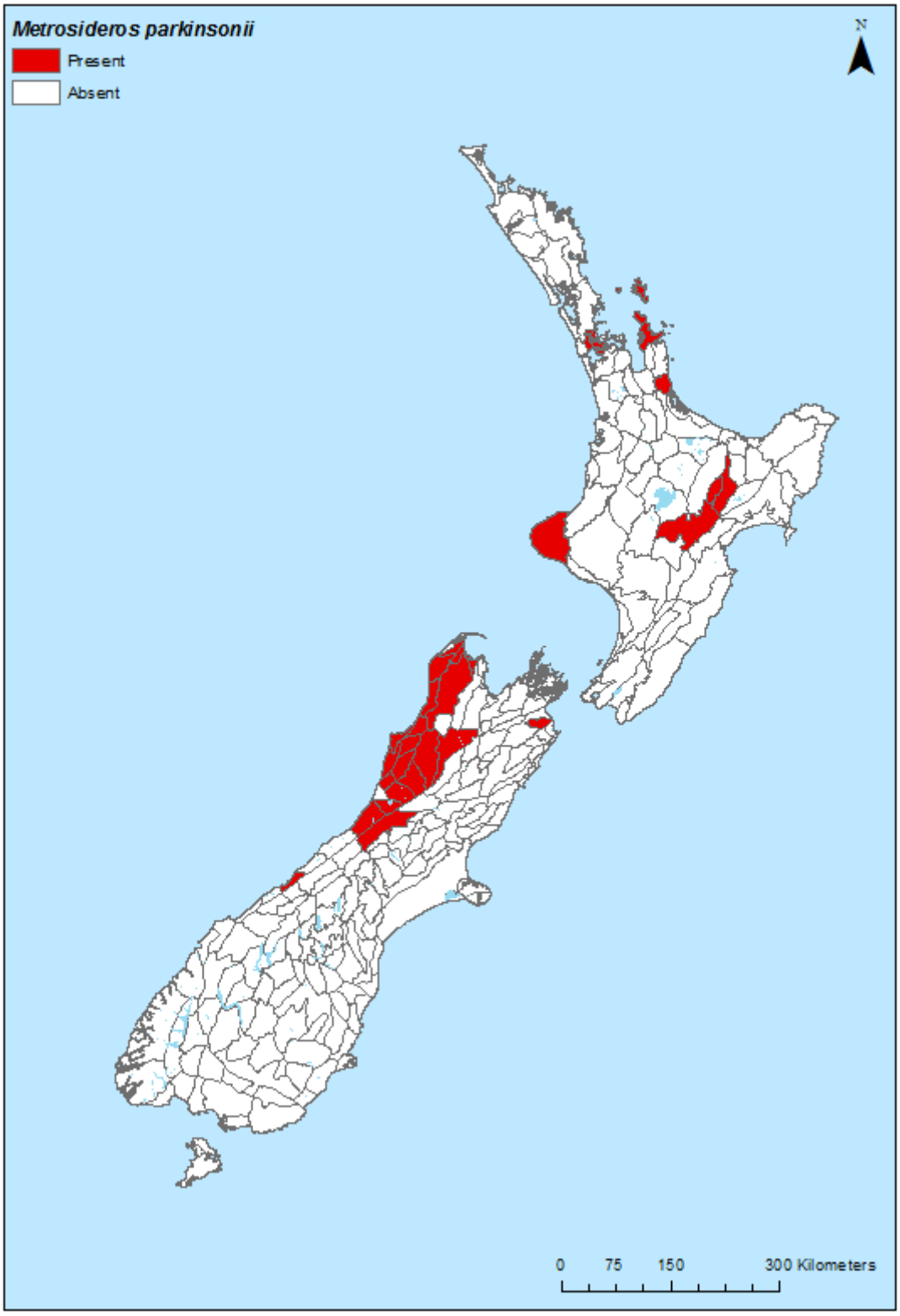

Figure 31 ***Metrosideros parkinsonii*** **– presence/absence**

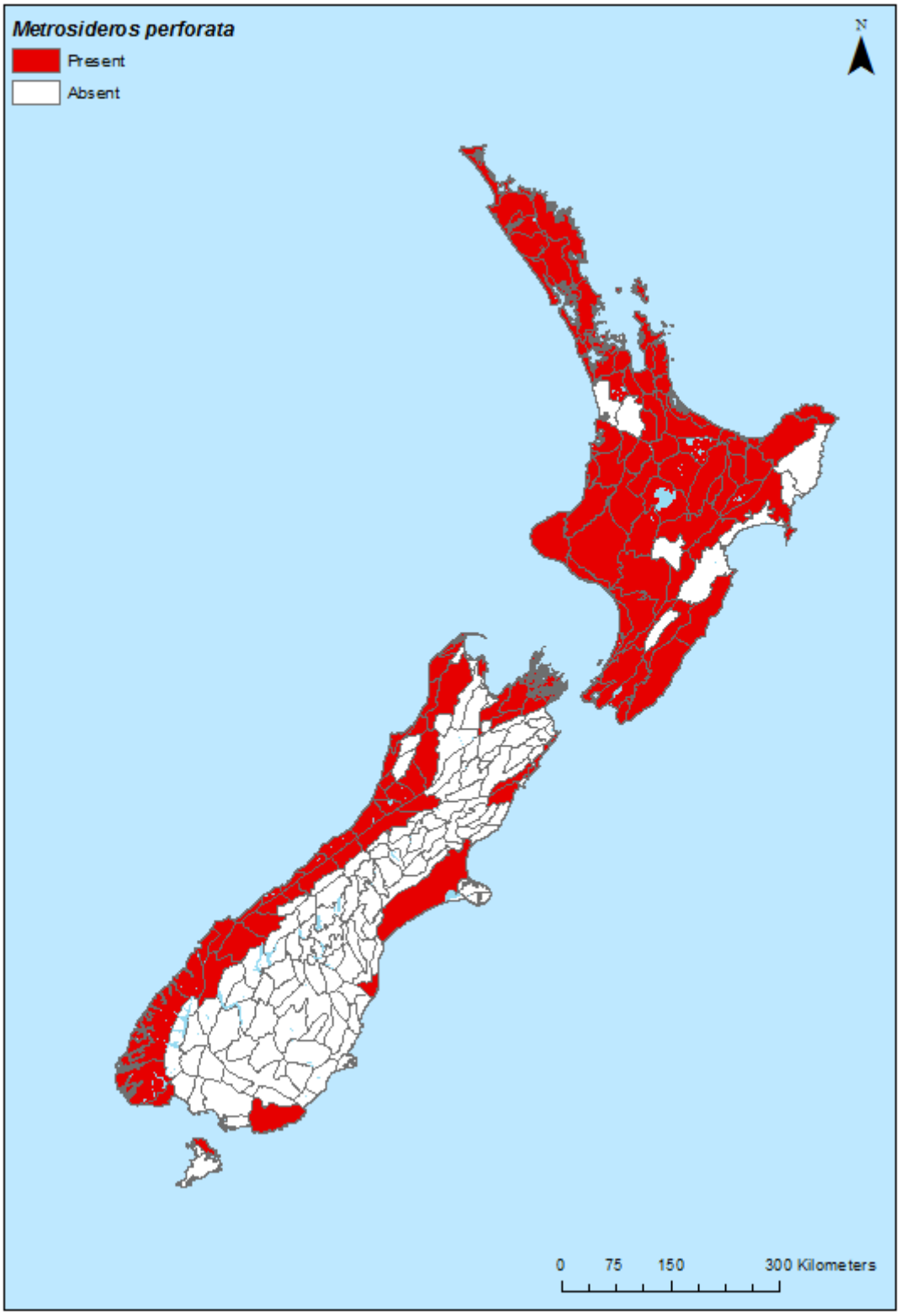

Figure 32 ***Metrosideros perforata*** **– presence/absence**

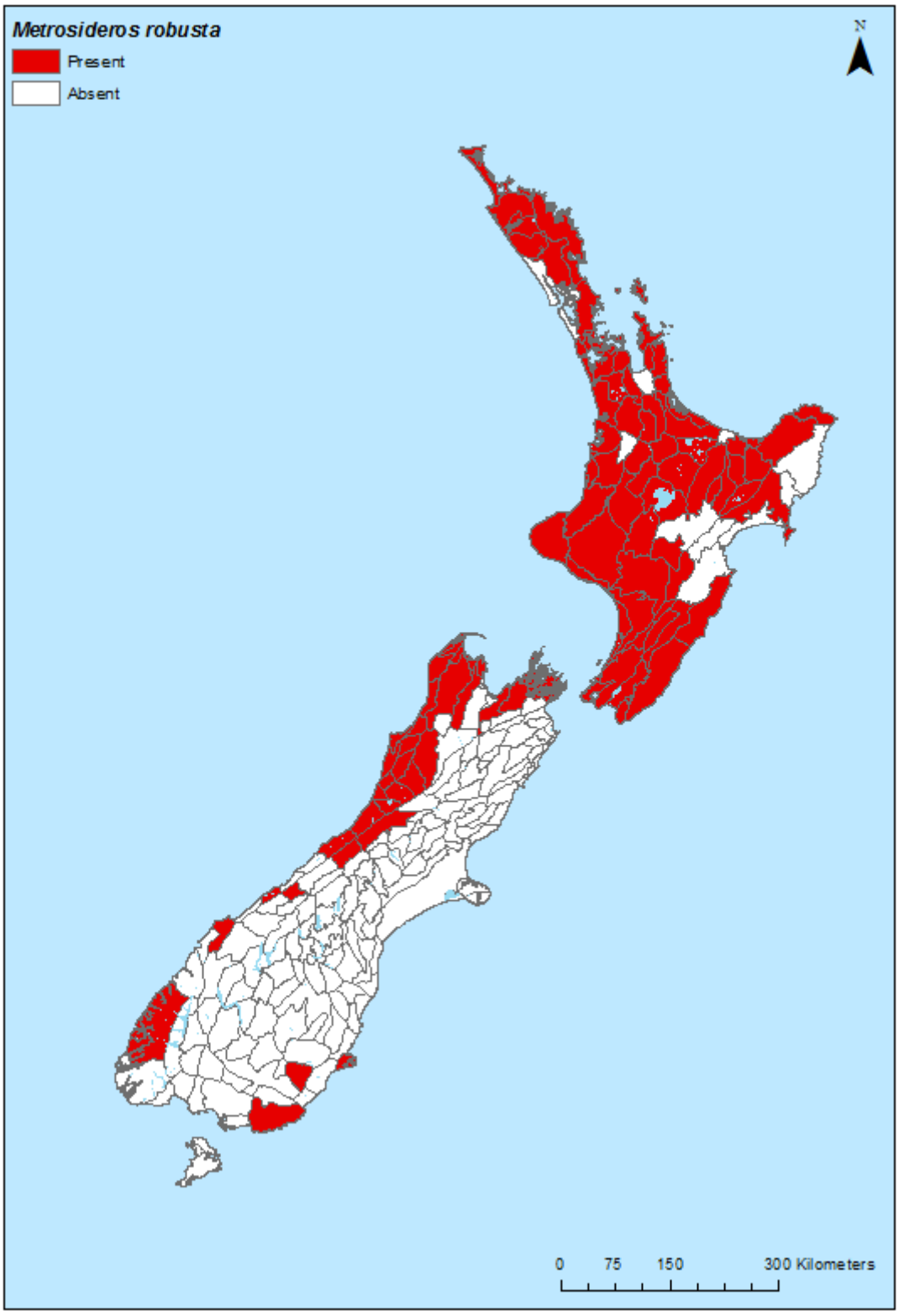

Figure 33 ***Metrosideros robusta*** **– presence/absence**

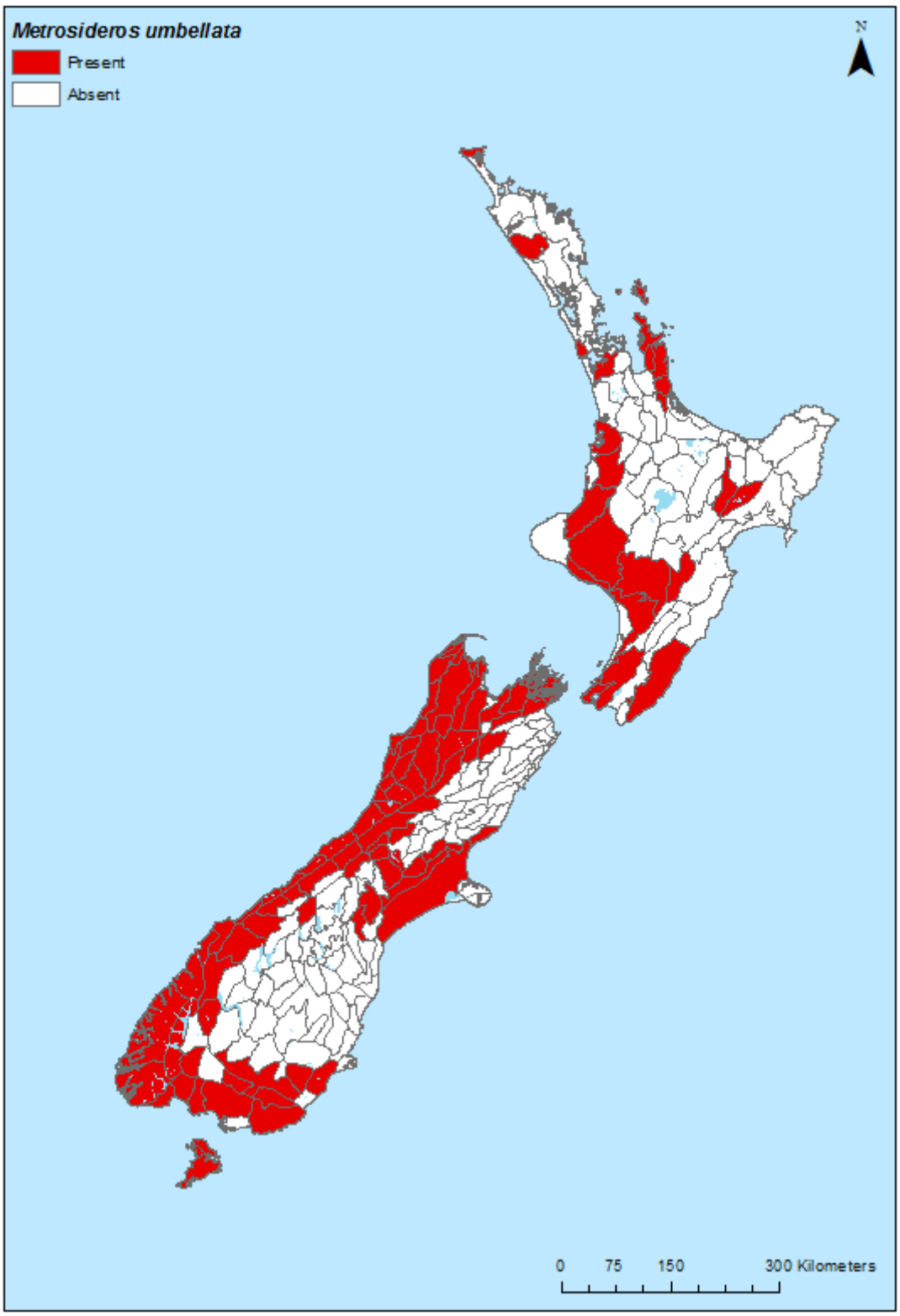

Figure 34 *Metrosideros umbellata* – presence/absence

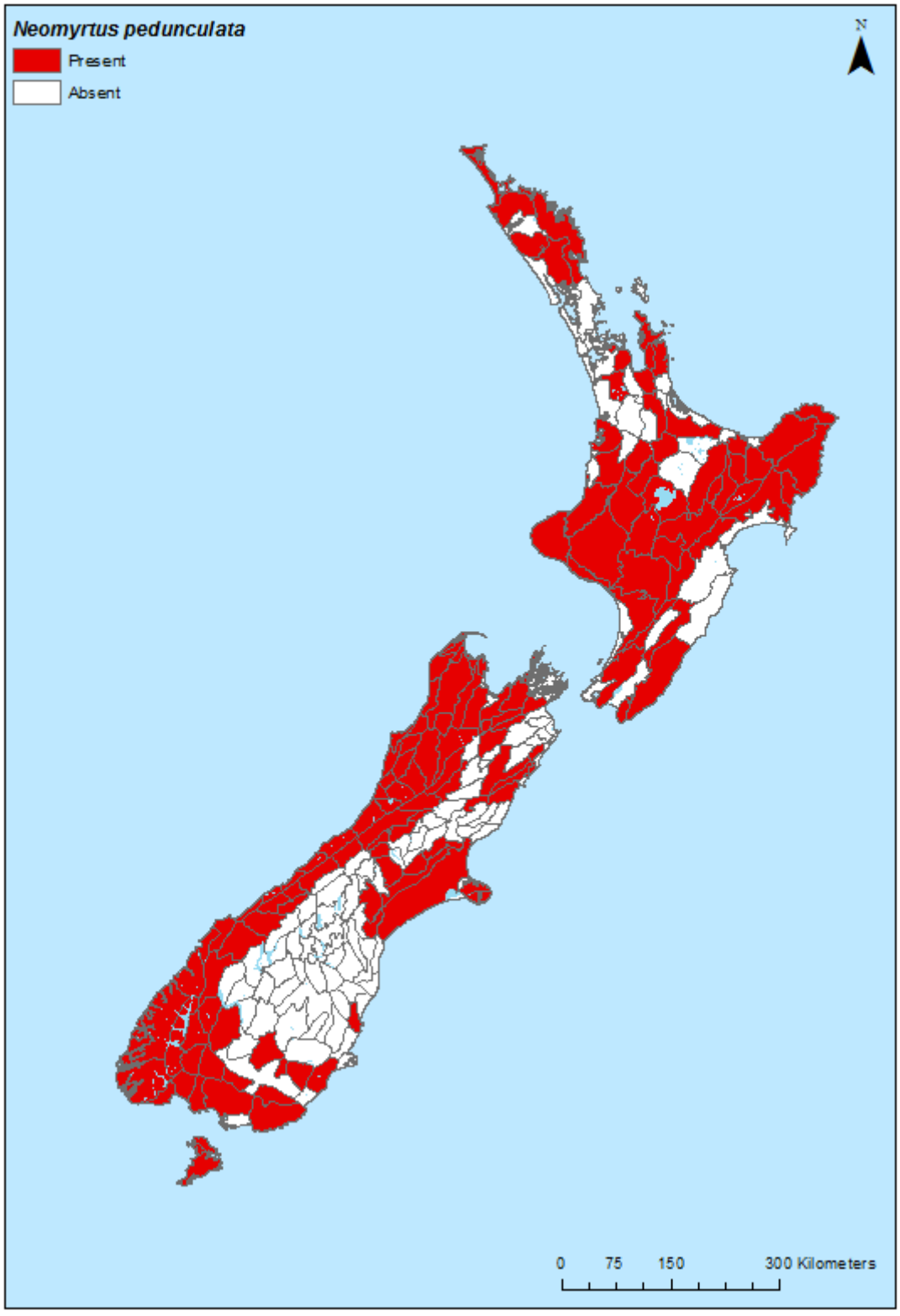

Figure 35 ***Neomyrtus pedunculata*** **– presence/absence**

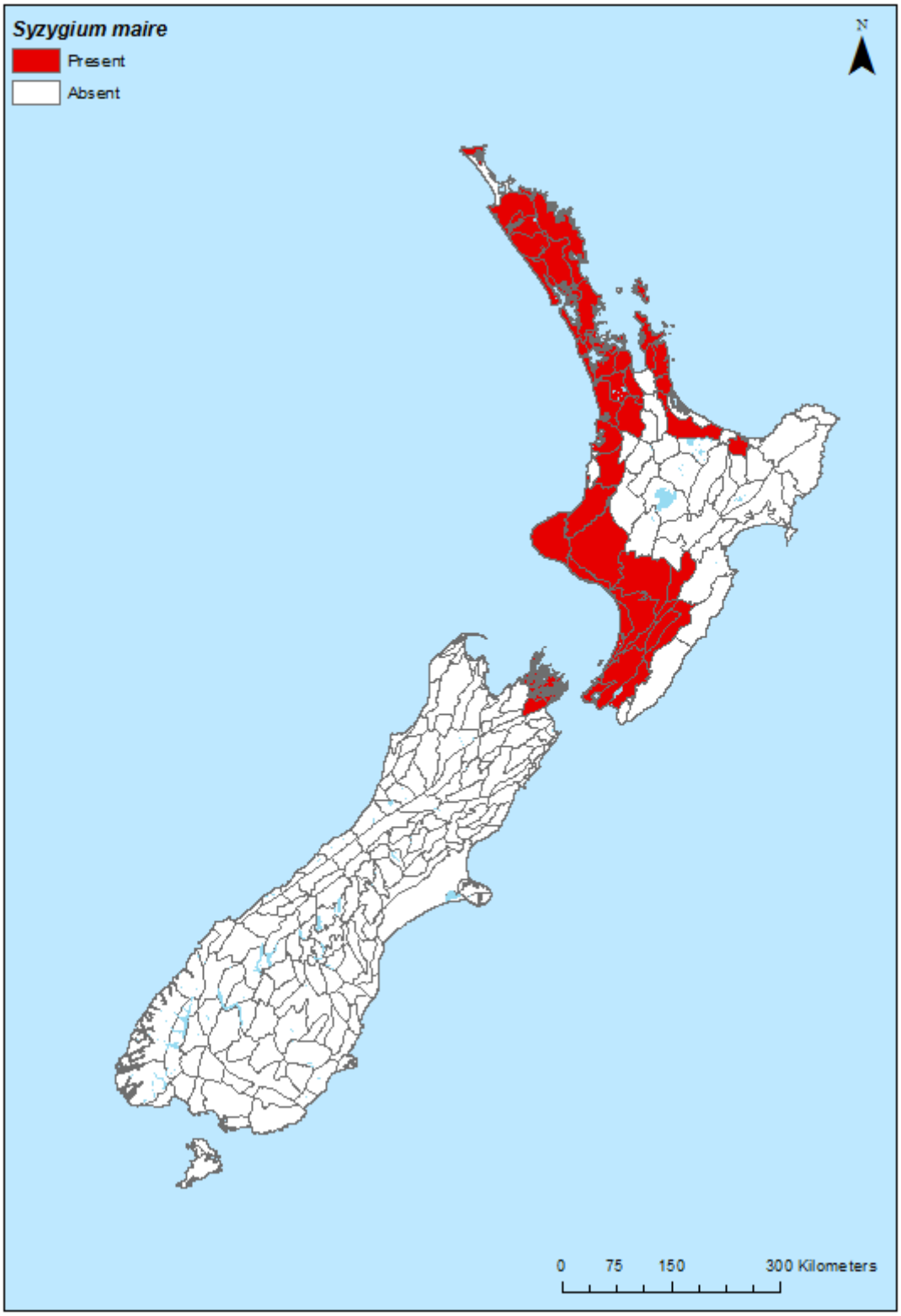

Figure 36 *Syzygium maire* – presence/absence

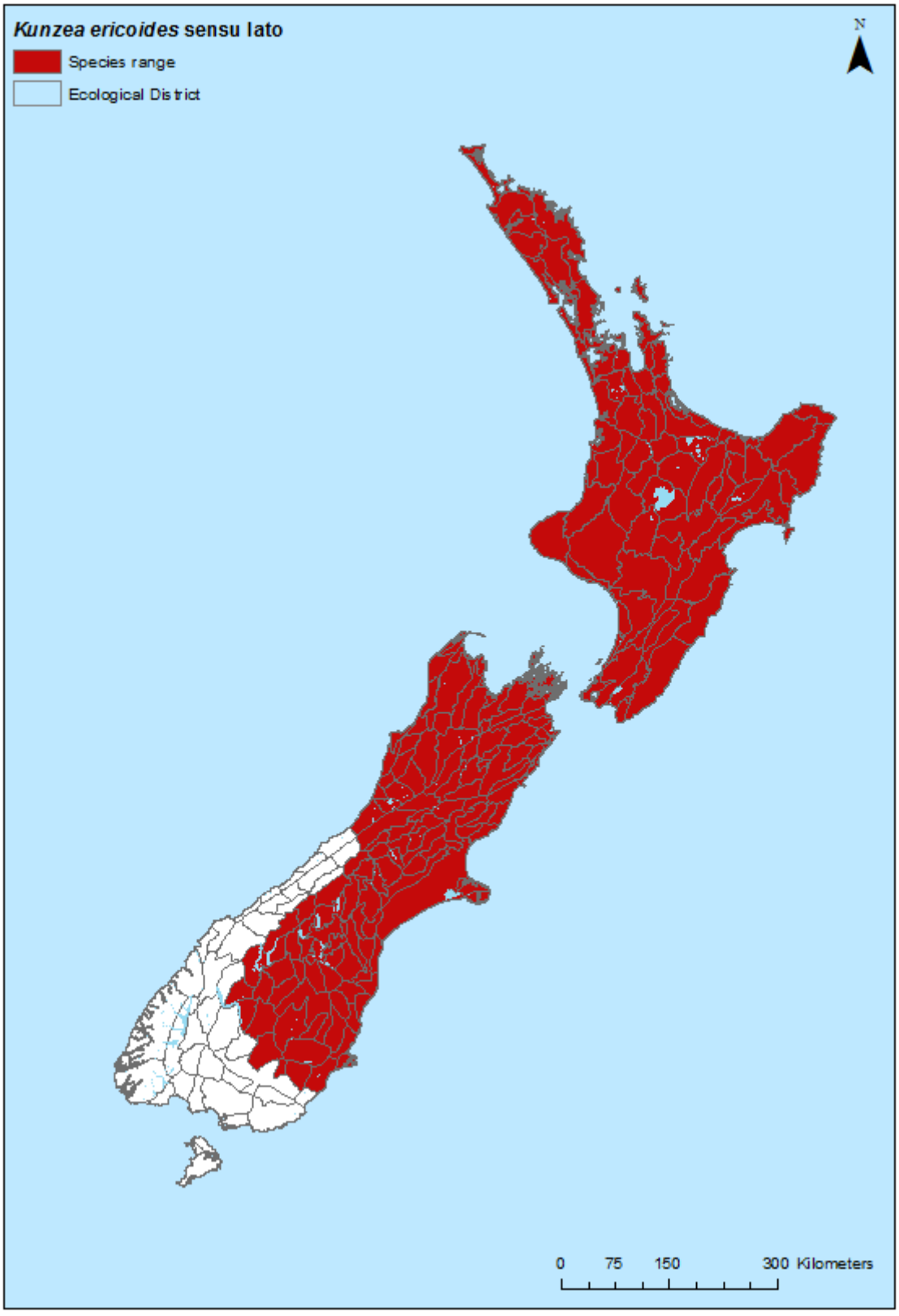

Figure 37 ***Kunzea ericoides sensu lato*** **– range**

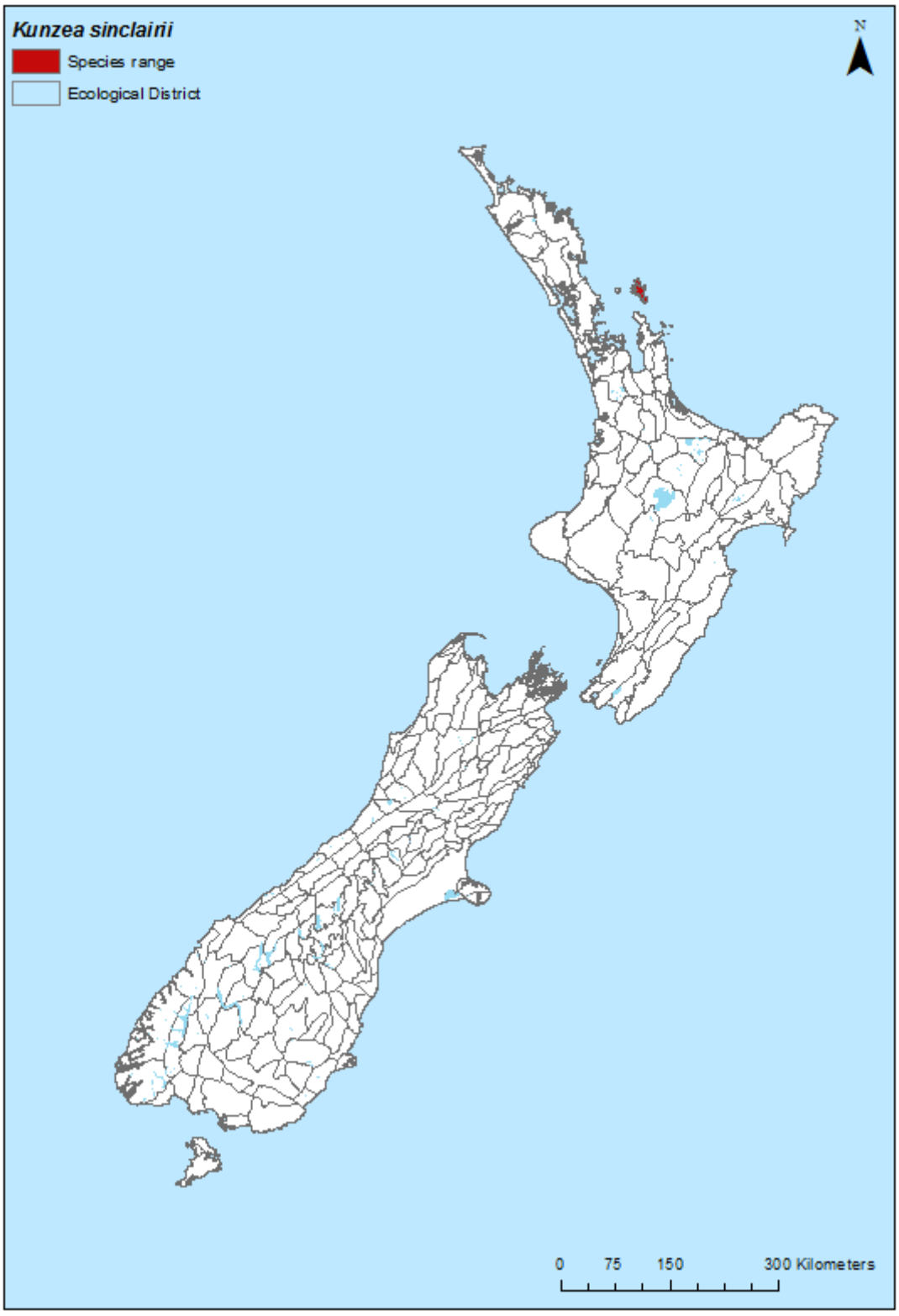

Figure 38 *Kunzea sinclairii* – range

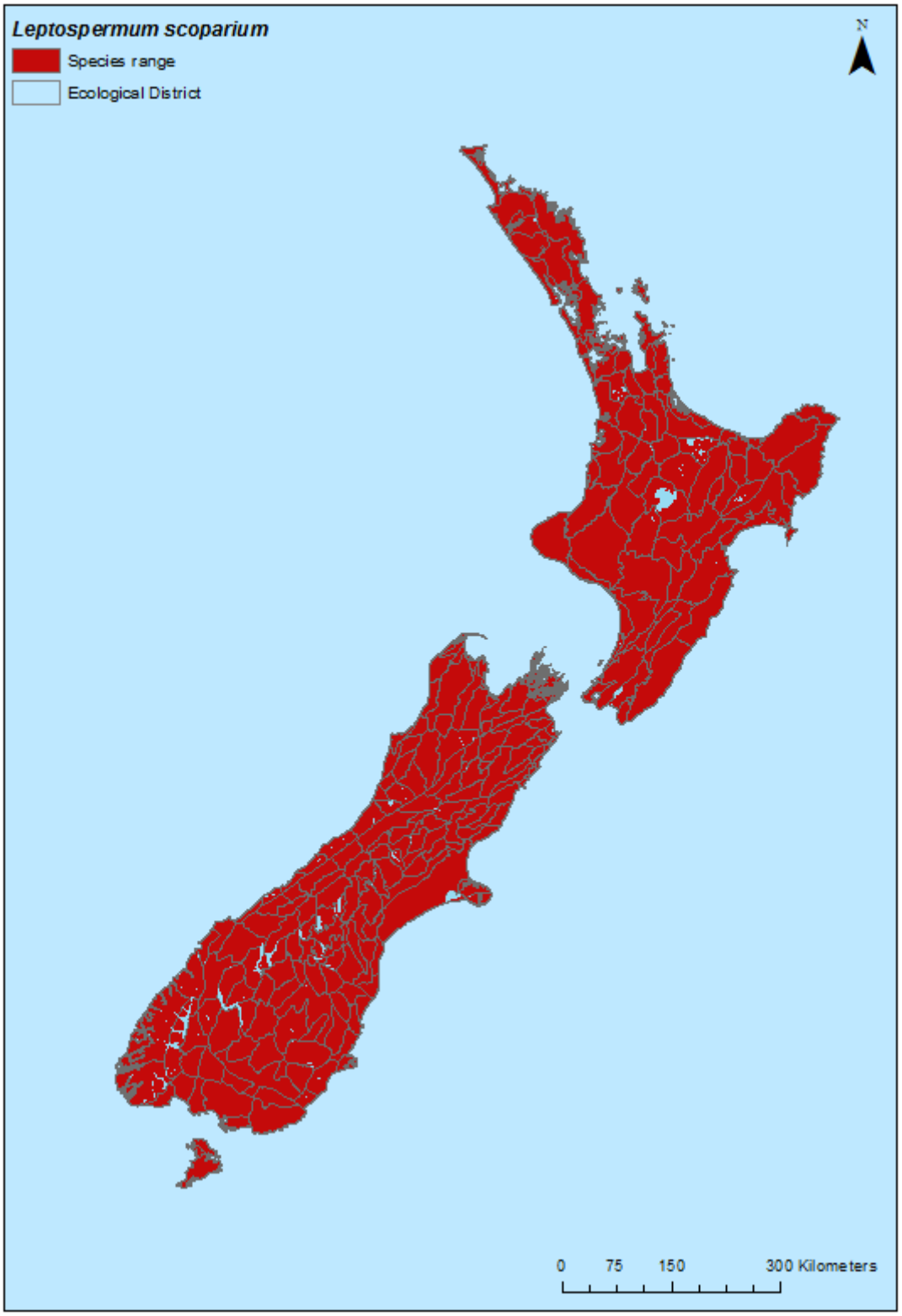

Figure 39 ***Leptospermum scoparium*** **– range**

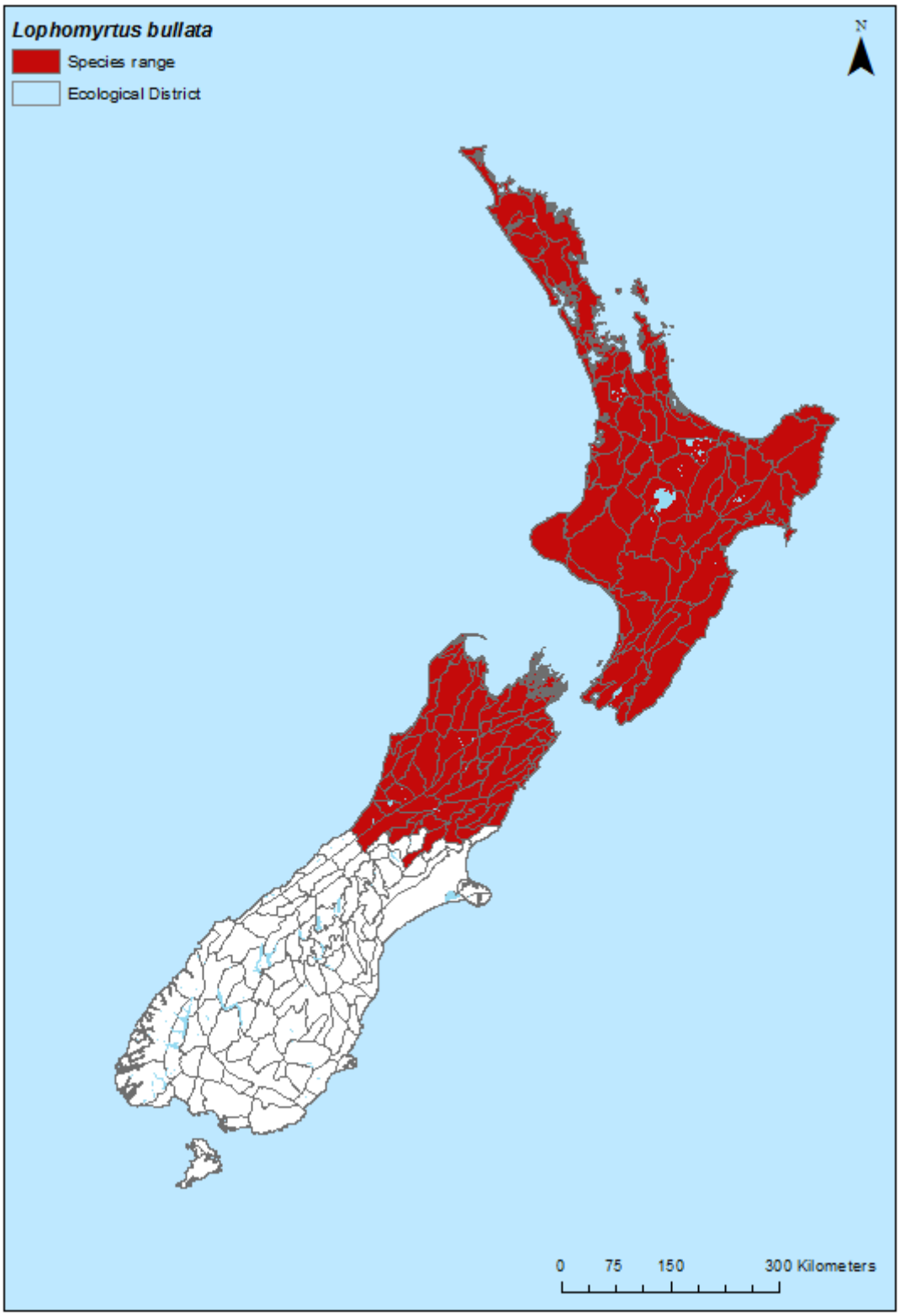

Figure 40 ***Lophomyrtus bullata*** **– range**

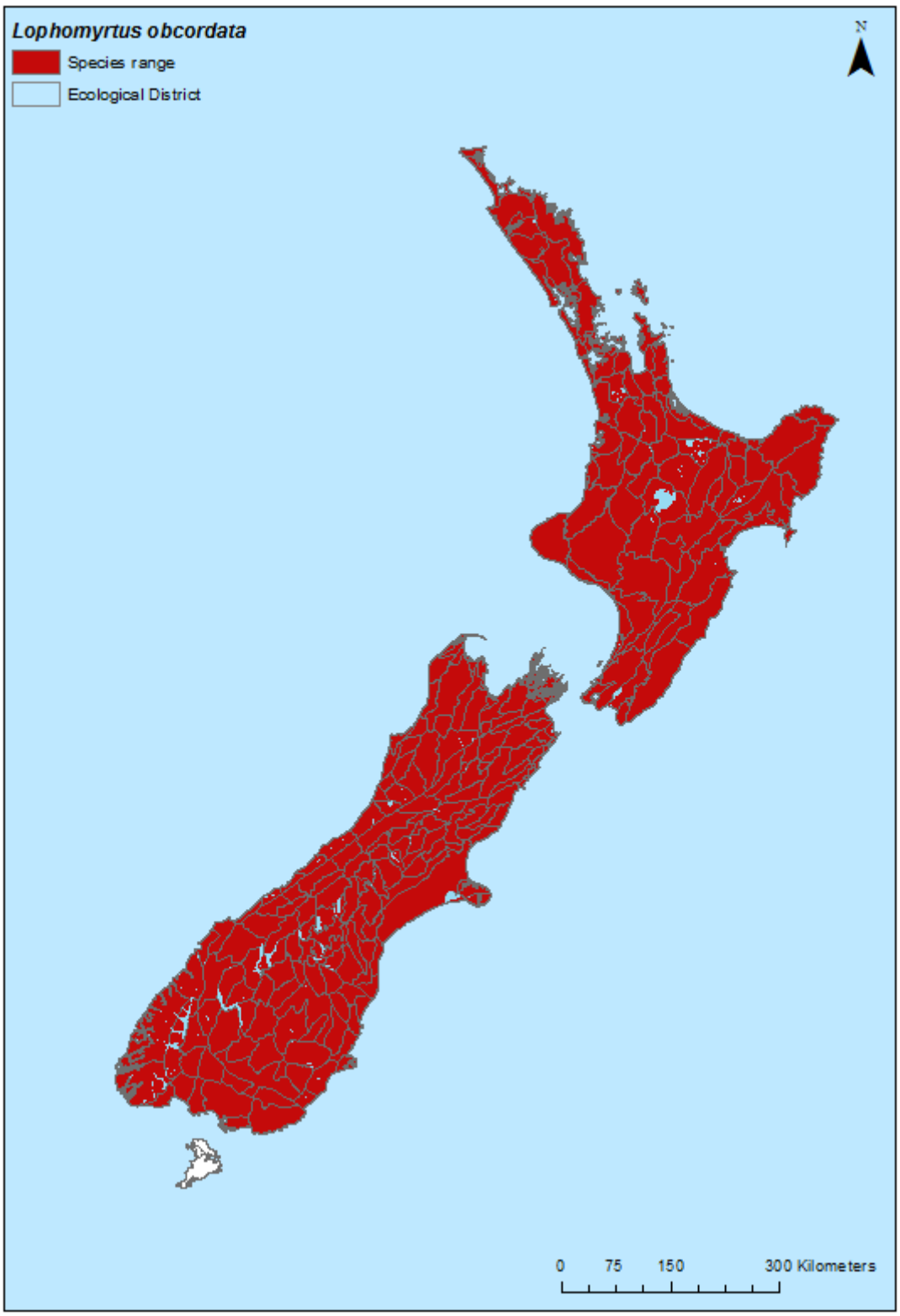

Figure 41 ***Lophomyrtus obcordata*** **– range**

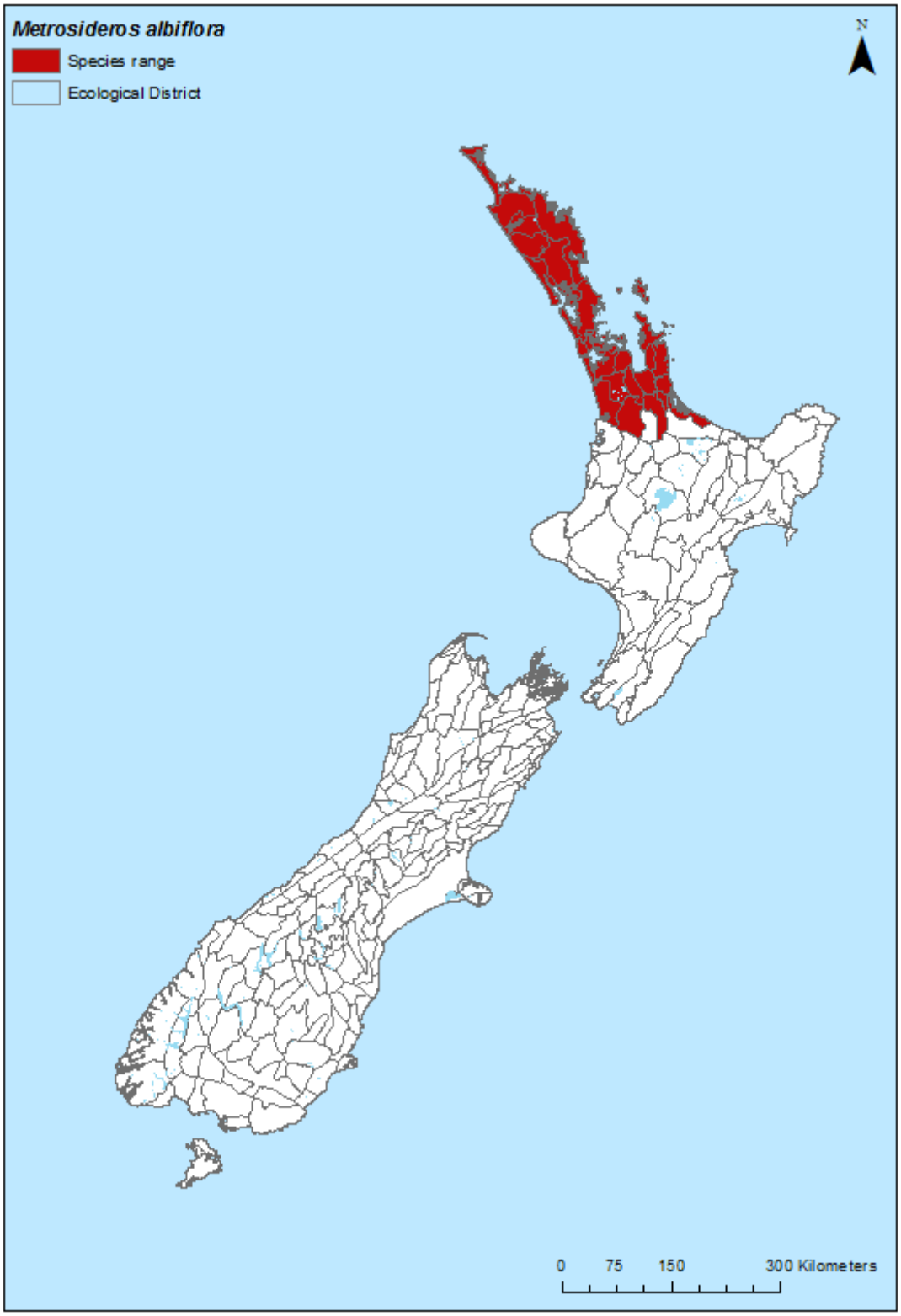

Figure 42 ***Metrosideros albiflora*** **– range**

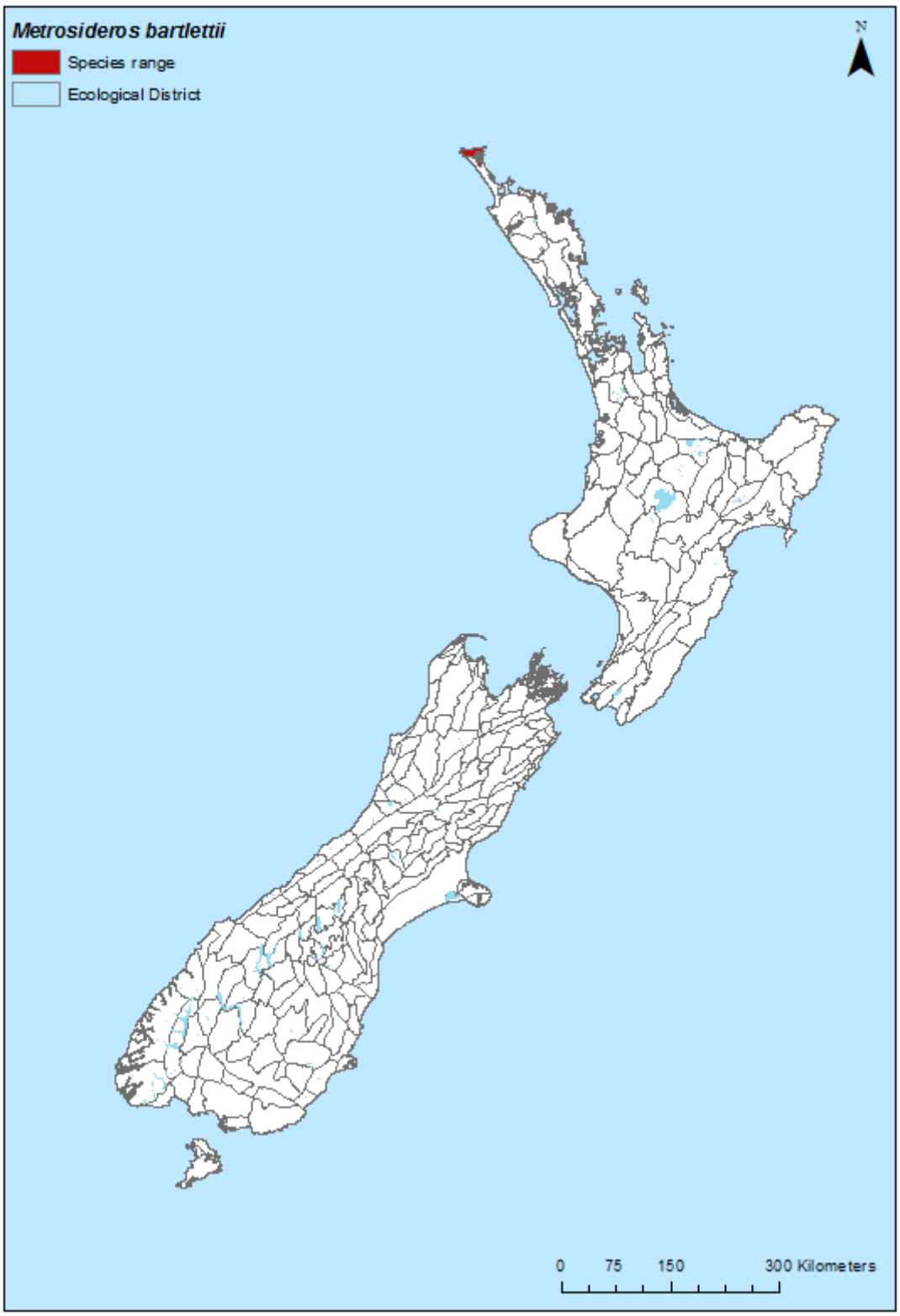

Figure 43 ***Metrosideros bartlettii*** **– range**

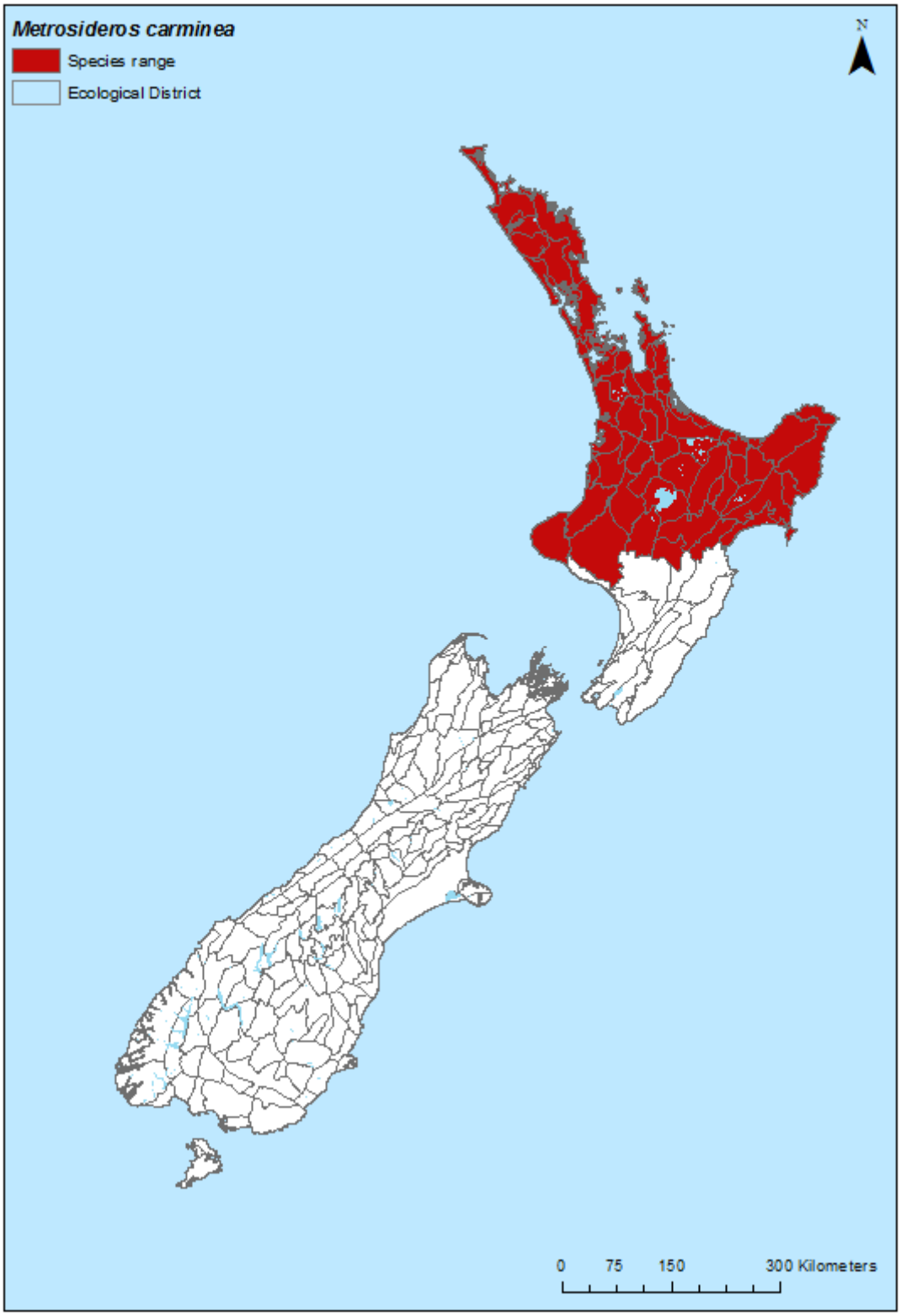

Figure 44 *Metrosideros carminea* – range

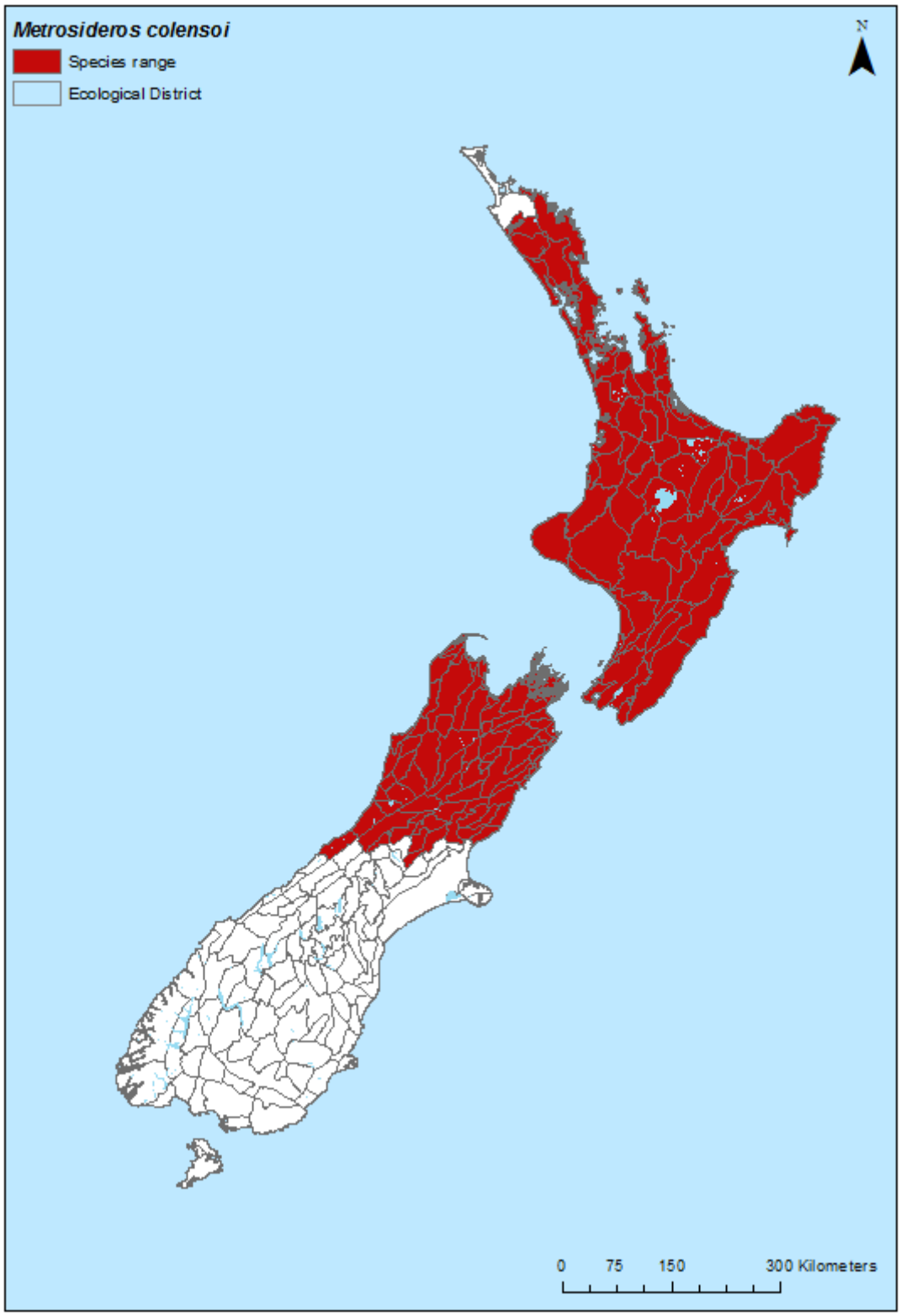

Figure 45 ***Metrosideros colensoi*** **– range**

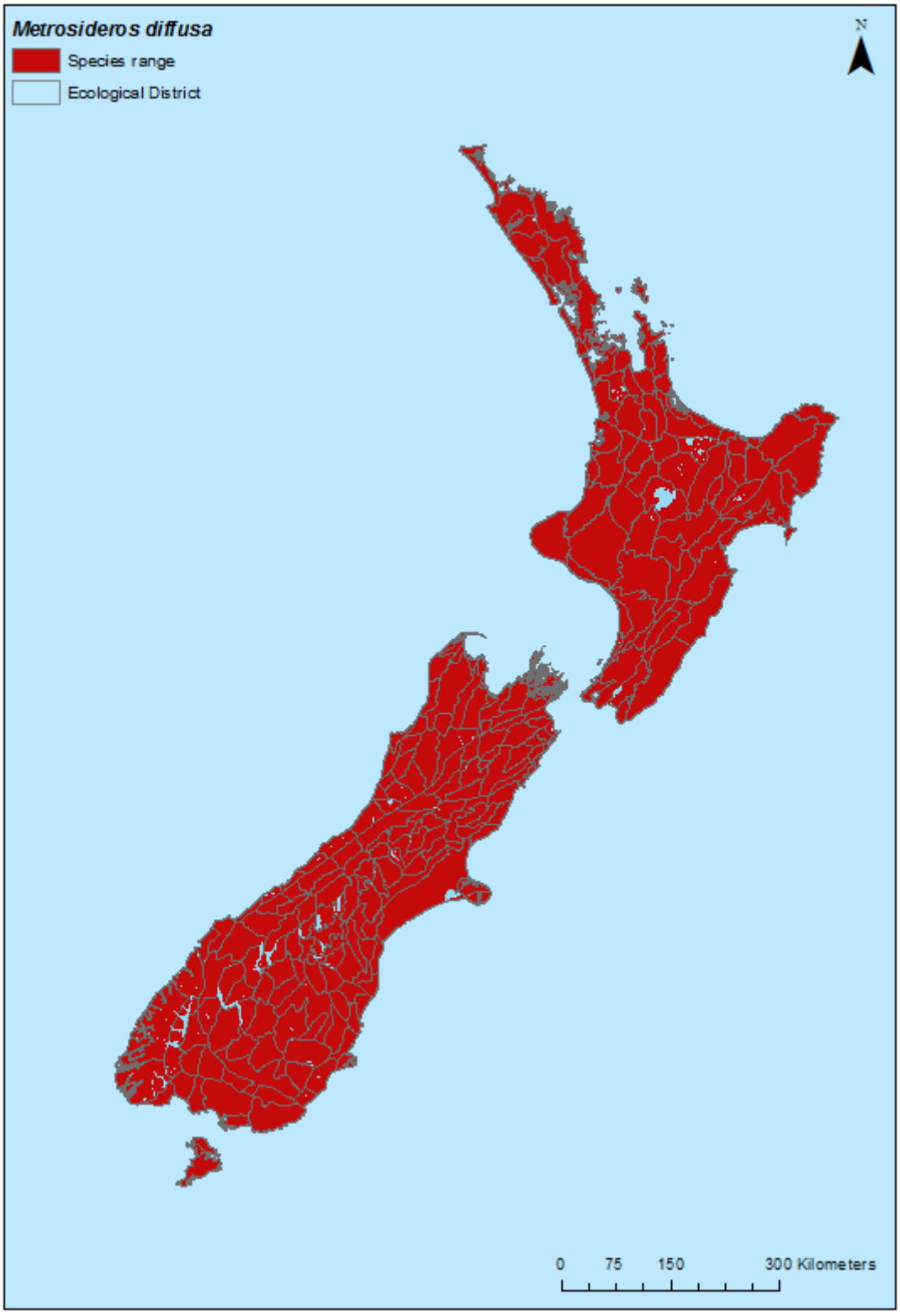

Figure 46 ***Metrosideros diffusa*** **– range**

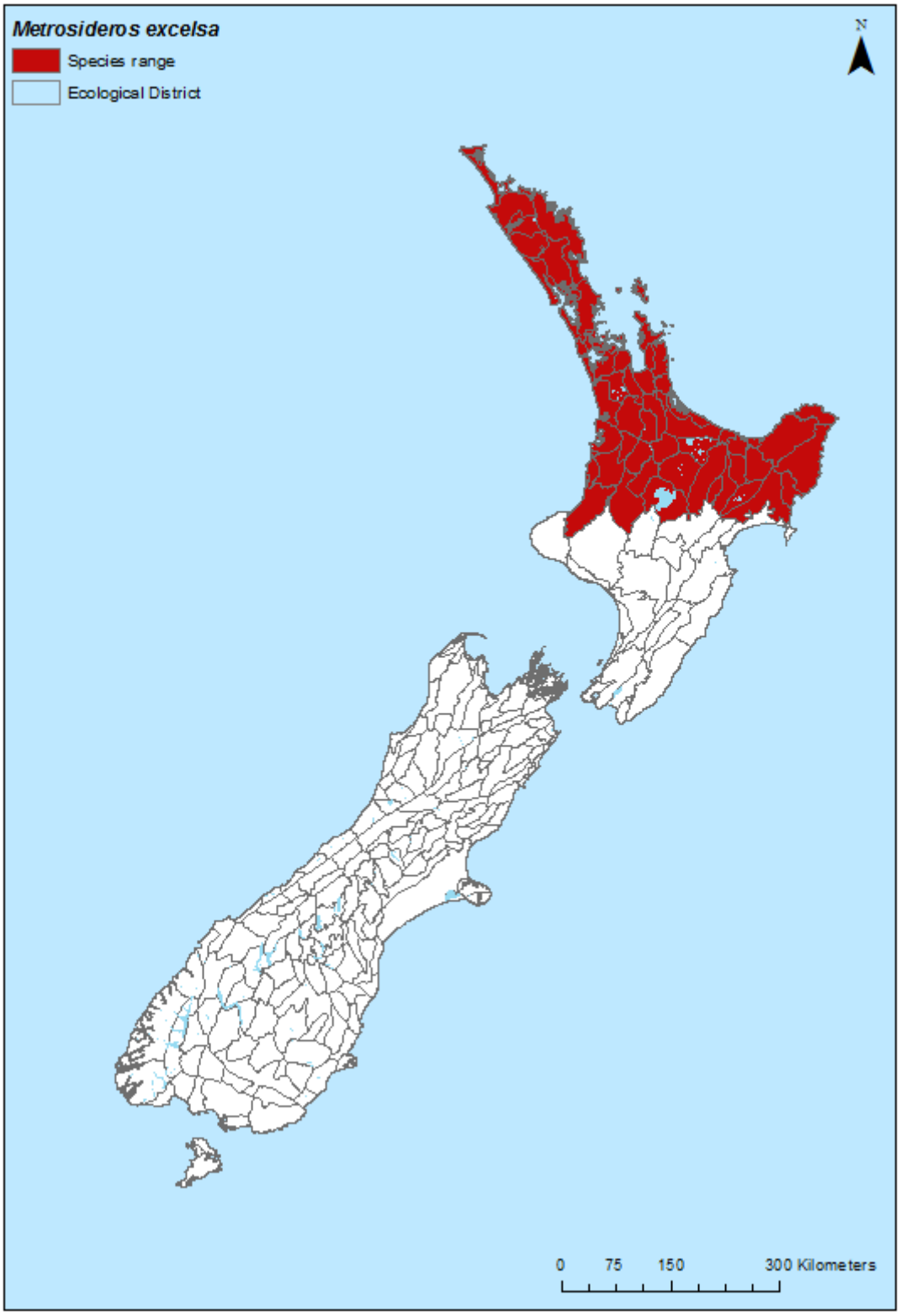

Figure 47 ***Metrosideros excelsa*** **– range**

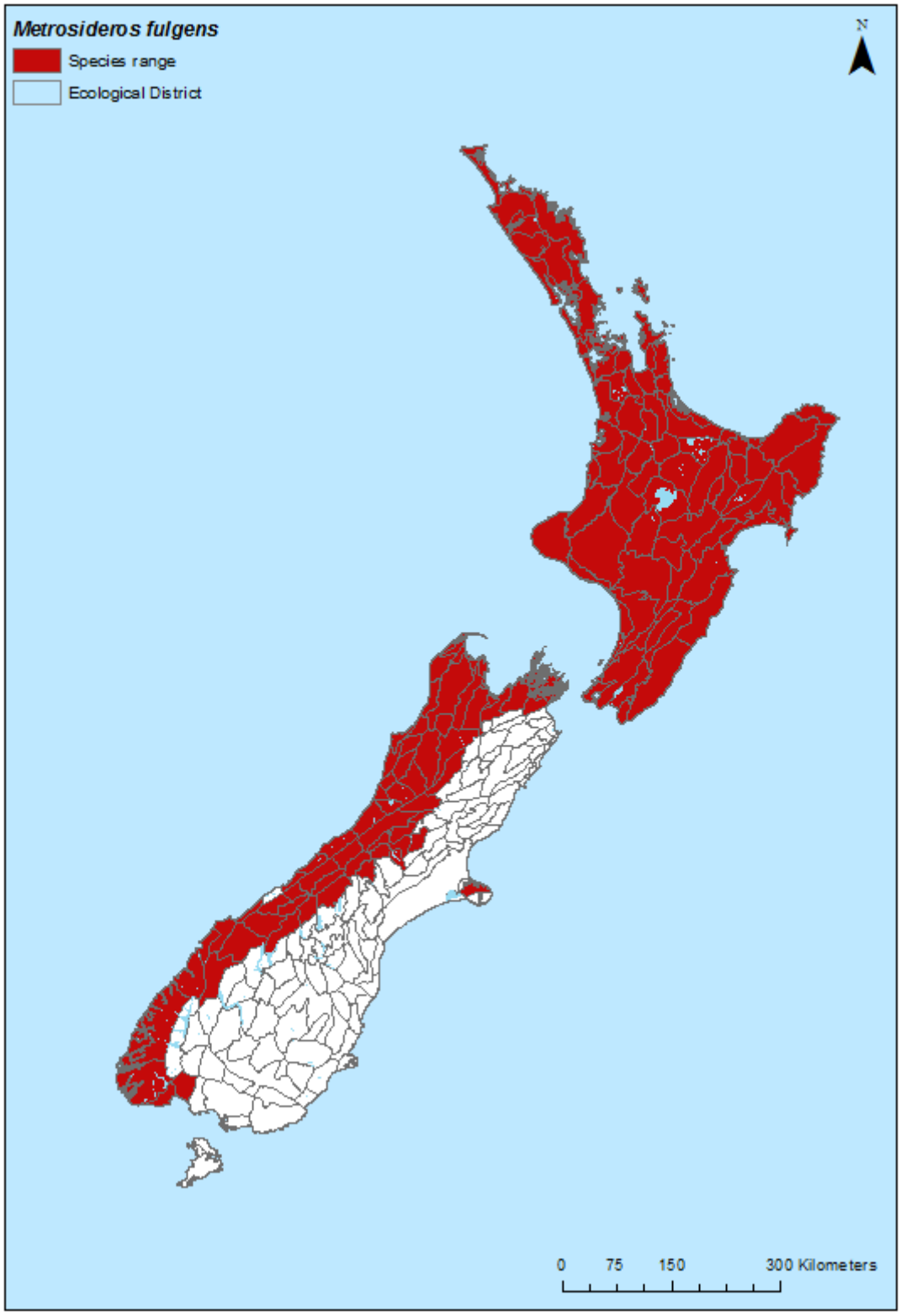

Figure 48 ***Metrosideros fulgens*** **– range**

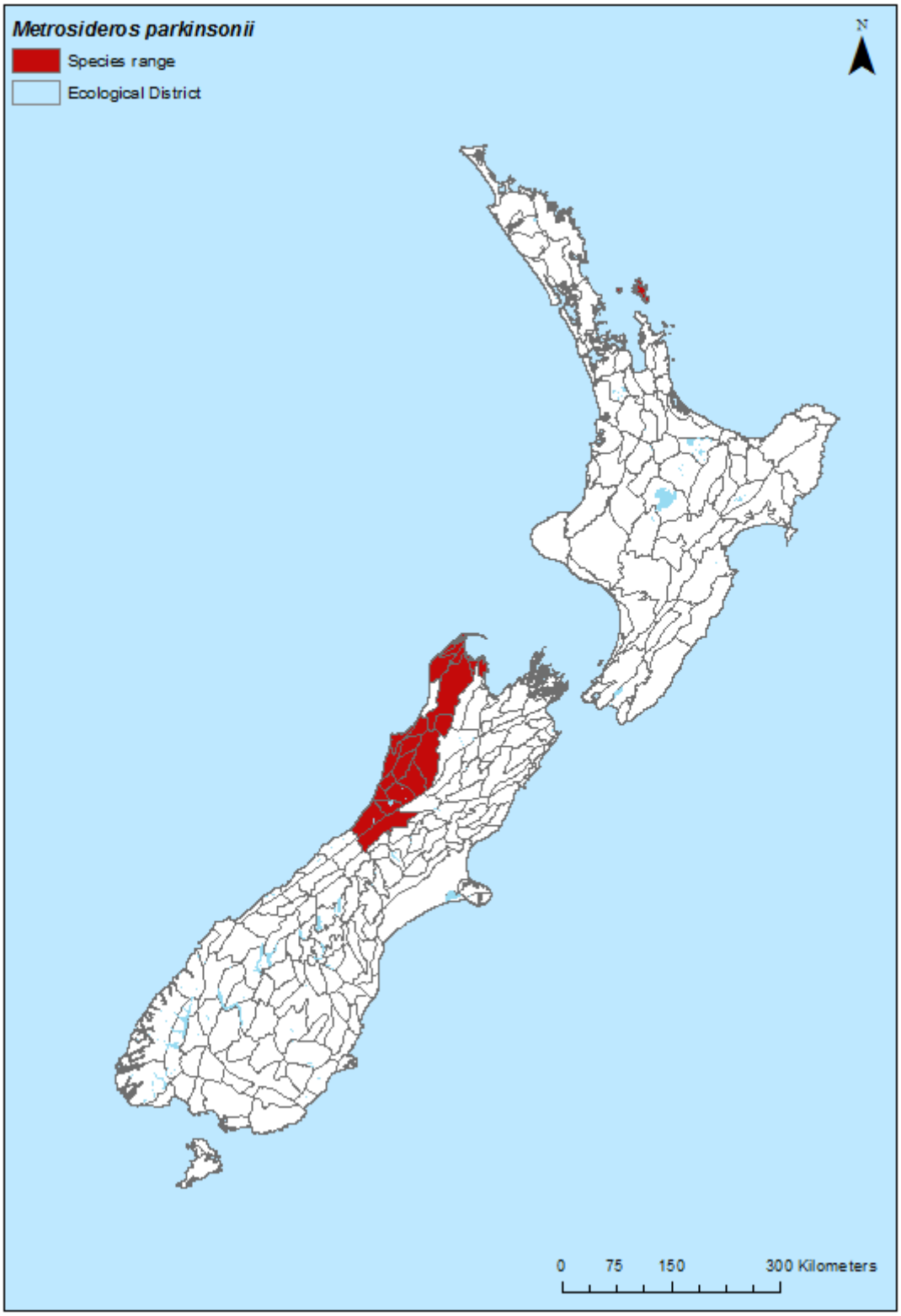

Figure 49 ***Metrosideros parkinsonii*** **– range**

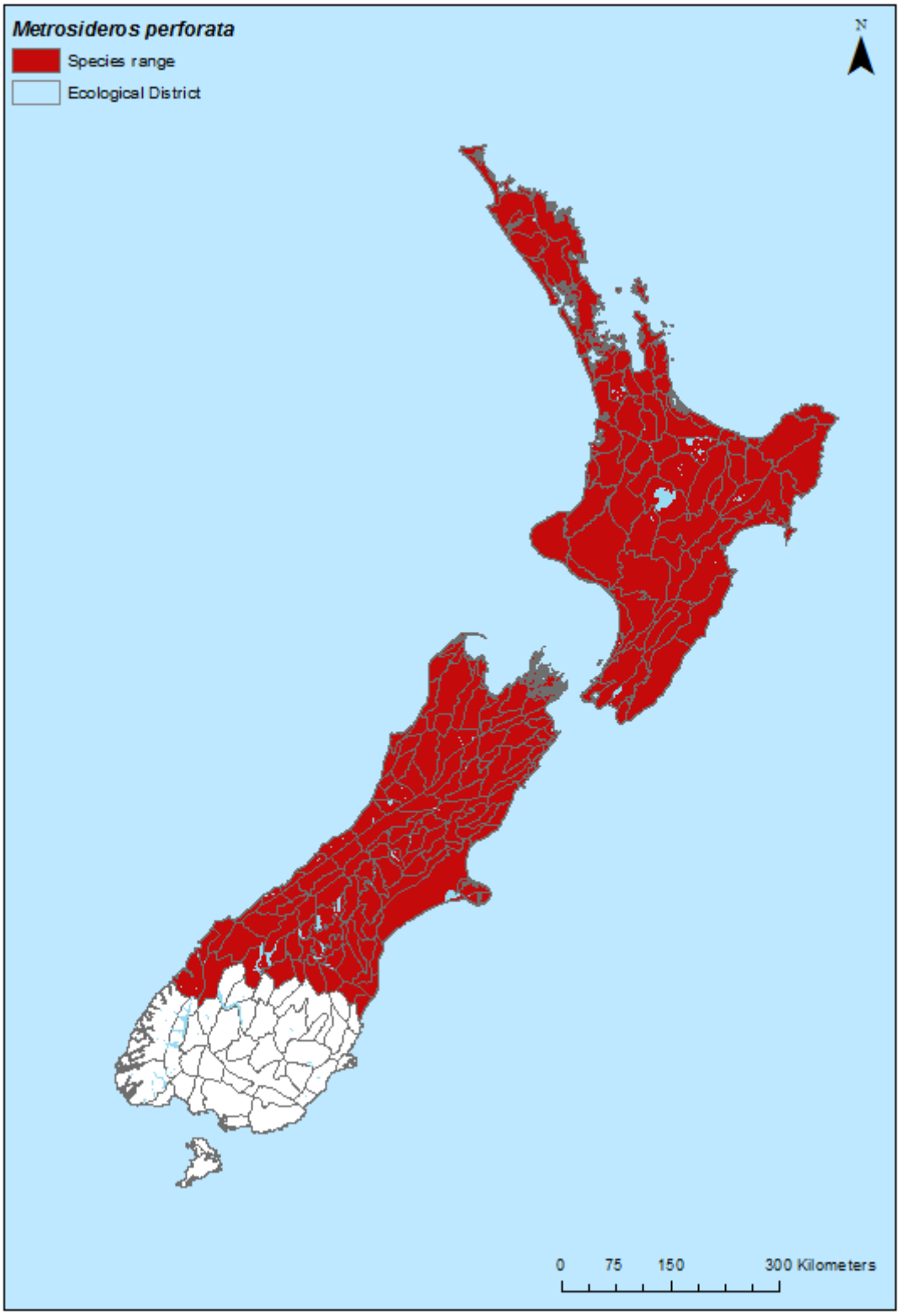

Figure 50 ***Metrosideros perforata*** **– range**

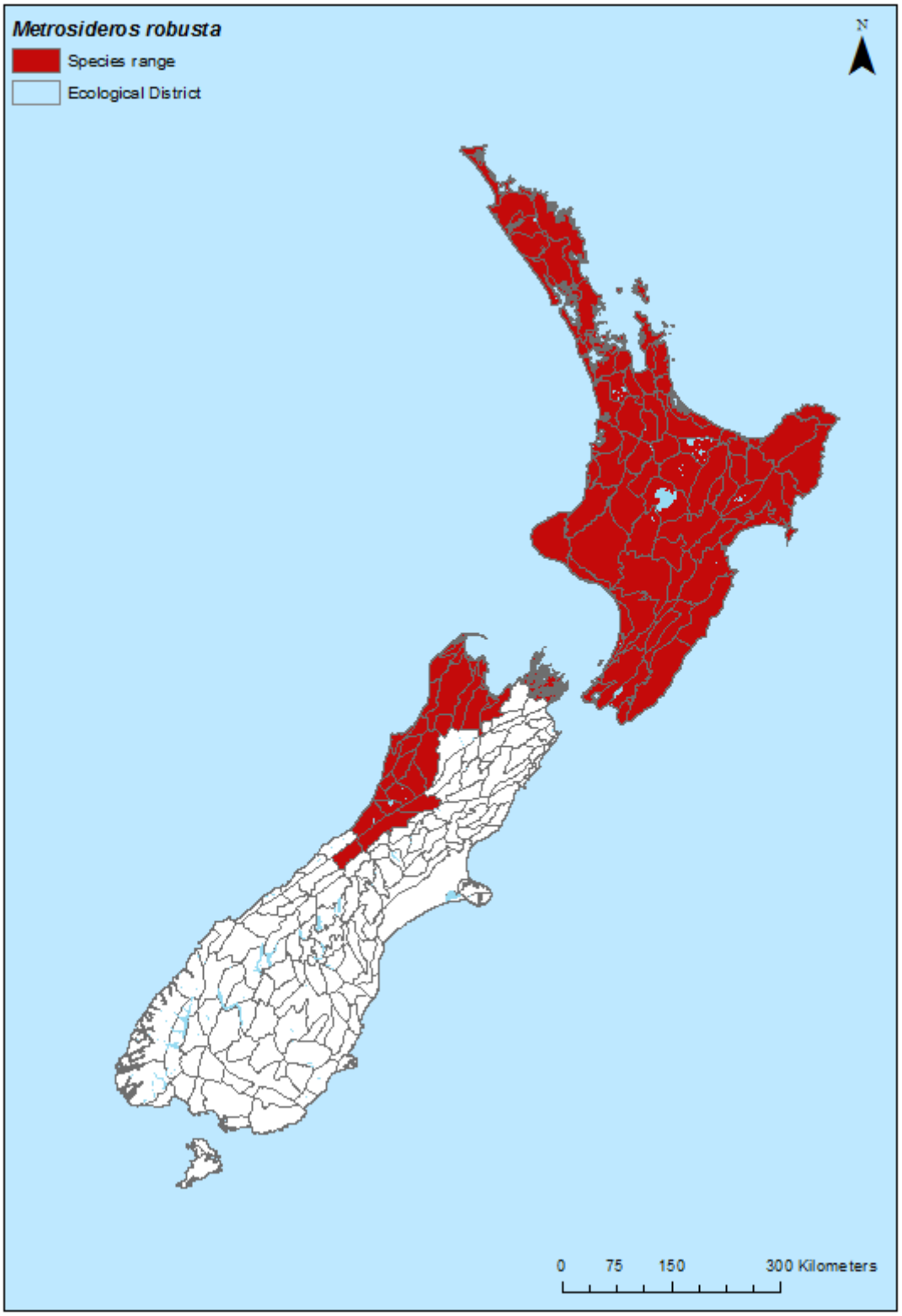

Figure 51 ***Metrosideros robusta*** **– range**

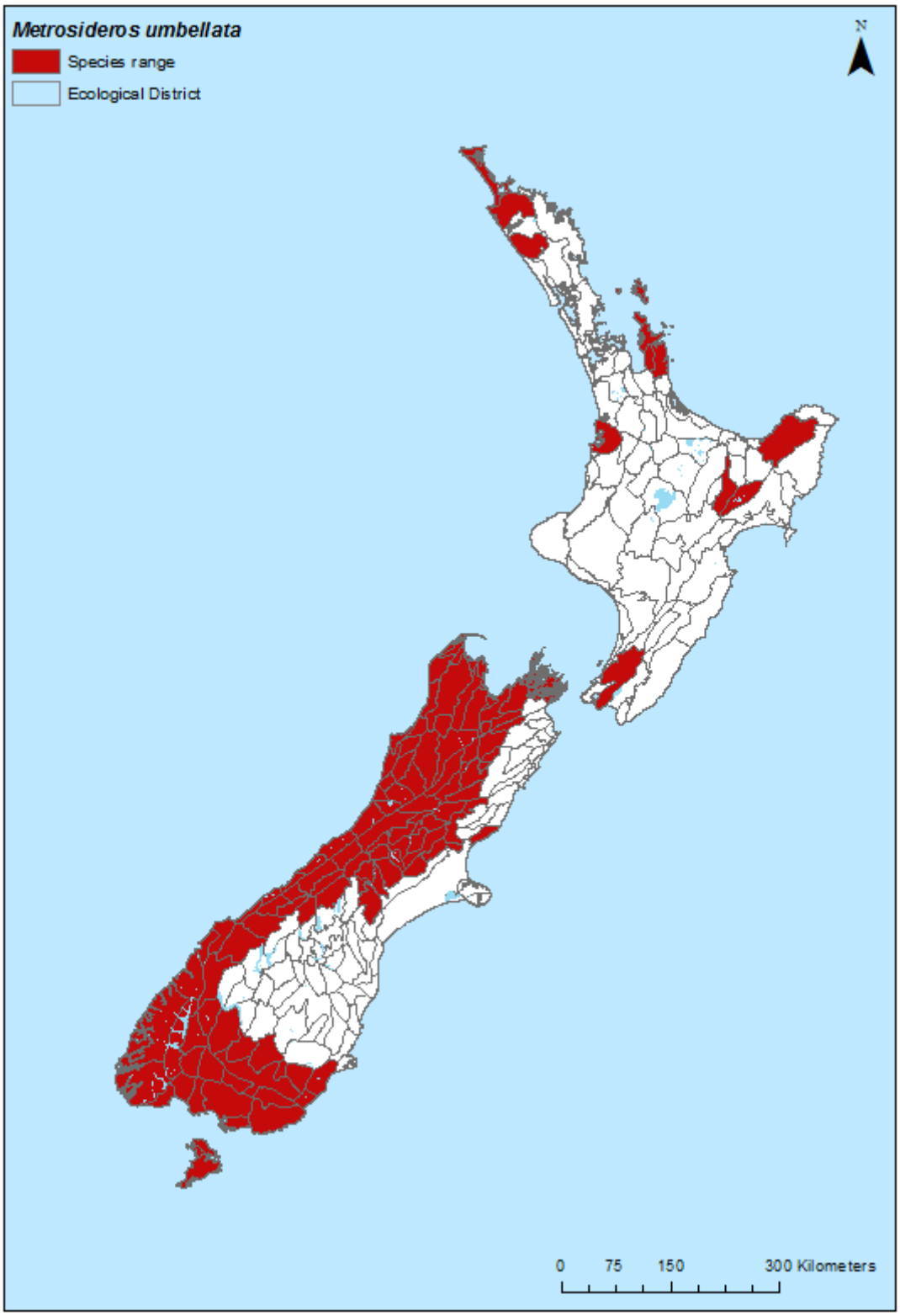

Figure 52 ***Metrosideros umbellata*** **– range**

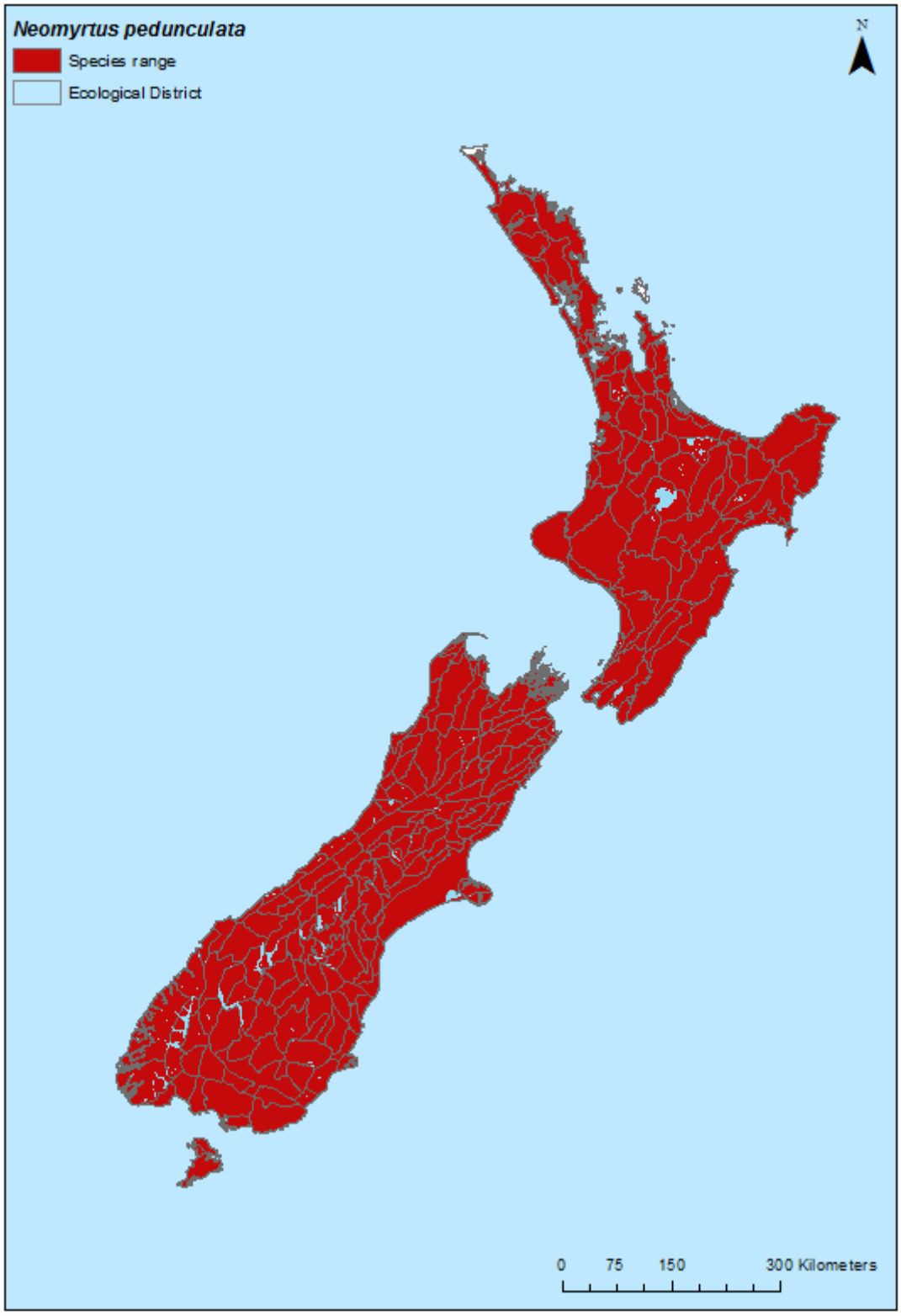

Figure 53 ***Neomyrtus pedunculata*** **– range**

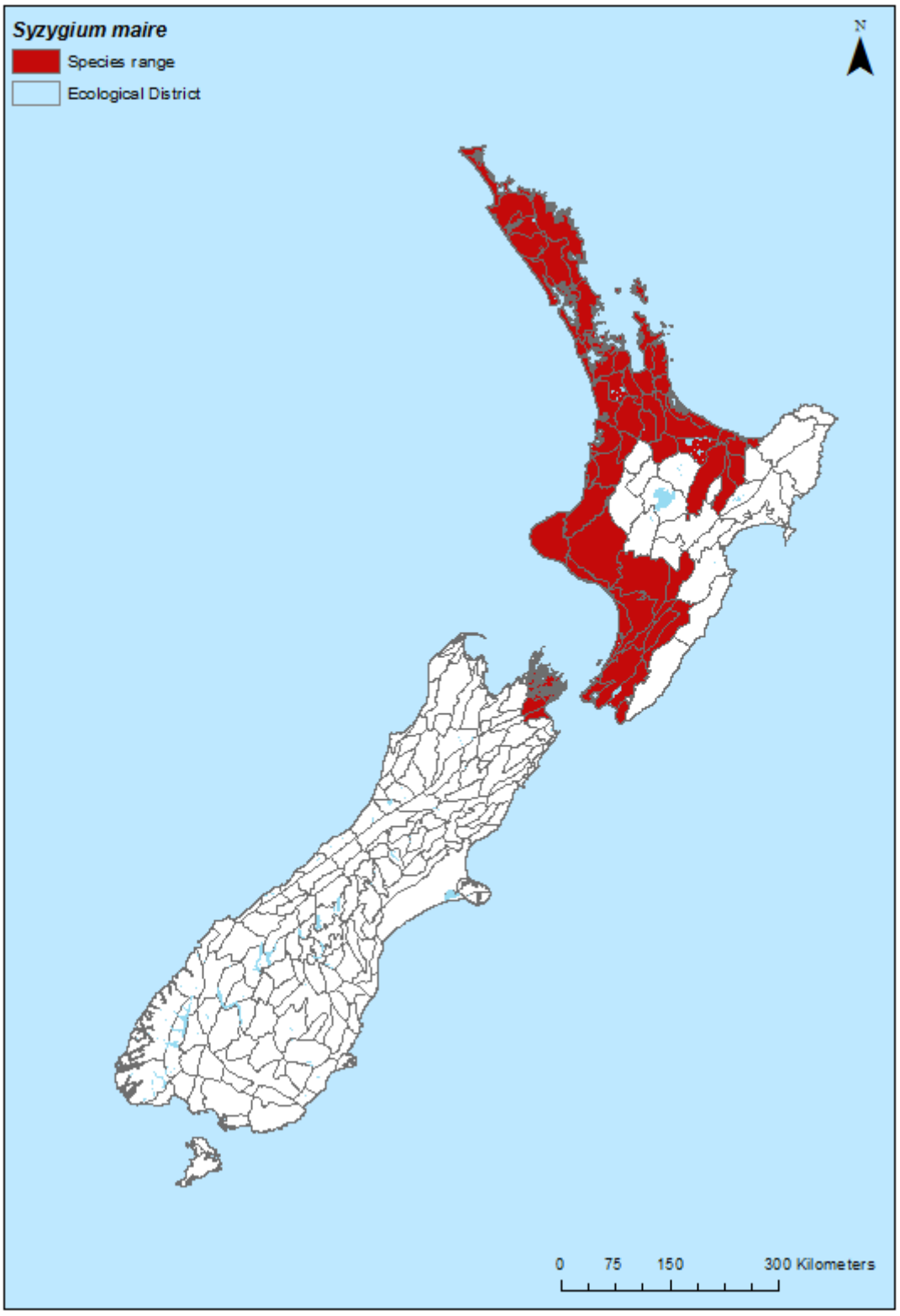

Figure 54 *Syzygium maire* – range

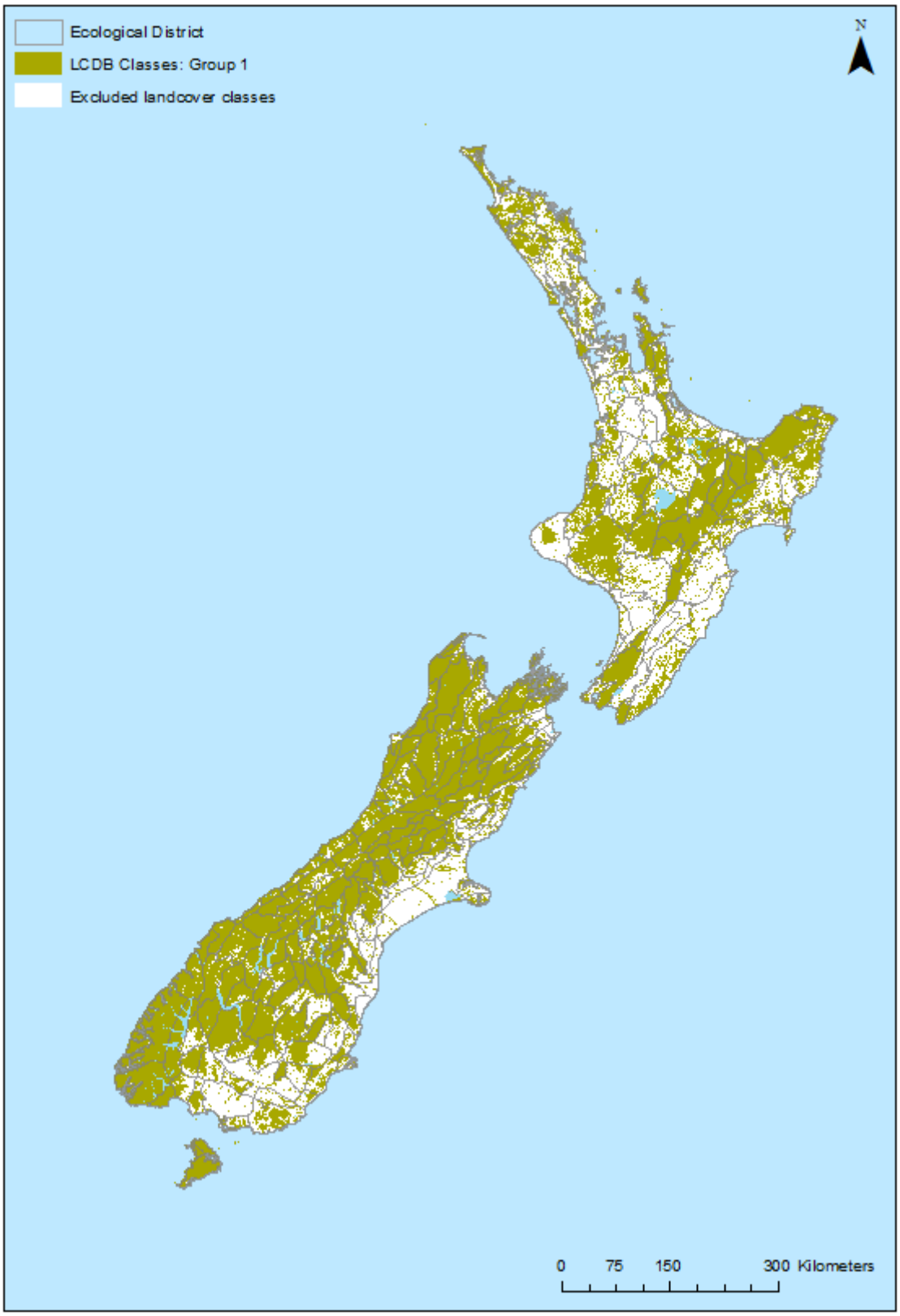

Figure 55 LCDB coverage – Group 1

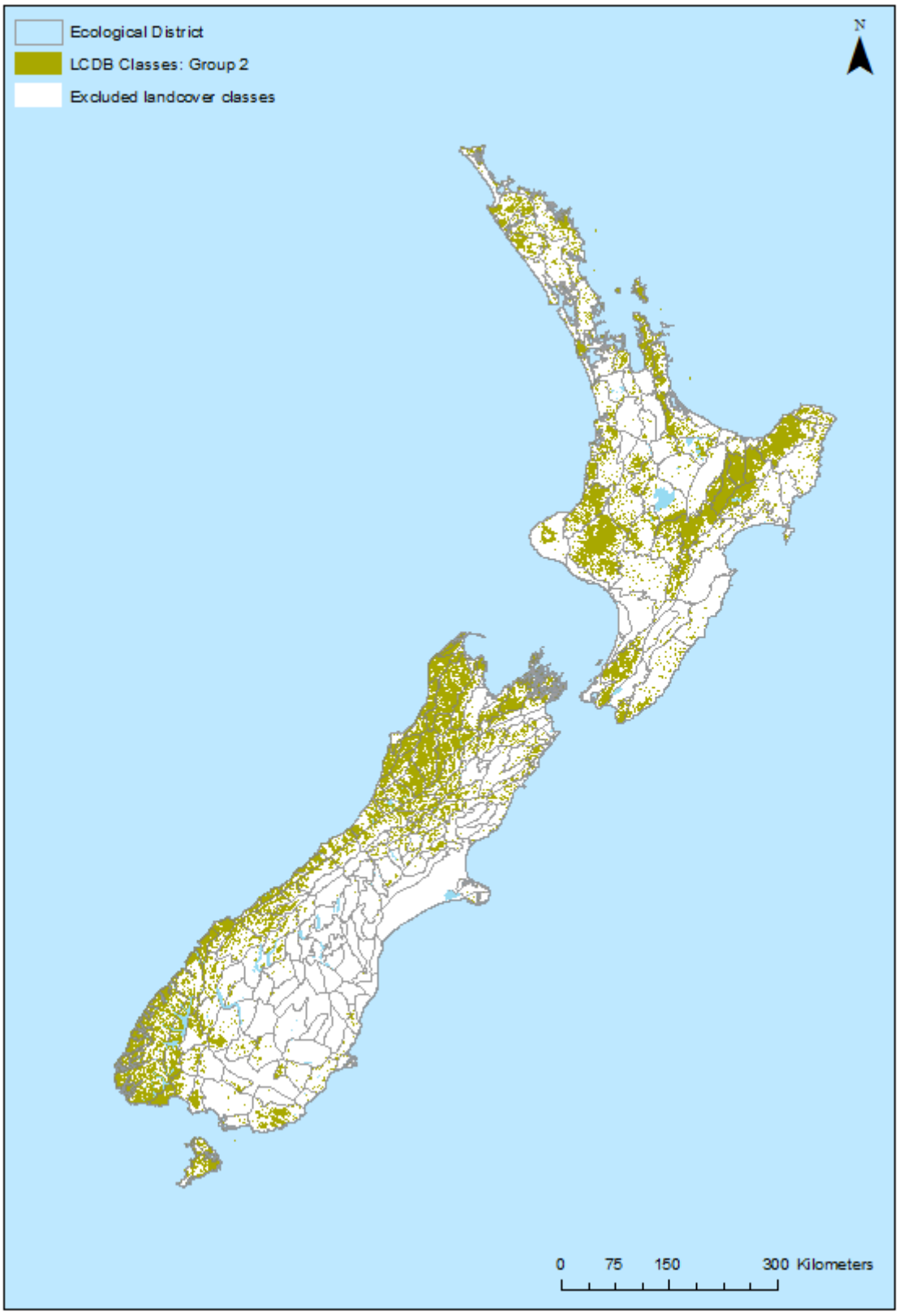

Figure 56 LCDB coverage – Group 2

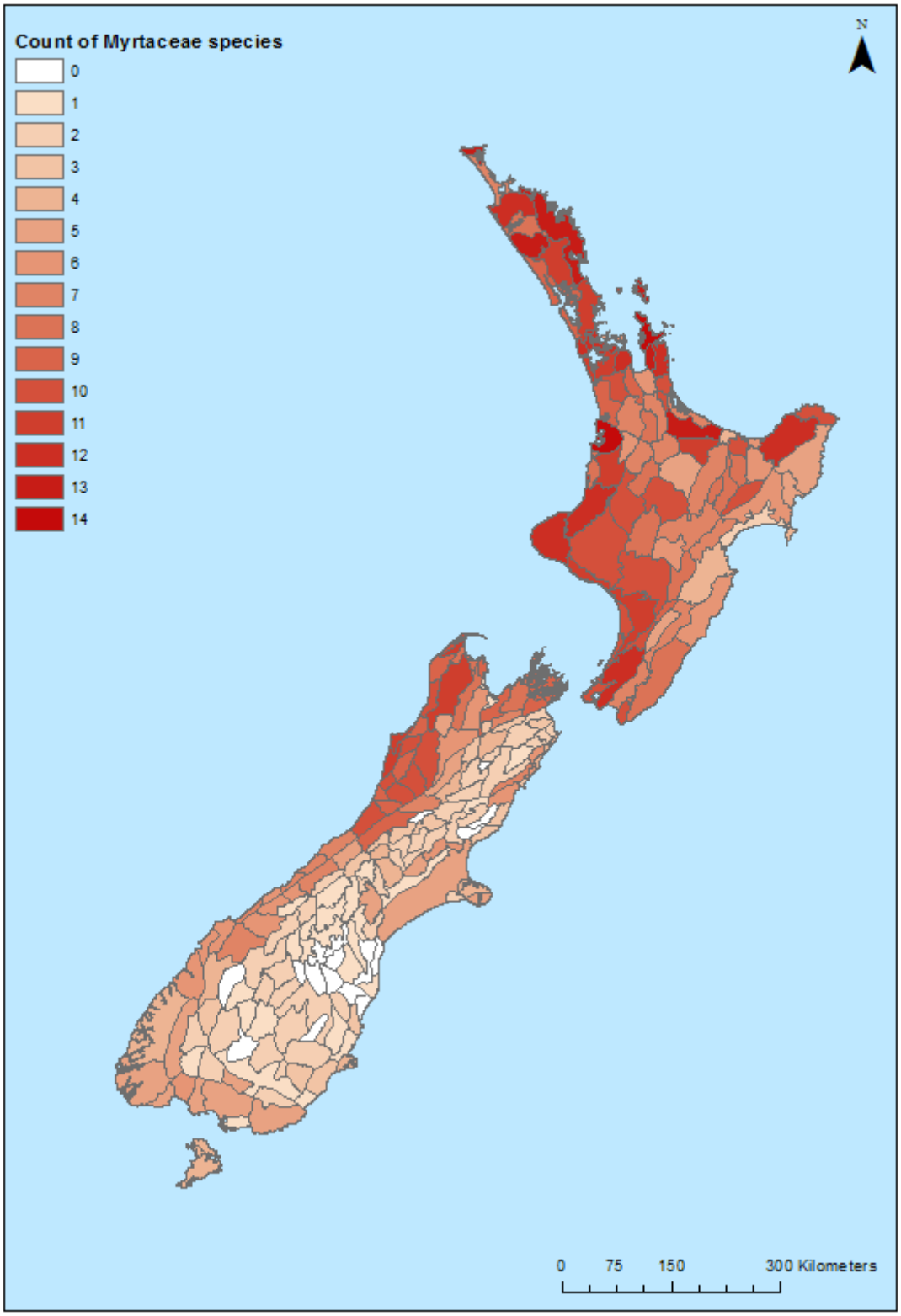

Figure 57 Myrtaceae species richness

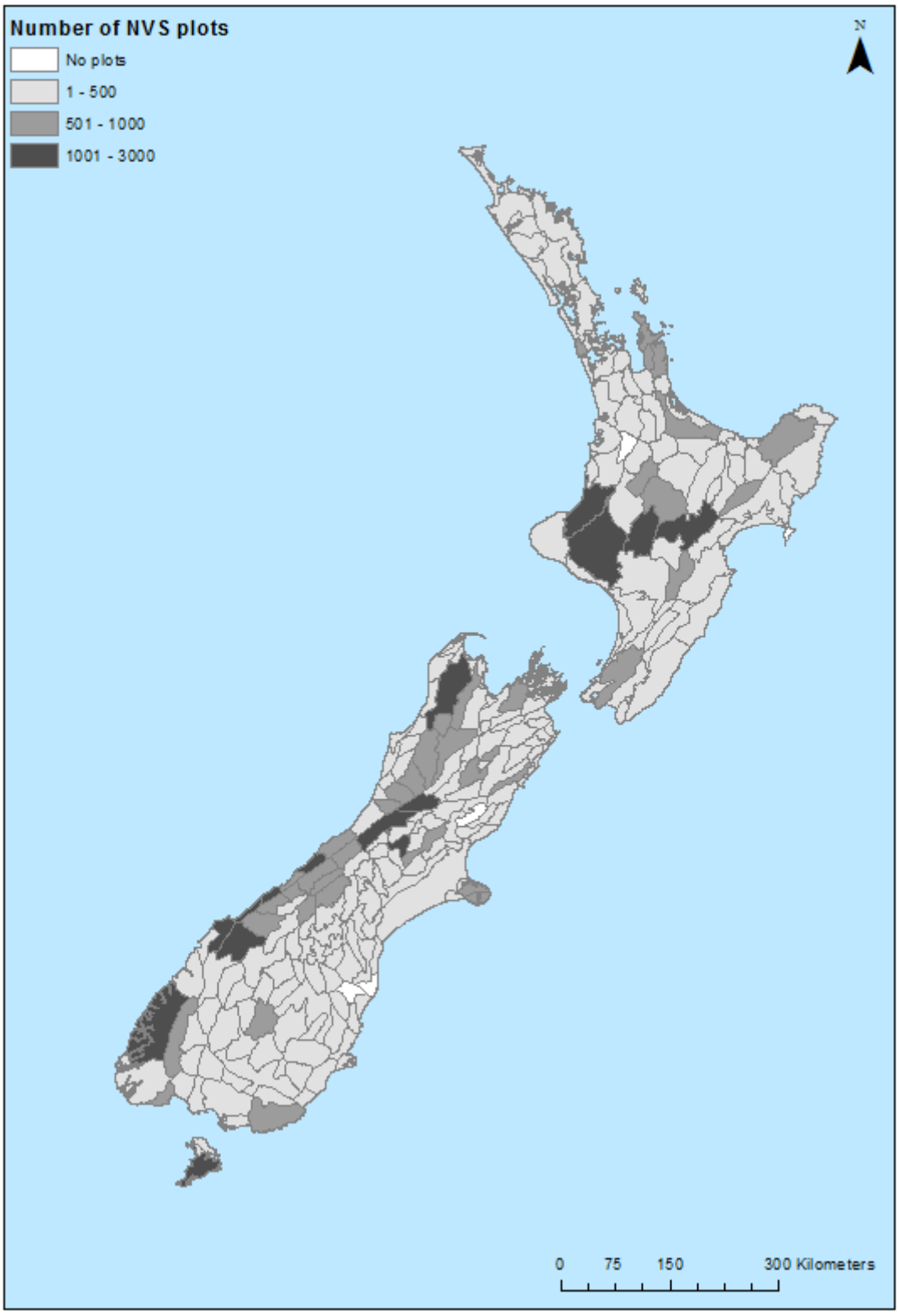

Figure 58 NVS Plot density

Limitations on interpreting these maps include:

- the difficulty identifying misidentifications and plants in cultivation (where they are not flagged as such in the data source)
- the estimation of species' natural range (we have used views from the published literature)
- the difficulty identifying duplicate observations of the same individuals, either as identical records from different sources, or observations at different times (the seemingly truncated geospatial resolution for some data sources, notably BioWeb and Scion/NZFRI, in particular, compound this difficulty)
- the use of the only source of presence/absence data (NVS) to provide a national measure of abundance when regional representation of that data is incomplete (e.g. there are no vegetation plots from Rangitoto Island digitised in the NVS databank).

Table 9 summarises the total range size for each taxon, based on our approach.

Table 9 Area (ha) for each Myrtaceae species using the LCDB mask

Species	*Area (ha)*
Kunzea ericoides sensu lato	11,507,782
Kunzea sinclairii	20,881
Leptospermum scoparium	15,245,882
Lophomyrtus bullata	2,147,121
Lophomyrtus obcordata	4,121,803
Metrosideros albiflora	498,461
Metrosideros bartlettii	19,526.
Metrosideros carminea	1,372,108
Metrosideros colensoi	2,872,941.
Metrosideros diffusa	6,494,587
Metrosideros excelsa	1,361,475
Metrosideros fulgens	4,818,728
Metrosideros parkinsonii	956,845
Metrosideros perforata	5,024,383
Metrosideros robusta	3,855,060
Metrosideros umbellata	4,133,547
Neomyrtus pedunculata	6,433,447
Syzygium maire	1,699,737

Notes: Areas are calculated by summing the area in the associated LCDB classes in the ecological districts where the species occurred. Any records from ecological districts judged to be beyond the species range were not included in these calculations.

Table 10 indicates the area and number of ecological districts of occurrence for each species.

Table 10 Area (ha) and number of ecological districts of occurrence for each species

Species	*Size of ecological district (ha)*	*Number of ecological districts*
Kunzea ericoides sensu lato	19,041,294	171
Kunzea sinclairii	29,090	1
Leptospermum scoparium	23,614,197	210
Lophomyrtus bullata	6,432,336.	48
Lophomyrtus obcordata	12,141,885	102
Metrosideros albiflora	1,547,512	13
Metrosideros bartlettii	36,529	1
Metrosideros carminea	4,560,178	42
Metrosideros colensoi	8,243,877	57
Metrosideros diffusa	16,843,485	138
Metrosideros excelsa	4,200,955	38
Metrosideros fulgens	10,906,789	92
Metrosideros parkinsonii	1,304,680	17
Metrosideros perforata	13,020,585	114
Metrosideros robusta	10,884,355	90
Metrosideros umbellata	8,551,166	86
Neomyrtus pedunculata	16,239,364	126
Syzygium maire	5,661,971	41

6 Conclusions

- The most narrowly distributed Myrtaceae species on the New Zealand mainland are *Kunzea sinclairii* and *Metrosideros bartlettii*. The most broadly distributed is *Leptospermum scoparium*.
- Nine species were designated as common in at least one ecological district. These are *Kunzea ericoides*, *Leptospermum scoparium*, *Lophomyrtus obcordata*, *Metrosideros diffusa*, *Metrosideros fulgens*, *Metrosideros perforata*, *Metrosideros robusta*, *Metrosideros umbellata* and *Neomyrtus pedunculata.*
- The ecological districts with the most Myrtaceae species are Colville and Kawhia (14 species each) and eastern Northland and islands (Great Barrier, Ōtānewainuku, Te Paki, Thames and Tūtāmoe (13 species each).
- Seventeen ecological districts appear to support no Myrtaceae species. These are Benmore, Duntroon, Grampians, Hawkdun, Kirkliston, Mākikihi, Minchin, Nokomai, Ōamaru, Richardson, Rock and Pillar, Sedgemere, St Bathans, St Mary, Waiau, Waikari and Waimate.

7 Recommendations

MPI has funded MWLR to map Myrtaceae distributions using species distribution models. This project will utilise presence-only and the presences from presence/absence (i.e. plots) data to produce models using the MaxEnt procedure, and using presence/absence data and the analytical approach of boosted regression trees. Predictor variables will be climate and other variables that can be derived from the layers underpinning Land Environments New Zealand. These models are scheduled for completion in September 2018, with a final report to be completed by May 2019. These maps will build on what we have delivered here and will result in a major improvement.

However, the underlying database could be markedly improved by:

- targeted digitisation of herbarium records from areas of known distribution but without extant data (most New Zealand herbaria are only partially digitised)
- targeted digitisation of vegetation plot records from areas of known distribution but no readily available vegetation plot data
- improving data quality from existing sources by, for example, flagging cultivated records, verifying potentially anomalous identifications, and quality assessment of georeferencing
- improving data quality in the NVS databank by further examination of spatial outliers
- gathering distribution data on the recently described new taxa of *Kunzea* and mapping these new taxa rather than the former broad concept of *Kunzea ericoides*
- incorporating records of occurrences in the pre-human record to improve current estimates of natural species ranges – these are currently qualitative and subjective.

8 Acknowledgements

We thank Chris Morse and Hamish Maule for checking the outlier NVS plot records, Jon Sullivan for providing access to the Naturewatch data, and Aaron Wilton for extracting NZVH data. We thank Ray Prebble for editing.

9 References

Allan HH 1982. Flora of New Zealand. 1: Indigenous tracheophyta – psilopsida, lycopsida, filicopsida, gymnospermae, dicotyledons. Wellington, P.D. Hasselberg.

Breitwieser I, Brownsey PJ, Heenan PB, Nelson WA, Wilton AD eds. 2010. Flora of New Zealand. http://www.nzflora.info. Accessed 25 October 2017.

Chamberlain S, Ram K, Hart T 2017. Package 'spocc': interface to species occurrence data sources. R package version 0.70.

de Lange PJ 2014. A revision of the New Zealand Kunzea ericoides (Myrtaceae) complex. PhytoKeys 40: 1.

Ministry of Agriculture and Forestry 2011. Risk analysis of the Puccinia psidii / Guava Rust fungal complex (including Uredo rangelii/Myrtle Rust) on nursery stock. Wellington, Ministry of Agriculture and Forestry.

Simpson PG 2005. Pōhutukawa & rātā: New Zealand's iron-hearted trees. Wellington, Te Papa Press.

Simpson PG 2011. Lake Sumner rata. Unpublished report for Hurunui Waiau Zone Committee.

Wardle P 1991. Vegetation of New Zealand. Cambridge, Cambridge University Press.

Wellington Botanical Society 2008. Native vascular plants of Great Barrier Island. Unpublished species list. https://www.thebarrier.co.nz/PDFs/ListNativeVascularPlantsGBI.pdf