DEPARTMENT OF PLANNING, INDUSTRY & ENVIRONMENT

Using representative sets of known species and habitat condition to inform change in biodiversity status

A case example for vascular plants

environment.nsw.gov.au

© 2020 State of NSW and Department of Planning, Industry and Environment

With the exception of photographs, the State of NSW and Department of Planning, Industry and Environment are pleased to allow this material to be reproduced in whole or in part for educational and non-commercial use, provided the meaning is unchanged and its source, publisher and authorship are acknowledged. Specific permission is required for the reproduction of photographs.

The Department of Planning, Industry and Environment (DPIE) has compiled this report in good faith, exercising all due care and attention. No representation is made about the accuracy, completeness or suitability of the information in this publication for any particular purpose. DPIE shall not be liable for any damage which may occur to any person or organisation taking action or not on the basis of this publication. Readers should seek appropriate advice when applying the information to their specific needs.

All content in this publication is owned by DPIE and is protected by Crown Copyright, unless credited otherwise. It is licensed under the Creative Commons Attribution 4.0 International (CC BY 4.0), subject to the exemptions contained in the licence. The legal code for the licence is available at Creative Commons.

DPIE asserts the right to be attributed as author of the original material in the following manner:
© State of New South Wales and Department of Planning, Industry and Environment 2020.

Authorship and citation

This document was prepared by David A Nipperess[1,2], Daniel P Faith[3], Kristen J Williams[4], Darran King[4], Glenn Manion[5,2], Chris Ware[4], Becky Schmidt[4], Jamie Love[2], Michael Drielsma[2], Stuart Allen[1] and Rachael V Gallagher[1]

[1] Department of Biological Sciences, Macquarie University, Sydney
[2] NSW Department of Planning, Industry and Environment, Sydney
[3] Australian Museum, Sydney
[4] CSIRO Land and Water, Canberra
[5] Cryzan, Armidale

This document should be cited as: Nipperess DA, Faith DP, Williams KJ, King D, Manion G, Ware C, Schmidt B, Love J, Drielsma M, Allen S & Gallagher RV 2020, *Using representative sets of known species and habitat condition to inform change in biodiversity status: A case example for vascular plants*, Biodiversity Indicator Program Implementation Report, Department of Planning, Industry and Environment, NSW, Sydney, Australia.

Published by:

Environment, Energy and Science
Department of Planning, Industry and Environment
59 Goulburn Street, Sydney NSW 2000
PO Box A290, Sydney South NSW 1232
Phone: +61 2 9995 5000 (switchboard)
Phone: 1300 361 967 (Environment, Energy and Science enquiries)
TTY users: phone 133 677, then ask for 1300 361 967
Speak and listen users: phone 1300 555 727, then ask for 1300 361 967
Email: info@environment.nsw.gov.au
Website: www.environment.nsw.gov.au

Report pollution and environmental incidents
Environment Line: 131 555 (NSW only) or info@environment.nsw.gov.au
See also www.environment.nsw.gov.au

ISBN 978-1-922318-50-3
EES 2020/0097
February 2020

Find out more about your environment at:

www.environment.nsw.gov.au

Contents

List of tables

List of figures

Acknowledgements

Tim Cooney, Joanne Wilson, Laura Babian, Jo White, Tony Auld, James Brazill-Boast, Alana Burley, Ron Avery, David Keith (DPIE), Karel Mokany (CSIRO) and Emily Nicholson (Deakin University) provided feedback during development and/or implementation. Glenda Wardle, Ayesha Tulloch, Aaron Greenville (University of Sydney), Thomas Newsome (CSIRO), Dan Rosauer (Australian National University), Tony Auld and James Brazill-Boast (DPIE) provided peer review comments on an earlier version of this document. Clare Gambley (CSIRO) supported the authors in the early translation of these key findings into the report card.

Thanks to Donella Andersen, Nature Edit, for her work editing and formatting this report.

List of abbreviations

ACT	Australian Capital Territory
AOO	Area of occupancy
APC	Australian Plant Census
CSIRO	Commonwealth Scientific and Industrial Research Organisation
EOO	Extent of occurrence
GDM	Generalised dissimilarity model
IUCN	International Union foor the Conservation of Nature
NARCliM	NSW and ACT Regional Climate Modelling project
NSW	New South Wales
SDM	Species distribution models
SoS	*Saving our Species*

Context

The New South Wales (NSW) Government has introduced new legislation for biodiversity conservation and native vegetation management, including the *Biodiversity Conservation Act 2016* (the BC Act), which commenced 25 August 2017. The goals of the BC Act include the conservation of biodiversity at bioregional and state levels, a reduction in the rate of species loss, and effective management to maintain or enhance the integrity of natural habitats. To contribute to assessing the performance of the new legislation, the former Office of Environment and Heritage NSW established the Biodiversity Indicator Program to report on the status of biodiversity and ecological integrity at regular intervals. The responsibility of implementing this program now rests with the Department of Planning, Industry and Environment.

Monitoring of *all* known biodiversity across New South Wales is impractical. Monitoring of biodiversity across New South Wales is a large, complex task that requires novel approaches to data collection and use, including the application of models to help measure status and track change. Surrogates intended to best represent biodiversity are often relied on to address these challenges. This indicator implementation report details how three biodiversity indicators for measuring status and change over time are developed and applied using existing and new approaches to analysing species location records and associated surrogate information.

The overarching monitoring framework, which outlines how indicators are related and derived, is presented in *Measuring Biodiversity and Ecological Integrity in New South Wales: Method for the Biodiversity Indicator Program* (OEH & CSIRO 2019). The method for the Biodiversity Indicator Program establishes a nested design within which all **indicators**, as they are developed, have a place. Each indicator is nested with others of its type in an **indicator family**, and each family is nested within one of five **themes** which are associated with either the biodiversity or ecological integrity **class** of indicators (as shown in Figure 1).

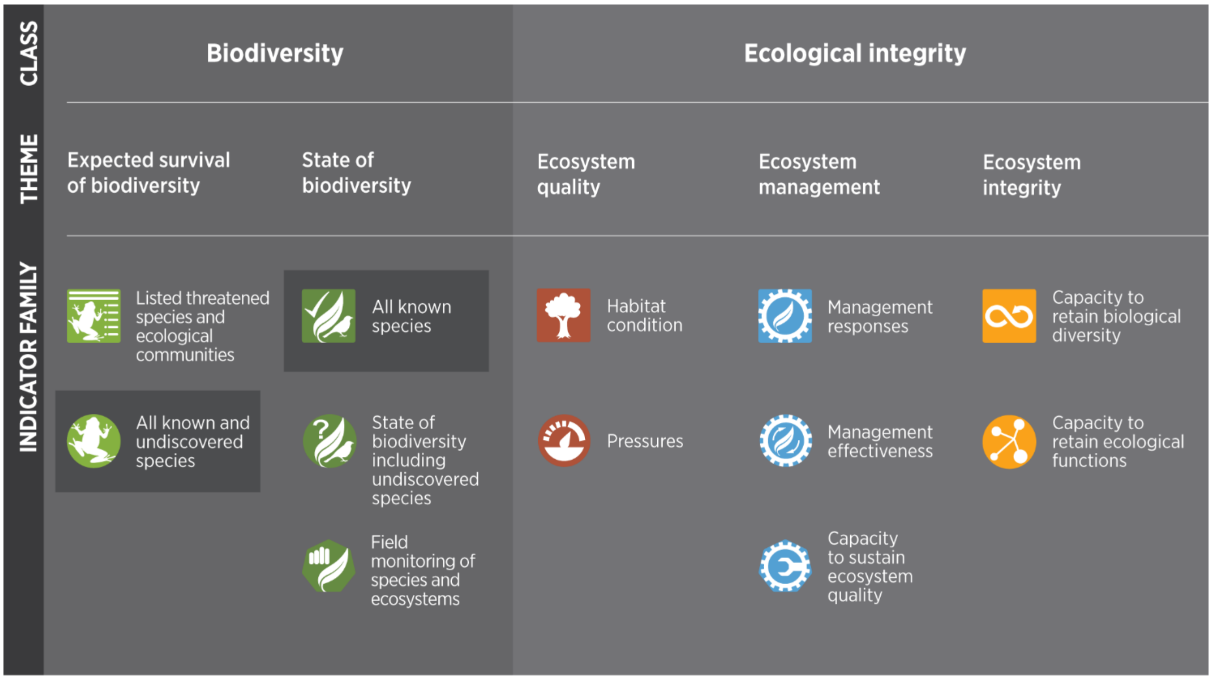


Figure 1 Nested structure used to arrange and link indicators for measuring biodiversity and ecological integrity in New South Wales. This implementation report covers indicators in two families (shown by the darker grey boxes).

The key results and highlights are presented in one of several report cards in the first *NSW Biodiversity Outlook Report* (DPIE 2020). The indicators detailed in this report sit within the nested framework as follows:

Class: **Biodiversity**

Theme: **1. Expected survival of biodiversity**

Indicator family: 1.2 Expected survival of all known and undiscovered species

Indicator: 1.2a Expected survival of all known species

Proportion of all known species expected to survive in 100 years, assessed for each biological group

Theme: **2. State of biodiversity**

Indicator family: 2.1 State of all known species

Indicator: 2.1a Within-species genetic diversity (for all known species)

The proportion of within-species genetic diversity of all species known to exist, assessed for each biological group

Indicator: 2.1b Extant area occupied (for all known species)

The average fraction of original habitat presently occupied by all known species, assessed for each biological group

The **expected survival of biodiversity** indicator family predicts extinction risk of all known and undiscovered species, beyond those formally assessed by the NSW Threatened Species Scientific Committee. It uses the IUCN Red List for Threatened Species assessment method for criteria B2 and B2(b)(ii) for estimated area of occupancy (AOO). Different measurement methods are applied to comprehensively assess biodiversity using this risk framework and to minimise sampling bias. It can be used as a measure of the overall effectiveness of ecosystem management in securing, maintaining or improving the survival of all known and undiscovered species in 100 years' time. The habitat condition family of indicators 3.1 support this assessment.

The **state of biodiversity** indicator family allows the assessment of the overall diversity of all known species that exist at present, including their genetic diversity, as a proportion of that which originally existed in New South Wales prior to the industrial era (c. AD1750). The habitat condition family of indicators 3.1 support this assessment.

Summary

The BC Act requires a program for monitoring biodiversity for the state of New South Wales. Consequently, the former Office of Environment and Heritage collaborated with Commonwealth Scientific and Industrial Research Organisation (CSIRO), Macquarie University and the Australian Museum to develop indicators for biodiversity at bioregional and statewide scales. The method implemented in this report establishes a first assessment for a suite of indicators that measure the rate of loss of biodiversity of vascular plants known to occur naturally in New South Wales.

Reduction in the geographic range size of a species due to habitat loss, alteration and fragmentation is well known to decrease within-species genetic diversity and increase extinction risk. Therefore, current range size and proportion of range lost from habitat loss, alteration and fragmentation were estimated for vascular plant species known to occur naturally in New South Wales. The area of effective habitat (i.e. high quality habitat able to support biodiversity) remaining for each species was estimated from two indicators (Love et al. 2020):

- ecological condition of terrestrial habitat
- ecological carrying capacity of terrestrial habitat.

From this basis, three indicators reported here were developed for vascular plants:

- expected survival of all known species
- within-species genetic diversity (for all known species)
- extant area occupied (for all known species).

Because most species in New South Wales have not been formally assessed for possible threatened status (i.e. at heightened risk of extinction), a provisional risk assessment using a limited set of criteria was completed for all NSW vascular plant species for which adequate data were available. This assessment formed the basis for the **expected survival of all known species** indicator. The risk assessment classified species into four categories: lowest risk, lower risk, higher risk and highest risk. By assigning each category a probability of survival, the proportion of NSW vascular plant species expected to survive in 100 years was estimated.

Extrapolating trends in the rate of biodiversity loss requires that the list of species used in analyses are representative of the overall biodiversity of New South Wales. A subset of NSW vascular plant species that uniformly represent the full variety of natural habitats for vascular plants in New South Wales (called the **representative species set**) was selected to be representative of all vascular plant species, including those yet to be discovered. This species set may also be indicative of the risk status of all vegetation-dependent biodiversity, including some invertebrate taxonomic groups that are not well known in New South Wales.

This first assessment reports on values for each indicator up to 2013 and is based on the assessment of ecological condition and ecological carrying capacity (Love et al. 2020). Condition and carrying capacity are used to infer the amount of habitat loss, alteration and fragmentation associated with each species' **original range** (based on occurrence observed since 1950). Results are summarised in the key findings section below.

Uncertainties exist in the following factors underlying the indicators:

- the true original range size of each species
- the variation among species in the relationship between current range size and habitat loss, alteration and fragmentation
- the variation among species in the relationship between current range size and within-species genetic diversity
- the variation among species in the relationship between current range size, decline in range size and extinction risk.

These indicators can be updated as new data become available on the condition of terrestrial vegetation in New South Wales and as new occurrence records improve our understanding of the original geographic distribution of NSW vascular plant species.

Key findings

The habitat-based assessment of the three biodiversity indicators reported here has been implemented for known species of vascular plants in New South Wales.

2.1b Extant area occupied (for all known species)

- Most vascular plant species known from New South Wales have experienced significant habitat loss and alteration across their original range.

- On average, about 56% of NSW plant species' original habitat remains effective, based on records of occurrence observed since 1950 and ecological condition of biodiversity habitats assessed up to 2013.
- For the representative species sets, an average of 57% of habitat remains effective.
- When considering ecological carrying capacity, which additionally accounts for the effects of fragmentation on ecological processes influenced by connectivity, an average of about 49% of a species' original habitat remains effective, both overall and for representative species sets.

2.1a Within-species genetic diversity (for all known species)

- Based on the proportion of remaining effective habitat in each species' original range, an average of 79–91% of original within-species genetic diversity is retained, both for all NSW vascular plant species (for which data are available), and for the representative species set. A range of values is given because each species will respond to loss of range size differently, depending on factors like dispersal ability and degree of adaptation to local environmental conditions, and these differences are not precisely known.
- When considering ecological carrying capacity, an average of about 74–89% of original within-species genetic diversity remains, both overall and for representative species sets.

1.2a Expected survival of all known species

- A provisional risk assessment suggests that 12% of all NSW vascular plant species are at the lowest level of extinction risk. This result implies that as much as 88% of the more than 6000 vascular plant species known to occur naturally in New South Wales are potentially at a heightened risk of extinction. Given most of these have not yet been formally assessed, this represents a significant knowledge shortfall regarding the true risk status (and, as such, the expected survival) of biodiversity in New South Wales.
- For the representative species sets, an average of 19% of vascular plant species are at the lowest level of extinction risk.
- Based on the provisional risk assessment, about 73% of NSW vascular plant species, and of representative species sets, are expected to survive in the next 100 years. This precautionary *lower* limit is an indicator, not a prediction. Such uncertainty does not impact on the information value of the indicator as long as range size and habitat loss, alteration and fragmentation are reliable correlates of extinction risk.
- Extrapolation of the rate of historical discovery of NSW vascular plants indicates that most species have been described and the inventory is very nearly (i.e. ~ 99%) complete.

Future assessments of the indicators

- These indicators can be updated as new data become available on the condition of natural vegetation in New South Wales and as new occurrence records improve the understanding of the geographic distribution of NSW vascular plant species, or other taxonomic groups are equivalently assessed.
- Potential refinements of the methods reported here include using species distribution modelling to estimate range area, revising data cleaning, estimating new parameters for modelling genetic diversity, and revising the risk assessment process to include additional criteria.

1. Introduction

The former Office of Environment and Heritage NSW collaborated with the Commonwealth Scientific and Industrial Research Organisation (CSIRO), Macquarie University and the Australian Museum to develop a method for the collection, monitoring and assessment of biodiversity information in New South Wales at regional and statewide scales (OEH & CSIRO 2019). The technical implementation of the method specifically detailed in this report establishes a first assessment, prior to commencement of the *Biodiversity Conservation Act 2016* (the BC Act), for a class of indicators that measures expected survival of biodiversity in 100 years, including those species not yet known to be at risk of extinction, in New South Wales.

Consistent with the BC Act, 'biodiversity' is here defined as 'the variety of living animal and plant life from all sources, and includes diversity within and between species and diversity of ecosystems'. This definition deals exclusively with composition (variety of life) and does not include structural and functional attributes that are often conflated with composition in a broader definition of biodiversity (c.f. Noss 1990). It is important to note that the biodiversity of New South Wales includes all those populations, species and ecosystems yet to be discovered, catalogued and assessed for their conservation status. This definition of biodiversity does encompass within-species diversity, that is, the variety of morphological, physiological and behavioural traits exhibited by individuals and populations of the same species, and the genetic variety that is a basis of this trait diversity.

A core purpose of the Act is to 'support conservation and threat abatement action to slow the rate of biodiversity loss and conserve threatened species and ecological communities in nature' (s. 1.3(h)). This implies an indicator is needed to measure the rate of biodiversity loss (Pereira & Cooper 2006). Successful conservation programs should slow or even reverse the rate of loss at the state level. The goal of slowing the rate of biodiversity loss is a globally recognised requirement for conservation programs and is a central goal for the Convention on Biological Diversity (Secretariat of the CBD 2011).

Expected diversity (Weitzman 1992) provides a means of estimating the rate of loss by determining the amount of biodiversity expected to still exist at some point in the future (e.g. 100 years from now), derived from the estimated survival probabilities of species. Rate of loss is simply the difference in diversity between now and that expected in the future. The key indicator of conservation success is then a reduction in the rate of loss. The expected diversity framework resembles the Red List Index (Butchart et al. 2004; Butchart et al. 2007), in that species are weighted according to their relative risk of loss, but differs in being a more directly probabilistic approach that is extensible to a variety of measures of biological diversity.

An expected diversity indicator is constrained by available information on biodiversity, requiring survival probabilities estimated from risk assessments of each species to be included in the calculation. Ideally, population viability analyses (Boyce 1992) for individual species would form the basis for determining survival probabilities. As these data are not generally available, a useful proxy to use are formal lists of species assigned to threat categories (Mooers et al. 2008), which typically utilise a standardised scientific assessment of extinction risk (Mace et al. 2008). Only those species described, catalogued and assessed to determine their threatened status could then be included in the calculation.

Information on the extinction risk of NSW species is provided by the formal listing of threatened species on schedules of the BC Act, but no comprehensive assessment of all members of a major taxonomic group has been attempted. Under the Act, nominated species are listed as either vulnerable, endangered, critically endangered or extinct, or are not listed due to not being at risk of extinction or because available data are too deficient for a determination. The listing of a species requires a rigorous assessment of available evidence, with assignment to a threatened category requiring qualification for specific criteria which follow those of the International Union for Conservation of Nature (IUCN), the *Guidelines for Using the IUCN Red List* (IUCN 2017). However, the NSW threatened species list is a small subset of all known NSW species and may not be representative of the true pattern of extinction risk. For example, less than 2% of NSW threatened species are invertebrates despite comprising about 67% of described Australian species (Chapman 2009; Nipperess 2015).

In the absence of a formal threatened status for many species, a provisional risk assessment process that uses similar criteria to a formal scientific determination could be applied to all species (or a large proportion) in a major taxonomic group. Commonly used criteria for formal conservation assessments, such as the *IUCN Red List of Threatened Species* (IUCN 2018) are a species' range area and trends in range area (Mace et al. 2008) because data on species' distribution are more likely to be available than abundance or genetic data. Further, trends in range area can be used as a surrogate for corresponding trends in abundance (Gaston et al. 2000) and within-species genetic diversity (Alsos et al. 2012; Mimura et al. 2017). Thus, monitoring range area and trends in range area for all (or most) species in a major taxonomic group has the advantage of being both tractable and likely to be symptomatic of the pattern of true extinction risk in that group, although not all species at risk will necessarily be detected by this means.

An indicator derived from a sample of species from a taxonomic group, often generated from the haphazard accumulation of available information, is not necessarily representative of the entire group, or of overall biodiversity. There is a need, therefore, to employ clear criteria to construct a set of species that is broadly representative of its taxonomic group and could be used as a surrogate for patterns of extinction risk in lesser-known groups. Given the general principle that biodiversity responds to such key environmental habitat gradients as temperature and rainfall (Faith et al. 2004), a set of species that is representative of environmental diversity is expected to be a good indicator set for overall biodiversity. The environmental diversity method (Faith & Walker 1996) has been shown to be an effective way to link degree of representation of environmental space to degree of representation of species diversity (e.g. Beier & de Albuquerque 2015). A set of species that represents the space as indicated by the environmental diversity approach will, therefore, also represent other species.

This report provides technical detail on the methods and results for a suite of indicators related by method and informing on the status of biodiversity in New South Wales. The indicators report on the average proportion of range area remaining for vascular plant species as a group in New South Wales, the genetic diversity remaining for that group of species, and the proportion of species within that group expected to survive in New South Wales in 100 years. The indicators are designed to be representative of all known vascular plant species in New South Wales and of those plant species inferred to be present but not yet discovered. **This first assessment reports on the status for each indicator up to 2013** (prior to commencement of the BC Act on 25 August 2017). The 2013 status values are based on the assessment period for the ecological condition and ecological carrying capacity indicators (Love et al. 2020), which are used to infer the amount of effective habitat remaining within each species' original range.

Technical details on the following indicators are included in this report, and summarised below (see Figure 1 for relationship to other biodiversity and ecological indicators):

- **Expected survival of all known species**

  Proportion of all known and undiscovered species expected to survive in 100 years, assessed for each biological group

  This indicator estimates extinction risk of all biodiversity (both known and undiscovered species) beyond those formally assessed by the NSW Threatened Species Scientific Committee. Species from a biological (i.e. taxonomic) group are sampled to uniformly represent the full range of natural habitats for that group. The representative species are provisionally assigned to risk of extinction categories based on the estimated proportion of their original habitat that remains intact. This is a limited, provisional assessment of risk using commonly available species occupancy data. The method uses species occurrence observations since 1950 in 2-kilometre map grids (each being 4 km^2) and area of occupancy (AOO) thresholds specified by the two criteria to discriminate four risk of extinction categories. Each species is further assessed for a reduction in AOO determined from the ecological condition indicator as a measure of habitat condition. The reduction in AOO in four classes (<30%, 30–50%, 50–80% and >80%), and the AOO thresholds (also using ecological condition) provide the dimensions of the risk categorisation. Each category is given a probability of survival which is applied to all representative species in that category. The indicator is calculated by summing the probabilities of survival for the representative species across all categories and is expressed as a proportion of the total number of species representing the biological group. This serves as an indicator for all known species within the biological group expected to survive in 100 years and, by logic, extends to undiscovered species in that group. Change in the value of the indicator reflects a change in survival probability due to a change in habitat condition. If sufficient habitat is lost or degraded for a particular species, its extinction risk category will also change.

- **Within-species genetic diversity (for all known species)**

  The proportion of within-species genetic diversity of all species known to exist, assessed for each biological group

  This indicator uses a representative sample of species to assess the proportion of within-species genetic diversity that still exists, after considering loss of suitable habitats. Genetic diversity is inferred from species diversity using geographic range and occupancy. Species from a biological (i.e. taxonomic) group are sampled to uniformly represent the full range of natural habitats for that group. The occurrence data in 2-kilometre map grids define each species' AOO. Reductions in AOO due to declines in habitat condition are determined using the ecological condition measure and, for comparison, the ecological carrying capacity measure. A power curve then relates the intact fraction of a species' AOO to the respective fraction of genetic diversity remaining. Two forms of the curve are used: one that simulates spatially high genetic diversity due to high rates of population divergence, and the other low. The two curves equate to an upper and lower estimate of fractional within-species genetic diversity. The indicator is calculated by separately summing the upper and lower fractions of genetic diversity remaining for all species representing the biological group. This serves as an indicator of within-species genetic diversity for all known species within the biological group and, by logic, extends to undiscovered species in that group. It is also used to show the variation in genetic diversity loss across the categories of species survival for indicator 1.2a (i.e. expected survival of all known species). Change in the value of the indicator reflects a change in habitat condition.

- **Extant area occupied (for all known species)**

  The average fraction of original habitat presently occupied by all known species, assessed for each biological group

  This indicator assesses the extant area occupied by all known species as a proportion after considering loss of original suitable habitat. Species from a biological (i.e. taxonomic) group are sampled to uniformly represent the full range of natural habitats for that group. The occurrence data in 2-kilometre map grids define each species' AOO, and fractional reductions in AOO due to declines in habitat condition are determined from the ecological condition indicator and, for comparison, the ecological carrying capacity indicator. The indicator is the average fraction of original habitat occupied by the representative species. This serves as an indicator for all known species within the biological group and, by logic, extends to undiscovered species. It is also used to show variation in reductions in AOOs across the categories of species survival for indicator 1.2a (i.e. expected survival of all known species). Change in the value of the indicator reflects a change in habitat condition.

2. Method design

2.1 Range area

The geographic distribution (or range) of a species is one of its most fundamental traits and is a useful indicator of extinction risk (Gaston & Fuller 2009). The area of a species' range and temporal trends in that area are often used in assessing risk of extinction as it is a useful surrogate for abundance information (which is unavailable for most species) (Gaston et al. 2000; Mace et al. 2008). Apart from this relationship with abundance, the particular case of a small range size is itself an extinction risk, as it increases the exposure of the entire species' range to threatening processes such as habitat modification (Cardillo et al. 2008).

Ultimately, a species' range must be inferred from occurrence records, that is, the set of individual observations which locate a species in space and time (Kissling et al. 2018). The *IUCN Red List Categories and Criteria* (IUCN 2012, referred to here as the '*IUCN Red List* Categories') derives two forms of range area from occurrence records: extent of occurrence (EOO) and area of occupancy (AOO). EOO is the area within the boundary connecting the outermost occurrence records (i.e. the convex hull). AOO is the summed area of regular grid cells intersecting with the occurrence records (Figure 2).

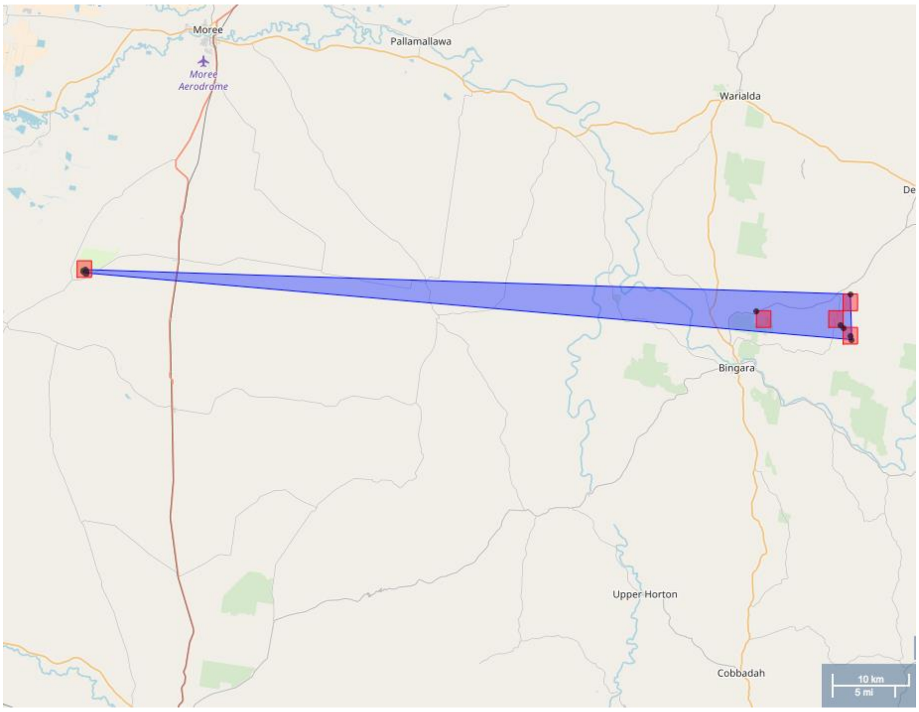


Figure 2 Range of the Myall Creek wattle (*Acacia atrox* Kodela) showing the intersection of occurrence records (black points) with 4-km^2 grid cells (in red) and the convex hull (in blue) connecting the outermost points

Area of occupancy (AOO) is the sum of the area of the grid cells intersecting with records. Extent of occurrence (EOO) is the area of the convex hull. Map generated from the spatial portal of the Atlas of Living Australia. Base map is provided by OpenStreetMap.

The EOO encompasses all known populations of a species but it is unlikely that the entire extent contains suitable and occupied habitat. AOO, therefore, is the area within the EOO that is known to be occupied. While this means that AOO might be preferred in risk assessments as representing confirmed occupied habitat, comparison of AOO and EOO provides additional information on the spatial separation of known populations, which is relevant to dispersal and gene flow and the spatial configuration of threatening processes. Clearly, AOO is sensitive to the size of the grid cell chosen and so, as recommended in the *Guidelines for using the IUCN Red List Categories and Criteria* (IUCN 2017, referred to here as the '*IUCN Red List Guidelines*'), this is usually standardised to 4 km^2 (i.e. 2 kilometre x 2 kilometre grids).

The temporal trend in range area provides a useful biodiversity indicator but is unlikely to be adequately tracked by the haphazard accumulation of occurrence records. By monitoring changes in range area for a suitably representative set of species, trends for overall biodiversity in New South Wales can be inferred. However, whether a species has an occurrence record in a particular place and in a particular year is heavily dependent on the intensity of the sampling effort and the innate detectability of the species (MacKenzie et al. 2002). All reliable records, collected over many decades, may therefore be needed to adequately estimate range area (IUCN 2017).

If range area inferred from occurrence records is too insensitive for finer-scale temporal monitoring, a better option is to instead monitor trends in habitat quality (loss, condition or carrying capacity) within the known range of a species. With the advent of satellite-based monitoring of land at high resolution and with sufficient time series, habitat quality can be inferred by calibrating remote-sensing data with field observations (Harwood et al. 2016). **Effective** range area can then be inferred from the overlap of original area (estimated from all reliable occurrence records) and habitat quality. Here we use the ecological condition indicator and the ecological carrying capacity indicator developed using a partially heuristic data-model fusion approach (Love et al. 2020). The calculation of effective range size can then be determined most simply in the case of AOO, where the area of each grid cell is down-weighted by multipying by the corresponding condition or carrying capacity index (which is scaled 0 to 1). This measure will be sensitive to changes in habitat loss through the use of satellite-sensed indicators of habitat condition over relatively short (e.g. annual) timescales and, given a suitably representative set of species, be able to interpret and track these trends in terms of their implications for overall biodiversity.

2.2 Estimating remaining genetic diversity from remaining range area

The genetic diversity of a species is related to the range area of that species. This is for essentially two reasons:

- limits to gene flow between populations that increase with spatial separation
- adaptation of populations to local conditions which become more variable over larger areas.

Thus, we expect that, as a species loses range area due to local extinction of populations, genetic diversity will be lost. Alsos et al. (2012) simulated this process by randomly sampling progressively smaller sets of populations for 27 northern hemisphere plant species and measuring the corresponding proportional loss of within-species genetic diversity. Using these same data, Mimura et al. (2017) showed that the resulting relationship between remaining area and remaining genetic diversity followed a power curve, although the shape of the curve depended on the initial genetic differentiation between populations of a species (Figure 3). Genetic differentiation is defined by the fixation index (F_{ST}), which is the proportion of genetic diversity in a species that is explained by population differences (Holsinger & Weir 2009).

The relationship between genetic diversity and range area can be defined as equation 1 where (G_0/G_1) is the proportional genetic diversity remaining, (A_0/A_1) is the proportion of area remaining, and z is the parameter determining the shape of the curve (Mimura et al. 2017). Area (A_0, A_1) is defined specifically as area of occupancy.

$$(\boldsymbol{G_0}/\boldsymbol{G_1}) = \left(\frac{A_0}{A_1}\right)^{z} \qquad \text{equation 1}$$

It follows that a useful indicator of trends in genetic diversity of a set of species can be derived from trends in AOO for those species. This is a proxy indicator that applies in lieu of samples that directly quantify the actual genetic diversity of populations (such as Alsos et al. 2012). Because low genetic diversity is known to impact on the fitness of individuals through inbreeding depression (Hoffman et al. 2014), trends in AOO may also provide an indirect indicator of population health (Mimura et al. 2017). Note also that this indicator framework can be used for inferences about a broader universe of species, when combined with an approach that uses a representative sample of species (Mimura et al. 2017; Bruford et al. 2017). The logic is that changes in the indicator value for the representative species set indicate the changes that would have been seen for the broader universe of species. We adopt this rationale (see section 2.6 below).

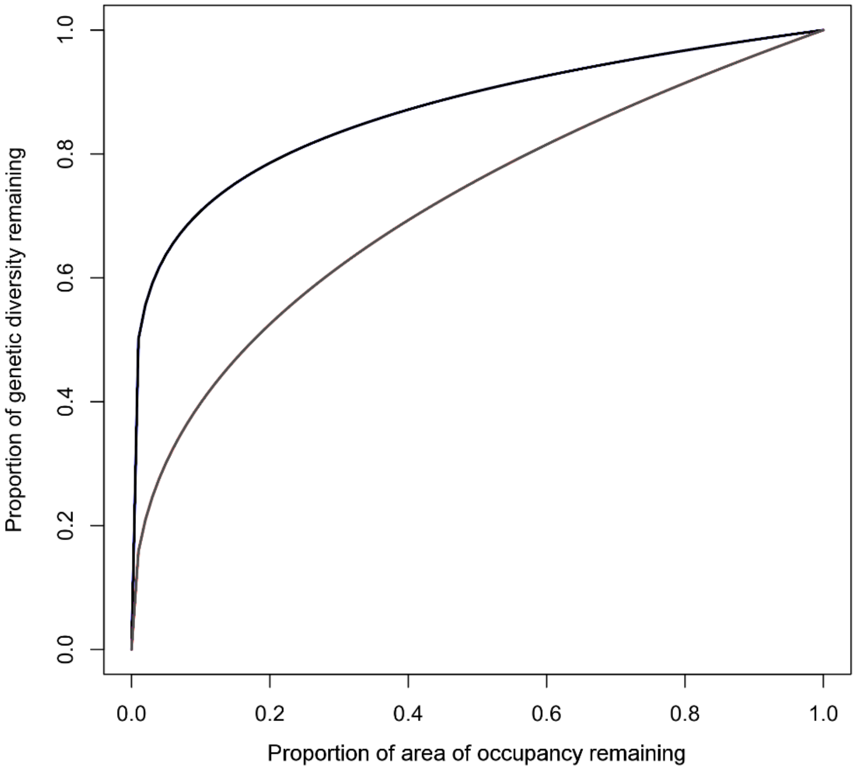


Figure 3 Proportion of genetic diversity remaining modelled as a power function of proportional area of occupancy remaining

The lower curve models the relationship for a species with a relatively high degree of population divergence (F_{ST} ~ 0.9) while the upper curve models a species with a relatively low degree of population divergence (F_{ST} ~ 0.3). Figure adapted from Mimura et al. (2017).

2.3 Expected diversity framework

Expected diversity (Weitzman 1992) measures the rate of loss of biodiversity by determining the amount of biodiversity expected to still exist at some specified point in the future (e.g. 100 years from now), derived from the estimated survival probabilities of species. Slowing the rate of loss is a central goal of the BC Act, as well as global conservation efforts. For species richness, expected diversity $(\mathbb{E}[D])$ is given by equation 2, and is simply the sum of the probability (q_i) of survival of each species (i) across a set of species (S) where $i \in S$ (Weitzman 1992).

$$\mathbb{E}[\boldsymbol{D}] = \sum_{\boldsymbol{i\in S}} \boldsymbol{q_i} \qquad \text{equation 2}$$

As a sum of species' weights (reflecting relative risk of loss), expected diversity resembles the Red List Index (Butchart et al. 2004; Butchart et al. 2007) but is a more directly probabilistic approach that is extensible to a variety of measures of biological diversity.

Rate of loss is simply measured as the change in expected diversity between two time steps(Weitzman 1992; Faith 2008), where each feature (i) may have a new probability $(\acute{q})$ of survival (equation 3).

$$\Delta(\mathbb{E}[\boldsymbol{D}]) = \sum_{\boldsymbol{i\in S}} \boldsymbol{\acute{q}_i} - \sum_{\boldsymbol{i\in S}} \boldsymbol{q_i} \qquad \text{equation 3}$$

2.4 Probabilities of survival of threatened species

While the true survival probabilities of individual species are not known, the IUCN does give estimates for probability of extinction for criterion E for specific threat categories (vulnerable, endangered and critically endangered). Criterion E gives critical thresholds for extinction probabilities (over specified time periods) determined from population viability analysis (Brook et al. 2000). Based on these estimates, Kindvall and Gärdenfors (2003) determined the per-year extinction probabilities for species and then extrapolated for other time periods (Figure 4). From this basis, the *Saving our Species Technical Report* (OEH 2013) derived survival probabilities (for a 100-year time period) in the development of a priority score for selecting projects (Table 1). The *Saving our Species* (SoS) program also included a 'secure' category for non-threatened species which was assigned a survival probability (over 100 years) of 0.95. Note that these survival probabilities do not in any way take into account changes in future conditions (i.e. in 100 years' time); they are simply extrapolations from the stated IUCN probabilities and assume no change in conditions.

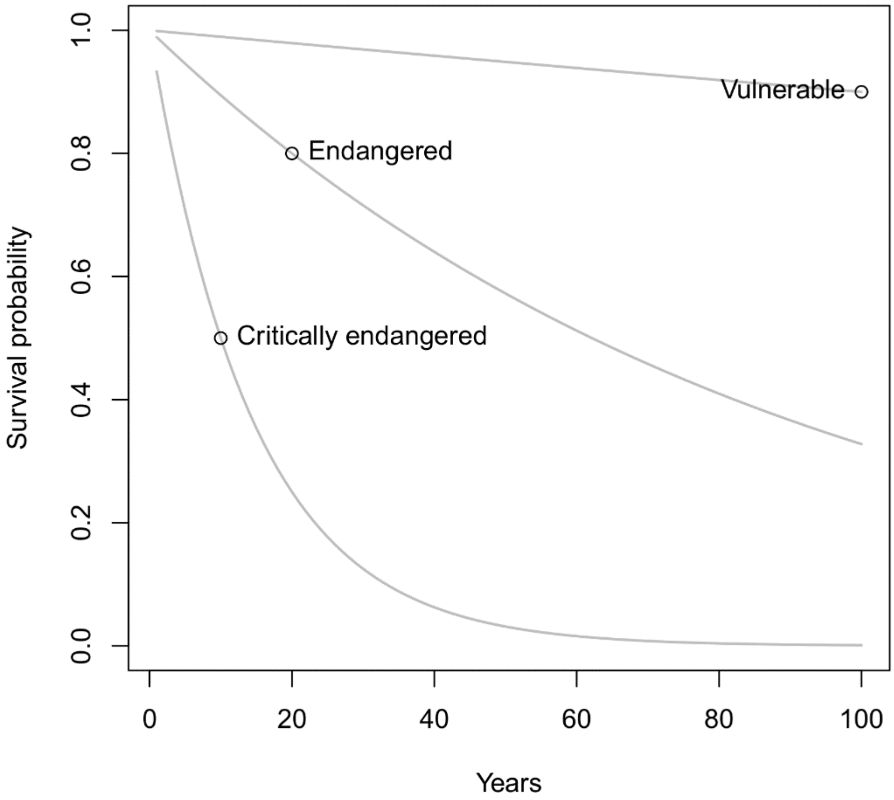


Figure 4 Extrapolation of survival probabilities of categories of threatened species over time derived from criterion E of the *IUCN Red List Categories* (IUCN 2012)

Original IUCN estimates are indicated as CR (critically endangered), EN (endangered) and VU (vulnerable). Calculations follow Kindvall and Gärdenfors (2003).

Table 1 Survival probabilities assigned to threat categories used by the *Saving our Species* program

Risk category	Probability (*q*) of survival in 100 years
Secure	0.95
Vulnerable	0.90
Endangered	0.30
Critically endangered	0.05

Note: Probabilities derive by extrapolation from IUCN criterion E (Kindvall & Gärdenfors 2003).

2.5 Provisional assessment of extinction risk

An expected diversity indicator is most informative when survival probabilities can be assigned to all known species. This requires comprehensive assessments which consider all known species in a major biological group. Such assessments are not within the remit or the resources of the NSW Threatened Species Scientific Committee, who assess species as they are nominated by the public, committee members or the Minister for the Environment.

In the absence of comprehensive assessments, a provisional system of assigning putative risk categories to species was adopted by adapting just one set of assessment criteria: criterion B2 and subcriterion B2(b)(ii) of the *IUCN Red List Categories* (IUCN 2012). The criterion assigns thresholds to AOO of a species to determine conservation status and can be easily applied to occurrence records. AOO is defined as the summed area of 4-km^2 grid cells that intersect with occurrence records for a species (Figure 2). To qualify for a threatened category under criterion B2, the Red List stipulates that a species must also fulfil at least two other conditions, including severe fragmentation of AOO or a restricted number of locations, continuing decline (in AOO or habitat quality) and extreme fluctuations (Table 2). Of these, continuing decline in AOO can be assessed in an automated way by intersecting AOO with a layer of habitat condition (from the ecological condition of terrestrial habitat indicator, see Love et al. 2020). The condition index scales from 0–1 and thus AOO is most simply adjusted by multiplication of area of each grid cell with the corresponding index. Given that a condition index of 1 corresponds to original (pre-industrial) condition, the adjusted AOO corresponds to the effective area of quality habitat. Therefore, we base our provisional risk assessment on current AOO and continuing decline in AOO (inferred from the condition model). Given that this provisional assessment is neither complete nor a formal listing process, we have adopted new risk categories (highest risk, higher risk, lower risk and lowest risk), which have similar criteria to the corresponding categories used by the *IUCN Red List Categories* and the BC Act, but are not entirely concordant with those formal categories.

Table 2 Criteria for assessing conservation status for criterion B2 of the *IUCN Red List Categories*

Red List Criteria	CR	EN	VU
B2. Geographic range in the form of area of occupancy (AOO)	<10 km^2	<500 km^2	<2000 km^2
And at least two of the following three additional criteria:			
a) Severely fragmented or number of populations	=1	≤5	≤10
b) Continuing decline observed, estimated, inferred or projected in any of: (i) extent of occurrence (EOO), (ii) area of occupancy (AOO), (iii) area, extent and/or quality of habitat; (iv) number of locations or subpopulations; (v) number of mature individuals			
c) Extreme fluctuations in any of: (i) extent of occurrence (EOO), (ii) area of occupancy (AOO), (iii) number of locations or subpopulations; (iv) number of mature individuals			

Notes: See Table 2.1 in IUCN (2017).
CR = critically endangered; EN = endangered; VU = vulnerable.
AOO is area of occupancy defined as the area of 4 km^2 containing species occurrence records.
EOO is extent of occurrence defined as the minimum convex polygon enclosing the AOO.
For comprehensive assessment of risk, additional considerations apply (a, b, c).

2.6 Representative species set

An indicator derived from a sample of species from a taxonomic group, often generated from the haphazard accumulation of available information, is not necessarily representative of the entire group. If a sample were truly representative of the group, a given index calculated on that subset of species would be a good indicator of the index value we would have obtained if we had been able to calculate the index for all species.

The environmental diversity framework (Faith & Walker 1996; Faith et al. 2004) provides a rationale and means for determining a set that is representative of the variety of physical environments that determine species composition across New South Wales and is thus inferred to be broadly representative of overall biodiversity (Bruford et al. 2017). First, a generalised dissimilarity model (GDM, Ferrier et al. 2007) is derived that predicts the response of biodiversity (changes in species composition) from changes in key environmental variables. Second, a set of ‘demand points’, equal to the number of representative species required, is determined. The locations of those demand points are determined to be uniformly distributed in the continuous environmental space defined by the GDM, and thus represent the diversity of ecological environments (i.e. of habitats occupied by species in the biological group used in the GDM). The distance between demand points is the predicted dissimilarity in species composition between those points. Third, occurrence records for candidate species are mapped to the environmental space. Finally, the *p*-median algorithm (the 'ED' method of Faith & Walker 1996) is employed to find the nearest species to each demand point, based on the average dissimilarity (distance in environmental space) of a species’ occurrences to that demand point. The result is that each part of environmental space is represented by a species, and no parts of environmental space are over-represented.

To implement this approach, a workflow extension of the environmental diversity – survey gap analysis (ED–SGA) software (Ferrier 2002; Manion & Ridges 2009) is used (Manion 2018). The ED–SGA software has been applied widely to guide the filling of gaps in biological surveys across New South Wales and elsewhere (e.g. Bell et al. 2014; Ferrier et al. 2007; Funk et al. 2005).

The process is as follows (see equation 4 below). For a set of *k* representative species, the first step corresponds to the normal procedure to create *k* demand points (Faith et al. in prep.). Then the algorithm sorts through the full available set of species (each species is associated with a defined list of locations where it has been observed). For each species, *j*, the algorithm finds its closest demand point, P_j (the distance from a species to a demand point is the average dissimilarity of all of species *j*’s observed locations to that demand point). If that demand point does not already have a species assigned to the set, we add that species to the representative set and continue, until all *k* demand points are represented by a species. Iteratively, we may update the assignment of species, so that, for a given demand point, a new species replaces the existing one if that new species is closer. The rationale for deriving this representative sample of species is consistent with the environmental diversity method (Faith & Walker 1996).

Suppose species *j*, has n_j areas, called a_t indexed by *t* where *t* goes from 1 to n_j. The distance, dist(*j*, p_i), from a species *j* to a demand point, p_i is:

$$\boldsymbol{dist(j, p_i)} = \left(\sum_{\boldsymbol{t=1}}^{\boldsymbol{n_j}} \frac{\boldsymbol{D(p_i, a_t)}}{\boldsymbol{n_j}} \right) \qquad \text{equation 4}$$

where *D* is a distance in our environmental space. We then use dist(*j*, p_i) values to assign, as described, each demand point a representative species. An important consideration is the size of the representative species set. As the size of representative set approaches that of the full list of candidate species, it will become more representative of the candidate list, rather than the broader biological group, because environmental diversity will be represented inefficiently by many species occupying the same environmental space (i.e. environmental diversity will be represented redundantly). If the representative set is too small, it will not adequately cover the full range of environmental diversity. Further, if a species selection algorithm is biased towards species with small ranges (because they are more closely associated with distinct areas in environmental space), then indicators derived from the set may be biased towards over-estimating extinction risk. Thus, it is important to generate large numbers of representative sets of different sizes and run diagnostic tests to explore the implications of choosing a set of a particular size (see section 3.4).

The representative set of species is then used in a variety of indicators. The principle being that the set of sampled species (for a given biological/taxonomic group) need to be selected once and only those species are used for the first assessment of indicators and to assess change in the future. Given the representative species set, the value, for any nominated indicator, is the average (or other summary) of the individual scores for the sample of *k* species.

2.7 Extrapolation of total biodiversity from discovery curves

The true number of species on Earth, in Australia, or in New South Wales remains unknown. Possibly more than 80% of the world’s species remain undescribed (Mora et al. 2011). Even in relatively well-known groups like mammals, new species continue to be discovered (Ceballos & Ehrlich 2009). For Australia, about 150,000 species have been described and named but there may be more than 500,000 (Chapman 2009). Discovery curves, which plot the cumulative number of species discovered over time, allow for an assessment of how complete the inventory of species might be for a taxonomic group. As an inventory nears completion, fewer new species should be discovered over time, causing a discovery curve to approach an asymptote (Figure 5). This asymptote can then be interpreted as the total number of species in the group. Statistical methods have been developed to estimate the asymptote from the rate of accumulation of species towards the end point of a discovery curve (Bebber et al. 2007). Thus, it is possible to predict the total number of species in a group, including those species yet to be discovered and described.

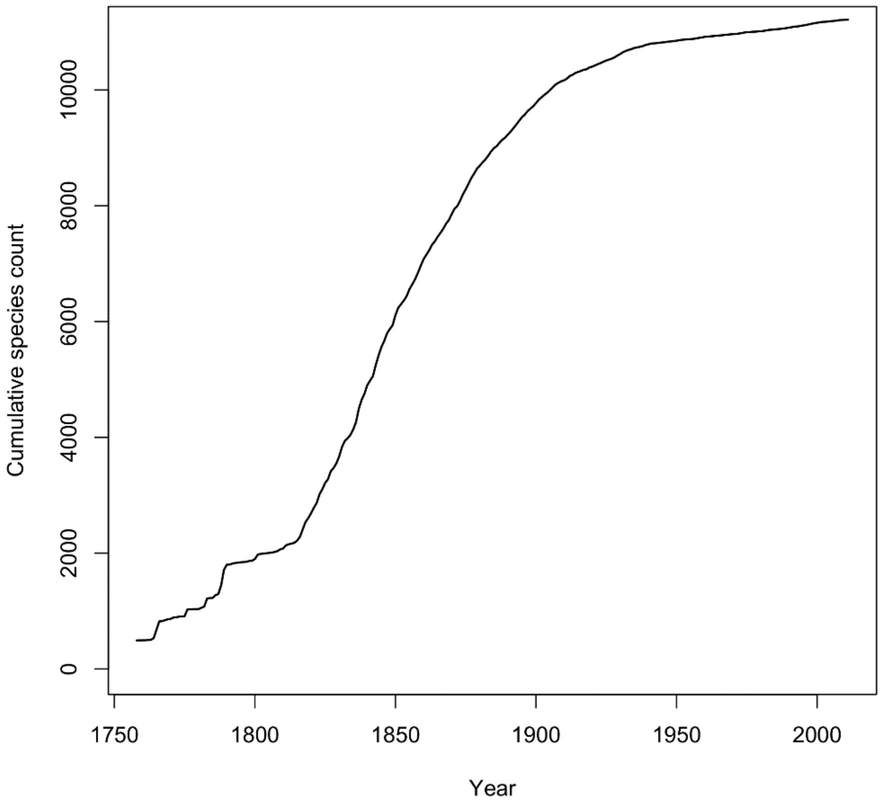


Figure 5 Discovery curve of the world’s birds, plotting year of taxonomic description against cumulative count of species discovered to that date

Data sourced from the Birdlife International Taxonomic Checklist (http://datazone.birdlife.org/species/taxonomy). The species inventory is complete when the curve reaches an asymptote.

3. Method implementation

3.1 Data sourcing, cleaning and manipulation

Range area was calculated from species occurrence records downloaded from the Atlas of Living Australia (Atlas of Living Australia 2018). First, a list of plant species for New South Wales was generated from all NSW occurrence records for plants (kingdom Plantae). Taxa were excluded from this list if they met any of the following criteria:

- not from one of the major subclass of vascular plants (i.e. *Cycadidae*, *Pinidae*, *Magnoliidae*, or a fern subclass)
- not a species-level or subspecific taxon
- not listed as a valid name in *both* the Australian Plant Census (APC, Council of Heads of Australasian Herbaria 2018) and the NSW Plant Information Network System (‘PlantNET’, Royal Botanic Gardens and Domain Trust 2018)
- not listed as native in either the APC or PlantNET.

Data cleaning reduced the original list of 16,501 unique names to 5528 species. All available occurrence records for this set of 5528 species of vascular plants were downloaded from the Atlas of Living Australia. Records were then removed if any of the following applied to them:

- they were not a preserved specimen
- spatial (latitude–longitude) or temporal (year of collection) data were incomplete
- they were a cultivated specimen
- they occurred more than 5 kilometres outside the coastline of Australia
- they were collected prior to 1950
- their coordinate (i.e. location) uncertainty, if known, was greater than or equal to 3000 metres.

Cleaning reduced the initial dataset of 7,802,849 records for 5528 species to 1,243,554 records for 5506 species. These included occurrences anywhere in Australia, of which 4855 species were represented by at least one occurrence within New South Wales (including the Australian Capital Territory, ACT, and Commonwealth properties). A project-specific unique identifier was assigned to each species to enable traceability in subsequent analyses.

Species with at least three occurrences within the NSW state border (excluding ACT and Commonwealth properties) were identified as the starting point for the selection of **representative species sets** (totalling 249,820 records in New South Wales). At least three occurrence records were selected to ensure the species were adequately known from New South Wales. This reduced the set for consideration to 4595 species. Only those occurrence records for the NSW species that overlapped the south-eastern Australia extent of the GDM model (see section 3.2) were used (516,689 records in total) in the representative species selection process (see section 3.4). Figure 6 shows the extent of location records for the 4595 species used in the selection process. NARCliM refers to the NSW and ACT Regional Climate Modelling project (Evans et al. 2014).

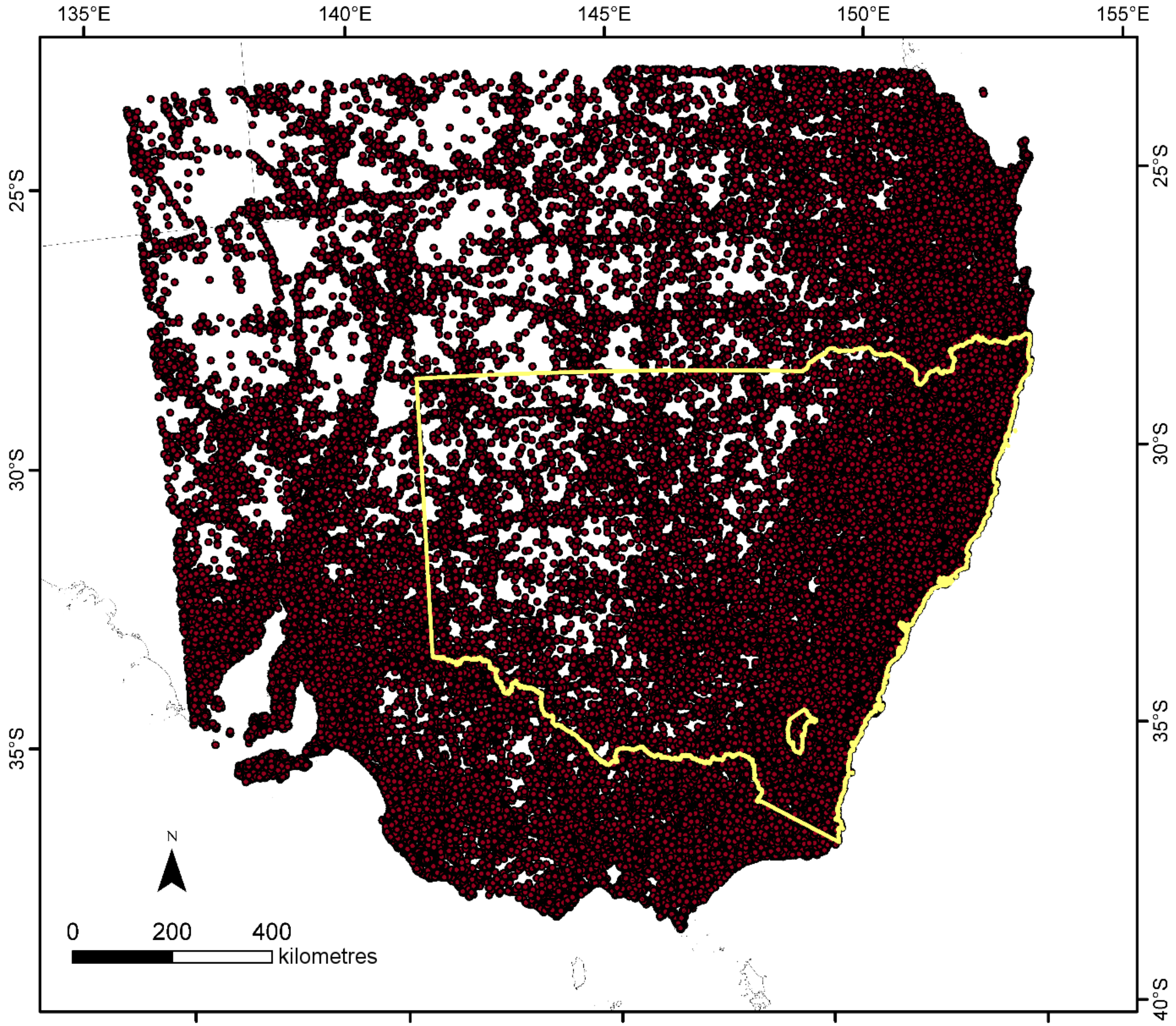


Figure 6 The south-eastern Australia (NARCliM) extent of location records for the 4595 species with at least three occurrences in New South Wales, that were used in the representative species set selection (see section 3.4 for details)

To estimate the total number of vascular plant species in New South Wales, a list of all native species in New South Wales was downloaded from PlantNET (Royal Botanic Gardens and Domain Trust 2018). The species list of 6062 species were matched to species names in the Atlas of Living Australia (which are standardised to the Australian Plant Census) and the earliest NSW occurrence record for the matched species recovered. Of the 6062 species in the PlantNET species list, 5304 were matched via original name or synonym to the Australian Plant Census, and of these, 5219 were matched to a year of earliest NSW record. The earliest records were used to generate a discovery curve from which an asymptote could be estimated.

3.2 Generalised dissimilarity model and condition model

An existing GDM for vascular plants (OEH 2016; see also Drielsma et al. 2017) was used to represent the diversity of ecological environments occupied by vascular plants in New South Wales. This model is the same as that applied by Drielsma et al. (2018) to derive an ecosystems classification for the assessment of the state of biodiversity persistence of ecosystems (including undiscovered species) indicator (see Figure 7). It was based on vascular plant survey data from across south-eastern continental Australia and used 25 environmental predictors (19 climate, 6 substrate) (OEH 2016). This south-eastern Australia region is commonly known as the 'NARCliM Domain' (Evans et al. 2014). It was designed to capture the cross-jurisdiction interactions between climate and potential change in species distributions over time, for use in modelling and adaptation simulation studies. The spatial extent of this model determined the extent of plant species records used in the selection of representative species (see sections 3.1 and 3.3). This dataset was derived using 250-metre gridded environmental variables for Australia.

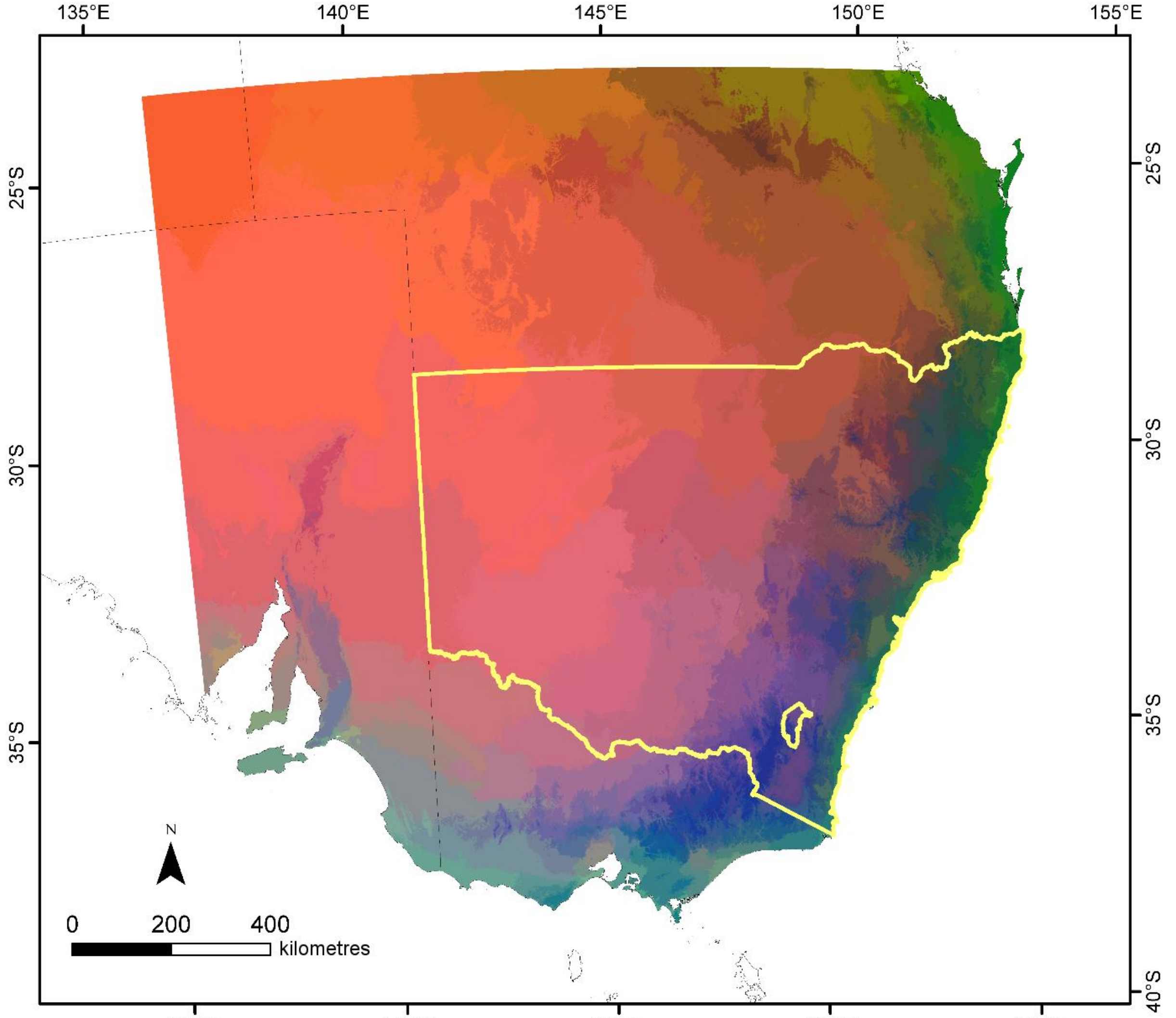


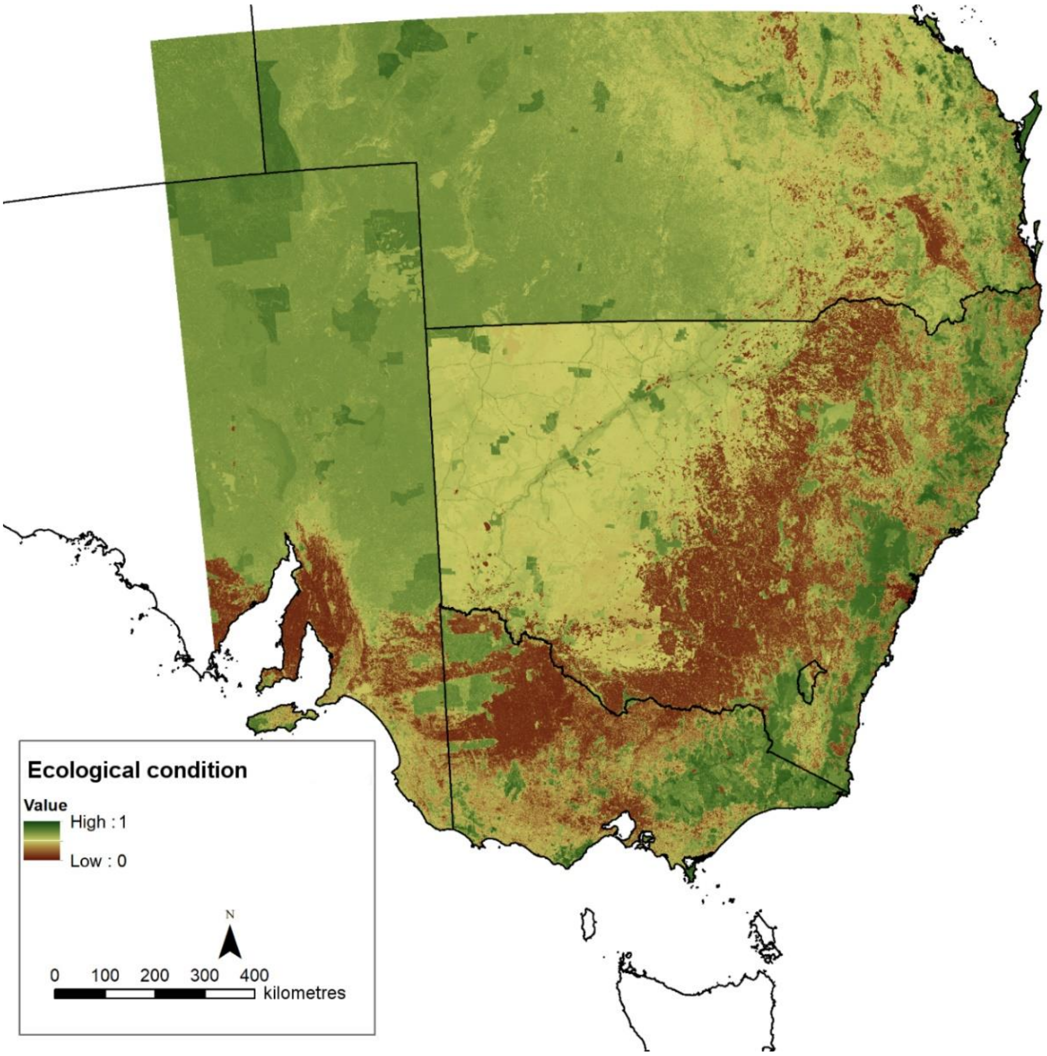
Figure 7 Predicted pattern of vascular plant species diversity based on a 250-group classification of the GDM model (OEH 2016)

Each class is coloured according to its species compositional overlap with other classes. Classes with similar colours are predicted to have similar plant species compositions. Source: OEH (2016); data as used in Drielsma et al. (2018).

The ecological condition indicator developed by Love et al. (2020) was used to infer declines in habitat condition and therefore reductions in occupancy for the representative species. This indicator measures the intactness and naturalness of terrestrial vegetation as habitat to support biodiversity, without considering the indirect effects of fragmentation or connections with surrounding suitable habitat. It is an extension and refinement of an earlier model of terrestrial native vegetation condition developed by Drielsma et al. (2012). The partially heuristic model is designed to directly estimate the amount and quality of generalised habitat for native species and ecosystems at each location, using relevant information such as remote-sensing time series of vegetation fractional cover, land-use and soil condition mapping. The source data currency ranges from 2003 to 2017 and is intended to be representative of habitat up to 2013. The data was developed using a 90-metre grid. This dataset was extended to cover the geographic extent of the GDM model using an earlier version of vegetation condition developed for mapping biodiversity benefits (OEH 2016; Drielsma et al. 2017) (see Figure 8-A). That dataset was developed using a 250-metre grid, to match the GDM model of equivalent extent.

The ecological carrying capacity indicator developed by Love et al. (2020) additionally accounts for the effects of fragmentation by adjusting the ecological condition of a location according to the quality of habitat within a neighbourhood (see Figure 8-B). This indicator is a more accurate reflection of the contribution each location makes to the potential carrying capacity of a landscape in supporting its original complement of biodiversity. A species whose range is fragmented has an increased extinction risk (Crooks et al. 2017) because movement processes, such as foraging, dispersal and migration, will be limited, increasing genetic divergence among populations and inbreeding within populations (Noël et al. 2006). The dataset currency is 2013 and was developed using a 90-metre grid, as for ecological condition (Love et al. 2020).

A. Ecological condition



B. Ecological carrying capacity

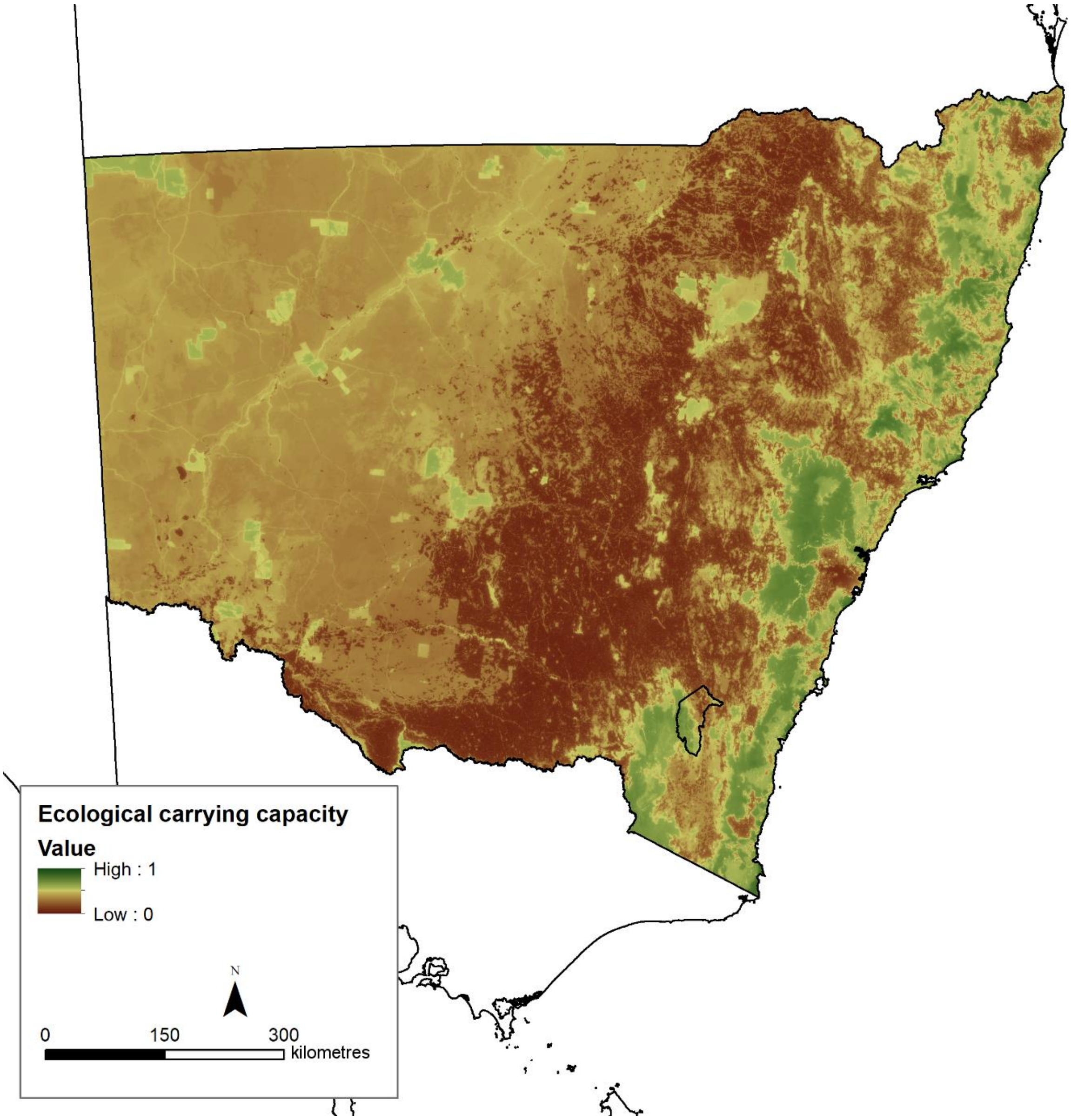


Figure 8 The ecological condition (A) and ecological carrying capacity (B) indicators by Love et al. (2020), with a currency approximating 2013

Both indicators apply within New South Wales only. The habitat condition layer developed by OEH (2016) was used to extend the ecological condition dataset (A) beyond New South Wales, so that it matched the NARCliM extent. Source for both datasets: Love et al. (2020).

3.3 Calculation of range area

For each of the 4595 species used as an input to the selection process (see section 3.1), we calculated range area in two ways (EOO and AOO) and across two regions (the state of New South Wales and the NARCliM project extent). We further calculated the reduction in AOO using two habitat condition indicators (ecological condition and ecological carrying capacity, described section 3.2). These calculations did not require knowledge of which would subsequently be selected as representative, other than to maintain the project-specific unique identifiers assigned to each species to enable relational joins.

As outlined in section 2.1, the AOO calculation for each species is standardised to 4 km^2 using 2 kilometre x 2 kilometre grids. However, multiples of the 90-metre grid ecological condition and carrying capacity datasets did not completely match the 2-kilometre grid, being between 21 and 22 times smaller (i.e. 1926 m and 2018 m respectively). We therefore conservatively generated the '2-kilometre' uniform grid to comply with the 21 times dimension of the condition datasets (specifically, 1926.03367404 metres), using the NARCliM extent of the GDM model as the spatial reference with the 'polygon to raster' conversion tool in ArcGIS$^{\text{TM}}$ 10.5 (ESRI 2016). This uniform grid dimension allowed aggregation and assignment of condition values to each species' AOO in a traceable manner. We could then also use the 90-metre grid extent of the habitat condition datasets in New South Wales to account for those portions of the uniform '2-kilometre' gridded AOO for each species within New South Wales relevant to the final indicator.

The uniform grid was converted to a point dataset (using the 'raster to point' conversion tool in ArcGIS$^{\text{TM}}$ 10.5) and the XY coordinates attached as attributes (using the 'add XY coordinates' data management tool in ArcGIS$^{\text{TM}}$ 10.5). Each row in the point dataset also has a unique number. This process then allowed the individual species locations to be spatially assigned the attribute values of the nearest centroid of the uniform grid (using the 'near' analysis tool in ArcGIS$^{\text{TM}}$ 10.5), as the AOO reference cell for each species' occurrence record. A summary analysis of the resulting dataset removed replicate occurrences of the same species with the same grid identifier. The number of unique AOO cells were then summed for each species to calculate the total AOO. This calculation was repeated for the NARCliM analysis extent (used in the selection of representative species) and the NSW state border extent (used in the calculation of the indicator).

The average habitat condition of each AOO grid cell was determined using 'zonal statistics as table' spatial analyst tool in ArcGIS$^{\text{TM}}$ 10.5, using the ecological condition and carrying capacity datasets as the value rasters, separately. Across the NSW border, the proportion of each grid cell within New South Wales was determined and a correction applied to the habitat condition values in those cells. The adjusted AOO was then calculated for each species as the weighted sum of the 4-km^2 area of unique AOO cells for each species, for the respective condition dataset. That is, the area of each grid cell in the AOO was multiplied by ecological condition or ecological carrying capacity index (which both scaled 0–1) to produce an effective area of habitat per grid cell, and these values were summed to get total effective area within a species' range.

The EOO of each species was simply calculated from the NARCliM extent of the original occurrence records using the 'minimum bounding geometry' data management tool in ArcGIS$^{\text{TM}}$ 10.5 with the 'convex hull' geometry option, followed by the 'add geometry attributes' data management tool to assign area in kilometres squared to each polygon. The NSW extent of each range polygon for each species was then determined by intersecting with the state boundary and recalculating the area.

The NARCliM extents of the EOO and AOO calculations for each species were then available for use in diagnostics tests of the representative species selection process or interpretation of the results, as described in section 3.4.

3.4 Representative species set selection

The process for selecting a representative species set first required an optimal number of demand points to be determined. An optimal number of points would avoid over-representation of environmental space (too many points) while also avoiding bias towards representative species with small ranges (too few points). Diagnostic measures of environmental representation and range size were used to make this choice (details below). To run the diagnostic tests, demand point generation and species selection were repeated for every possible number of points from 1 to 4595 (the number of species in the cleaned occurrence record dataset).

Initial tests showed that the species composition of the representative set produced by the selection algorithm was influenced by the starting order of candidate species, with early selections of the algorithm tending to be at the beginning of the list. So, to avoid potential bias, nine random orderings of candidate species were generated (in addition to the original, alphabetical ordering) and then used as input for the selection algorithm. These random orderings will henceforth be referred to as ‘random starts’. Thus, the species selection algorithm was run for every number of one to 4595 demand points, for each of the nine random starts using the ‘species selection’ operation mode (as shown on Figure 9). Variance in average range area and expected diversity among random starts was very low and so nine random starts was considered sufficient, especially considering the large computational load required to generate representative sets for every set size from one to 4595 for every random start.

A new option to enable a representative set of species to be selected, as described in section 2.6, was developed and embedded within the existing *.NET Survey Gap Analysis Tool* (Manion 2018). The new application tool ‘derive a representative sample of species from demand point set and site data via ED’ (Figure 9) is a new function in the drop-down tools menu and is accompanied by a data preparation tool ‘build full predictor table from (“spp_code, x, y”)’. The new tool requires two inputs:

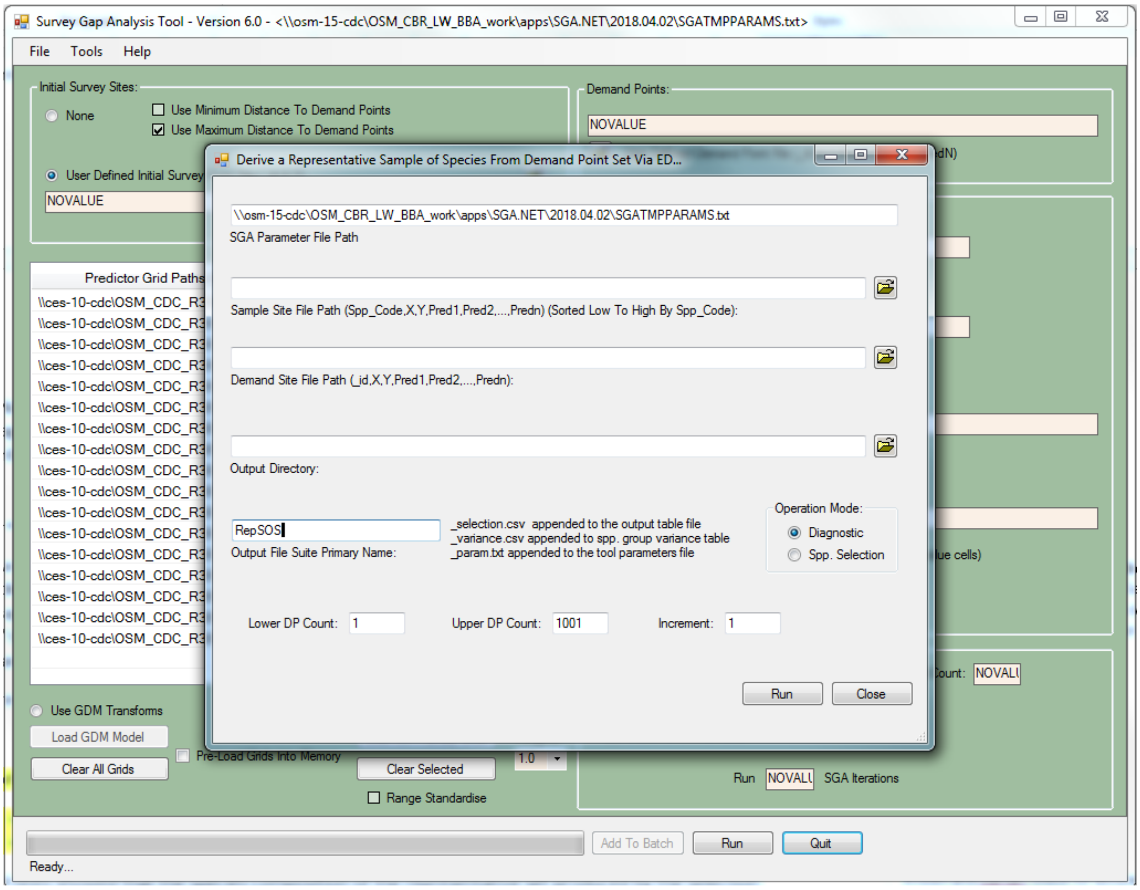
- the location records (XY coordinates) for the species attributed with the values of the environmental variables scaled by the GDM model (see section 3.2)
- a set of demand points equal in number to the total number of species available for selection.

Two operation modes are available:

- diagnostic, which is an exploratory function enabling the user to test different numbers of iterations and increments for rapid assessment of diagnostic to determine selection sufficiency
- species selection mode, for setting the selection of a certain number of species and output datasets listing the selected species.

The diagnostic outputs are also provided in the species selection mode, but the species selections are not provided in the diagnostic mode. Because the selection process is iterative, single increments through thousands of selections can take several days, depending on the dimensions of the search and analysis.

Before the species selection tool is applied, the optimal demand points must be generated. Demand point generation is also an iterative process. The 25 environmental variables scaled by the fitted GDM model (see section 3.2) define the hyper-dimensional space within which a uniform distribution of 4595 demand points is required. This number is determined to be equivalent to the number of species in the input file that have at least three occurrence records (different locations) within New South Wales (see section 3.1). The method uses a random seed location and then iteratively searches for the next location that is optimally the furthest away in the hyper-dimensional environmental space, until the maximum desired number has been achieved. This process can also take several days or weeks to complete, especially if all grid cells within the analysis region are used in the search. A faster operation is available that uses a spatially uniform sample of all grid locations. The sample table is ideally at least 1% of the total number of grid cells available for selection, and the XY coordinates are attributed with the values of the 25 GDM-scaled environmental variables, as the input. We used this option, based on one million samples from a data grid of 12,293,727 cells to derive 4595 demand points. The selection of demand points was further constrained to only those regions within New South Wales where the GDM model predictions were interpolated (i.e. not extrapolated) within the environmental range of the source data used to build the model (i.e. the original vascular plant species occurrences, described in OEH 2016). The geographic distribution of the 4595 demand points is given in Figure 10.


Survey Gap Analysis Tool - Version 6.0 - <\\osm-15-cdc\OSM_CBR_LW_BBA_work\apps\SGA.NET\2018.04.02\SGATMPPARAMS.txt>

Initial Survey Sites:
○ None
☐ Use Minimum Distance To Demand Points
☑ Use Maximum Distance To Demand Points
● User Defined Initial Survey
NOVALUE

Demand Points: NOVALUE

Derive a Representative Sample of Species From Demand Point Set Via ED...

SGA Parameter File Path: \\osm-15-cdc\OSM_CBR_LW_BBA_work\apps\SGA.NET\2018.04.02\SGATMPPARAMS.txt

Sample Site File Path (Spp_Code,X,Y,Pred1,Pred2,...,Predn) (Sorted Low To High By Spp_Code):

Demand Site File Path (_id,X,Y,Pred1,Pred2,...,Predn):

Output Directory:

Output File Suite Primary Name: RepSOS

_selection.csv appended to the output table file
_variance.csv appended to spp. group variance table
_param.txt appended to the tool parameters file

Operation Mode:
● Diagnostic
○ Spp. Selection

Lower DP Count: 1
Upper DP Count: 1001
Increment: 1

○ Use GDM Transforms
☐ Pre-Load Grids Into Memory
☐ Range Standardise
Run NOVALU SGA Iterations

Figure 9 The input interface for deriving a representative sample of species embedded within the existing *.NET Survey Gap Analysis Tool* (Manion 2018)

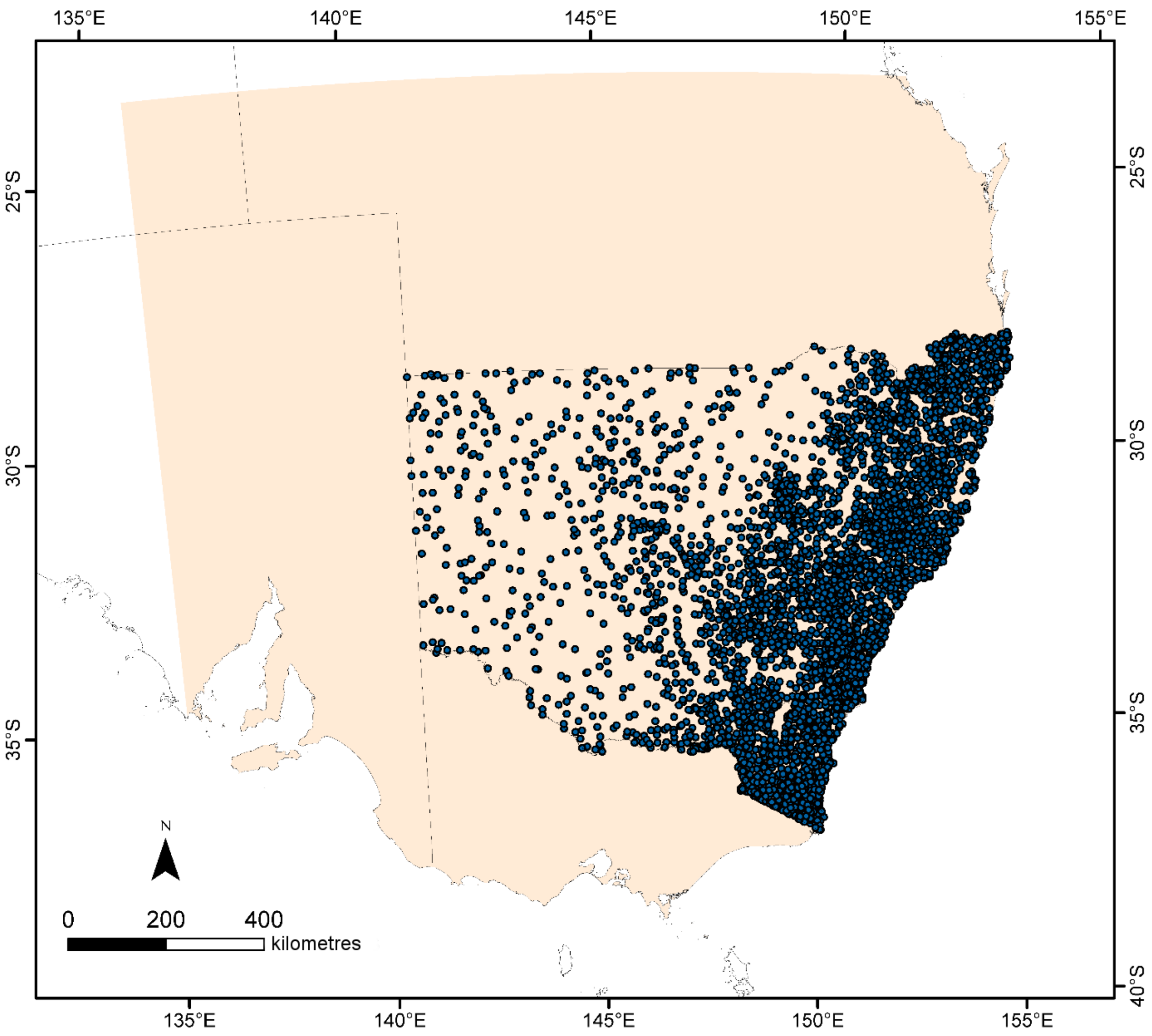


Figure 10 Geographic distribution of 4595 demand points selected to evenly sample the environmental space in New South Wales defined by a GDM of vascular plants based on 25 environmental predictors. The background shade defines the NARCliM extent.

Having selected the optimal set of demand points within New South Wales, the representative species selection tool is then run using the species occurrences records across the entire NARCliM extent attributed with the values of the 25 GDM-scaled environmental variables. The selection method then iteratively searches for the species with the shortest average distance of occurrences in environmental space to each demand point, and then retrospectively tests the optimality of this selection at each subsequent iteration after making the next selection.

Two diagnostics were used to inform the choice of the size of the representative species set:

- an index of unrepresented environmental space
- mean range size of the species of the set.

The index of unrepresented environmental space was calculated as the largest mean distance of a species' occurrences *not* included in the set to its nearest demand point. The diagnostic of mean range size was calculated as the mean EOO and AOO of species included in the set and included the area found outside of New South Wales. These diagnostics were calculated for representative sets of all sizes from 1 to 4595 species and for all nine random starts of the selection algorithm (Figure 11 and Figure 12). From the diagnostics, it was decided that a **representative species set of about 25% (i.e. 1149 species)** would be adequately environmentally representative without including many redundant species (shown as the long plateau in Figure 11) while also avoiding strong bias towards species with small ranges.

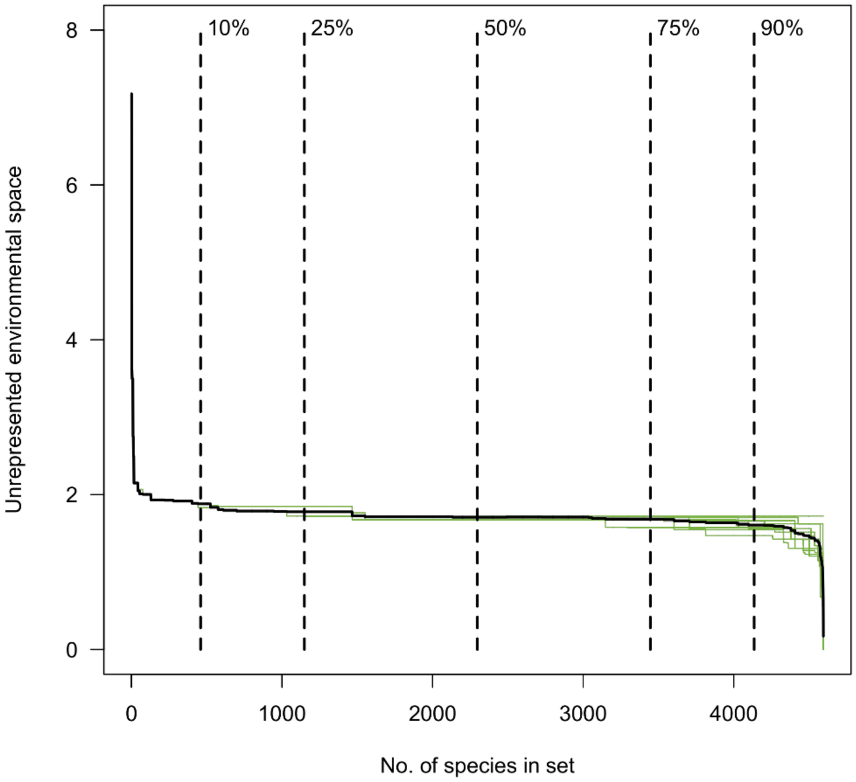


Figure 11 Diagnostic plot of environmental representativeness for the full range of sizes of representative species sets for nine random starts of the selection algorithm

Diagnostic index (y-axis) is the largest distance of a species not included in the set to its nearest demand point. The curve for each random start is given in green (the lighter colour). The black curve is the mean across random starts. Dashed vertical lines show the points where the indicated percentages of the species have been selected.

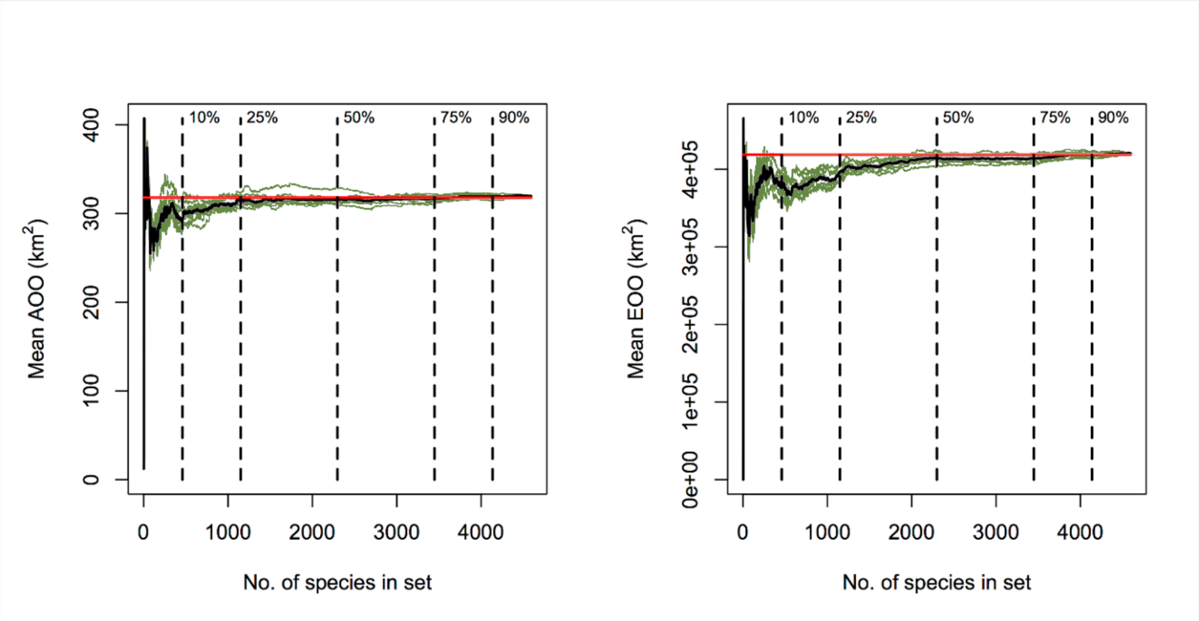


Figure 12 Diagnostic plot of mean range size for the full range of sizes of representative species sets for nine random starts of the selection algorithm

Diagnostic index (y-axis) is the average area of occupancy (AOO) on the left, or extent of occurrence (EOO) on the right, of species included in the set. The curve for each random start is given in green (the lighter colour). The black curve is the mean across random starts. The horizontal red line is the mean range size for all species in the occurrence dataset. Dashed vertical lines show the points where the indicated percentages of the species have been selected.

3.5 Indicator: Expected survival of all known species

3.5.1 Provisional risk assessment

All species in the occurrence record dataset were assigned a risk category. Following the *IUCN Red List Guidelines*, current AOO, after down-weighting by condition (ecological carrying capacity was not used for this indicator), and proportional decline in AOO were adopted as the dimensions of a simple risk assessment model (Figure 13). Critical thresholds for AOO were taken directly from criteria B2 of the *IUCN Red List Categories* (IUCN 2012). Critical thresholds for proportional decline in AOO were taken from criterion A of the *IUCN Red List of Ecosystems* (Bland et al. 2017), specifically for declines observed over the past 50 years. Critical thresholds for decline were adopted because the ecological condition index reports some degree of decline within the ranges of *all* NSW vascular plant species. If a literal interpretation of IUCN *Red List Guidelines* were adopted (where any amount of decline is acceptable for subcriterion B2(b)(ii)), then *all* NSW species would be threatened; an outcome that was neither plausible nor conservative. When a species had an AOO less than or equal to a critical threshold *and* a proportional decline in AOO greater than or equal to a critical threshold, then it was assigned a risk category and corresponding survival probability (Table 3). Risk categories were considered approximately equivalent to the vulnerable, endangered and critically endangered categories used by the IUCN Red List applied to species assessments (IUCN 2017) and were assigned the corresponding survival probability (see section 2.4).

Table 3 Risk categories adopted in this study along with the critical thresholds for current area of occupancy (AOO) and decline in condition of AOO, and the corresponding survival probability

Risk category	AOO	Decline in AOO condition	Survival probability
Lowest risk	> 2000 km^2	≥ 0%	0.95
Lower risk	≤ 2000 km^2	≥ 30%	0.90
Higher risk	≤ 500 km^2	≥ 50%	0.30
Highest risk	≤ 10 km^2	≥ 80%	0.05

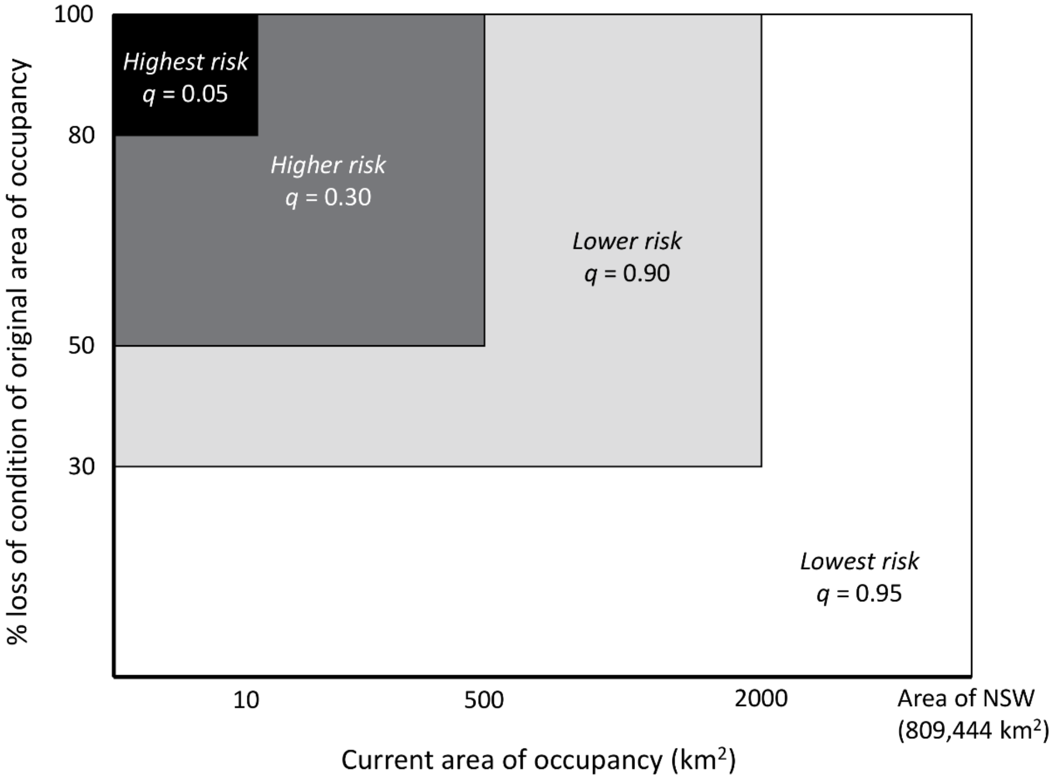


Figure 13 Criteria for determining a provisional risk assessment of a species based on size and trends in area of occupancy

When a species has an area of occupancy (AOO) and percentage loss of condition of AOO such that it falls within a shaded box, it is assigned a survival probability (over 100 years) as indicated. Species that fall in the white space outside the coloured boxes are considered lowest risk and assigned a survival probability of 0.95 (over 100 years).

3.5.2 First assessment value (up to 2013)

Values for expected diversity applicable up to 2013 were calculated for all 4855 plant species in the occurrence record dataset and for all nine representative species sets (each of 1149 species). Expected diversity was simply the sum of the survival probabilities for each species (see equation 2). Expected diversity is also reported as a proportion (of the number of species expected to be still surviving in 100 years).

3.5.3 Null model analysis

Expected diversity of the representative species set was compared to a null model where a set of 1149 species (the same size as the representative set) was selected at random. This was to test whether the representative set is different from a randomly selected set of species. The random draw was repeated 999 times to generate a distribution of expected values. The expected diversity of all nine representative species sets, and the mean across sets, were compared to the null distribution. An expected diversity value would be considered significantly different from the null model if it was lesser or greater than the 95% confidence interval of the randomisations (two-tailed test).

3.5.4 Extrapolation to known and undiscovered biodiversity

A discovery curve of the NSW vascular plant species was generated from the dataset of first NSW records for the PlantNET native species list. Following Bebber et al. (2007), the slope of the discovery curve was modelled as a relationship between the number of new species discovered at time *t* and the cumulative number of species discovered prior to time *t* (i.e. *t-1*). If a simple linear regression model can be fitted to this relationship, then the asymptote (total number of species, including undiscovered species) can be estimated from the intercept of the regression model. That is, given a linear regression with slope, b, and intercept, c, the asymptote is estimated as $-c/b$.

However, discovery curves typically have multiple ‘phases’ (Bebber et al. 2007; Nipperess 2015), meaning that a single linear fit is unlikely to describe the entire slope over time. Therefore, segmented regression analysis was used to fit separate linear models to different parts of the slope. In this case, only the last segment is relevant to estimating the asymptote. Models with one segment (simple linear regression), two segments and three segments were trialled.

Because earliest dates could be recovered for only 5219 species out of the 6062 in the PlantNET native species list, the asymptote was used to estimate the *completeness* of the species inventory (that is, the proportion of all species that have been discovered) (Haque et al., 2017). Assuming that the PlantNET list has a similar level of completeness, the total number of vascular plant species in New South Wales is 6062 divided by completeness (estimated from the sample).

3.6 Indicator: Extant area occupied (for all known species)

The indicator was calculated as the proportion of the original AOO that remains effective based on the ecological condition or ecological carrying capacity indicators assessed up to 2013. That is, the down-weighted AOO (area multiplied by the mean condition or carrying capacity within 2 kilometre x 2 kilometre grid cells, summed across grid cells) divided by the unweighted AOO (summed area of grid cells occupied since 1950). The indicator was calculated for all 4855 NSW plant species in the statewide occurrence dataset and for all nine representative species sets (each of 1149 species).

3.7 Indicator: Within-species genetic diversity for all known species

The indicator was calculated as a power transformation of the extant area occupied indicator. That is, the proportion of AOO that remains effective as habitat for biodiversity was raised to the power of the scaling parameter, z (see equation 1). Because plant species vary in the degree of genetic divergence among their constituent populations, two different values of z were used. For species with low genetic divergence ($F_{ST} \sim 0.3$), the scaling parameter was set as $z = 0.15$. For species with high genetic divergence ($F_{ST} \sim 0.9$), the scaling parameter was set as $z = 0.40$. These values match those used by Mimura et al. (2017) and are based on the range of genetic divergence values observed by Alsos et al. (2012). The indicator was calculated for all 4855 NSW plant species in the statewide occurrence dataset and for all nine representative species sets (each of 1149 species).

4. Results

4.1 Area of occupancy adjusted by ecological condition or ecological carrying capacity

The AOO, adjusted by ecological condition or ecological carrying capacity, was calculated for each species in the occurrence dataset. The frequency distributions for all species and for a selected representative species set are given in Figure 14.

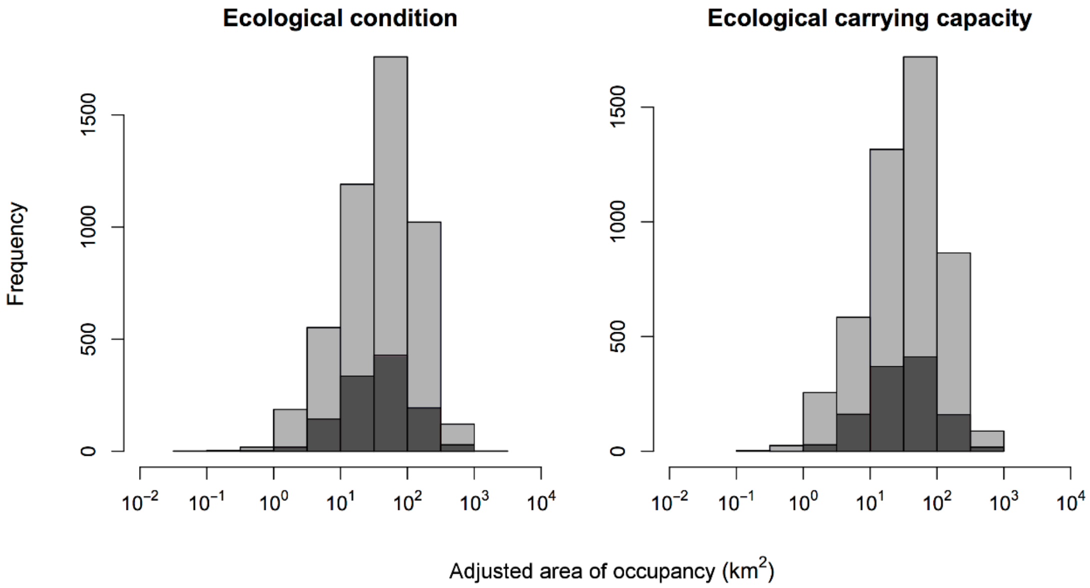


Figure 14 Frequency distribution of area of occupancy after down-weighting by (left) ecological condition of terrestrial habitat indicator; and (right) ecological carrying capacity from the ecological carrying capacity of terrestrial habitat indicator

The distribution of the full dataset of plant species is given by the height of the upper (light grey) columns and for a selected representative set (the second random start) by the lower (dark grey) columns.

4.2 Proportion of area of occupancy remaining

The proportion of original AOO remaining for each species in the occurrence dataset was calculated for both ecological condition and ecological carrying capacity. The frequency distributions for all species and for a selected representative species set are given in Figure 15.

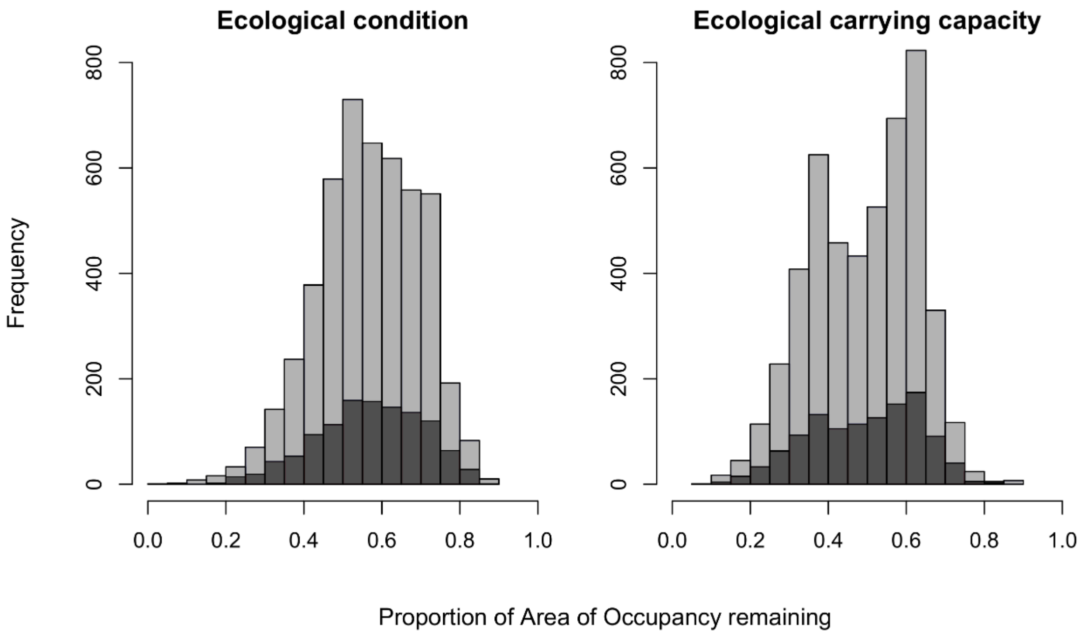


Figure 15 Frequency distribution of proportion of original area of occupancy remaining after down-weighting by (left) ecological condition of terrestrial habitat indicator, and (right) ecological carrying capacity from the ecological carrying capacity of terrestrial habitat indicator

The distribution of the full dataset of plant species is given by the height of the upper (light grey) columns and for a selected representative set (the second random start) by the lower (dark grey) columns. The sum of the values corresponds to the number of species known to exist indicator.

4.3 Provisional risk assessment

Species were assigned a provisional risk category based on their AOO (adjusted by ecological condition only) and proportion of AOO remaining, after adjustment by ecological condition only (Figure 16). Counts of species in each risk category are given in Table 4.

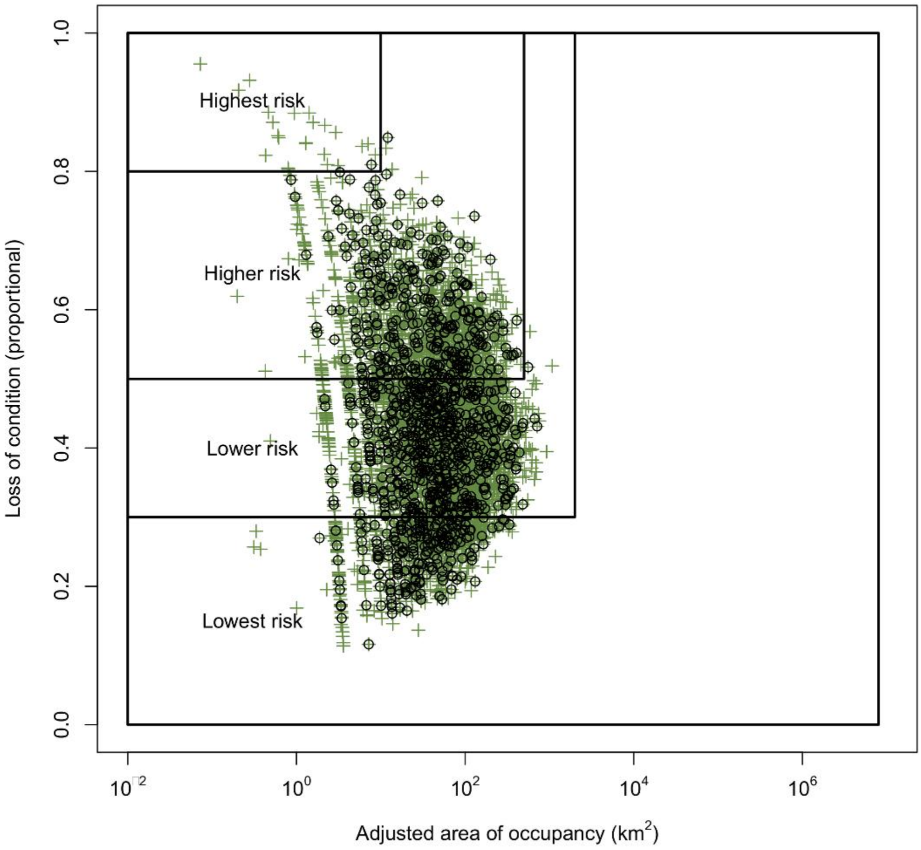


Figure 16 Location of plant species in a risk assessment space defined by area of occupancy (AOO), adjusted by the ecological condition of terrestrial habitat indicator, and proportional loss of condition of original AOO

Green crosses are all 4855 species in the occurrence dataset. Black circles are those species in a selected representative species set (the second random start). Right-hand boundary of Lowest Risk category is set to the total area of New South Wales (~ 8 x 10^6 km^2).

Table 4 Percentage of species in each putative risk category

Risk category	Percentage of species		
	All NSW plants in dataset	Representative species sets (mean)	Listed species in NSW dataset
Highest risk	1%	<1%	1%
Higher risk	30%	30%	32%
Lower risk	58%	51%	43%
Lowest risk	12%	19%	24%

Note: Proportions are given for all 4855 NSW vascular plant species in the cleaned occurrence dataset and for the mean of representative species sets. Proportions are also given specifically for the subset of NSW species that are listed as threatened under the BC Act.

4.4 First assessment for indicator: Expected survival of all known species

By assigning survival probabilities to each risk category, the number and proportion (up to 2013) of all species expected to survive in 100 years were calculated (Table 5). Expected survival was calculated for all NSW plants in the dataset and for each representative species set. The results for representative species sets were also compared to a null model (Figure 17) of 999 random selections of species from the full dataset of the same size (1149 species). Results for all representative sets (and the mean) were within the 95% confidence limits of the randomisations and thus none were significantly different from the null model.

Based on the provisional risk assessment, about 73% of all NSW vascular plant species and the representative species sets are expected to survive in 100 years.

Table 5 First assessment results (up to 2013) for the indicator: Expected survival of all known species

	Expect to survive in 100 years	
Set	**No. of species**	**Percentage of species**
All NSW plants in dataset	3527.5	72.7
Representative species sets (mean ± SD)	836.57 ± 8.11	72.8 ± 0.1

Notes: Indicator is given as the number and percentage of species expected to survive in 100 years.
The number of species for ‘All NSW plants in dataset’ is the number expected to survive out of the occurrence dataset of 4855 species and should not be read as an estimate of survival for all plant species (known and unknown) in New South Wales.
For the ‘Representative species sets’, results are for the mean and standard deviation across nine different sets generated from nine random starts of the species selection algorithm.

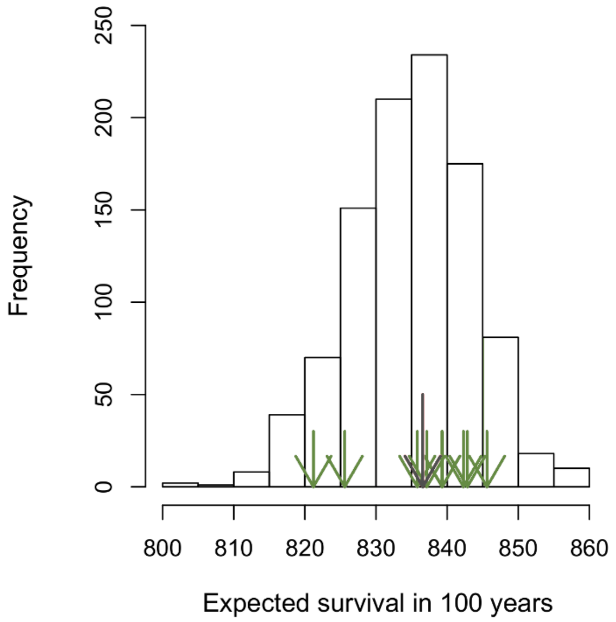


Figure 17 Frequency distribution of expected diversity values of 999 random selections of 1149 from 4855 species from the occurrence dataset

Expected diversity of each of nine representative species sets are indicated by the shorter green arrows. The mean across sets is shown by the longer arrow. Results for all representative sets fall within the 95% confidence limits of the randomisations.

4.4.1 Extrapolation to known and undiscovered biodiversity

Of the 6062 species listed as native in PlantNET, 5219 could be matched to an earliest record in the Atlas of Living Australia. The discovery curve for the matched species is given in Figure 18. Segmented regression analysis of the slope of the curve fitted three linear segments of which the third (most recent) was used to estimate the asymptote of the discovery curve (Figure 19 and Figure 20). Variance explained on the complete model was poor (adjusted $r^2 = 0.15$) but the fit to the final segment (from which the asymptote was estimated) was good (adjusted $r^2 = 0.80$).

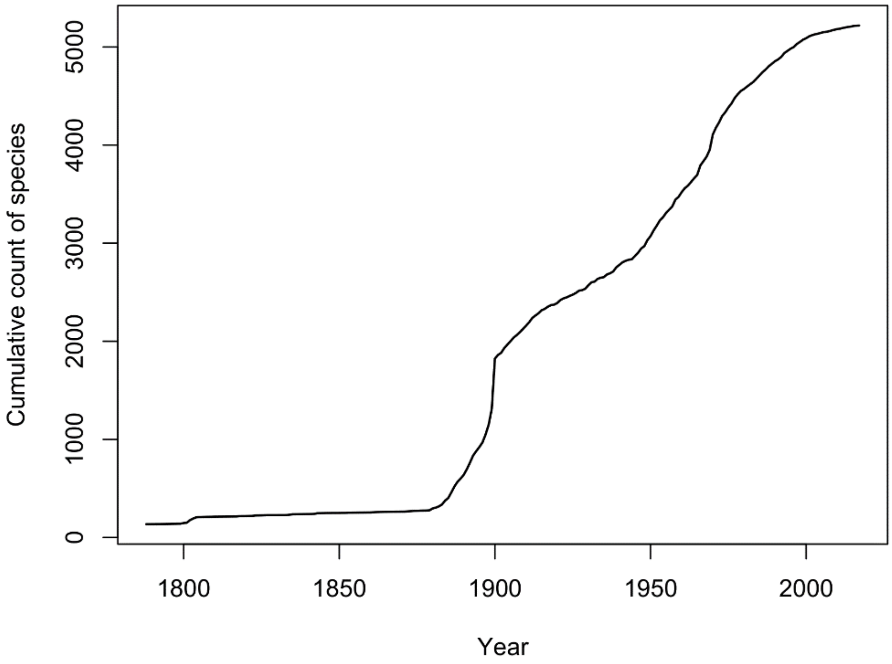


Figure 18 Discovery curve showing the accumulation of described plant species occurring in New South Wales over time

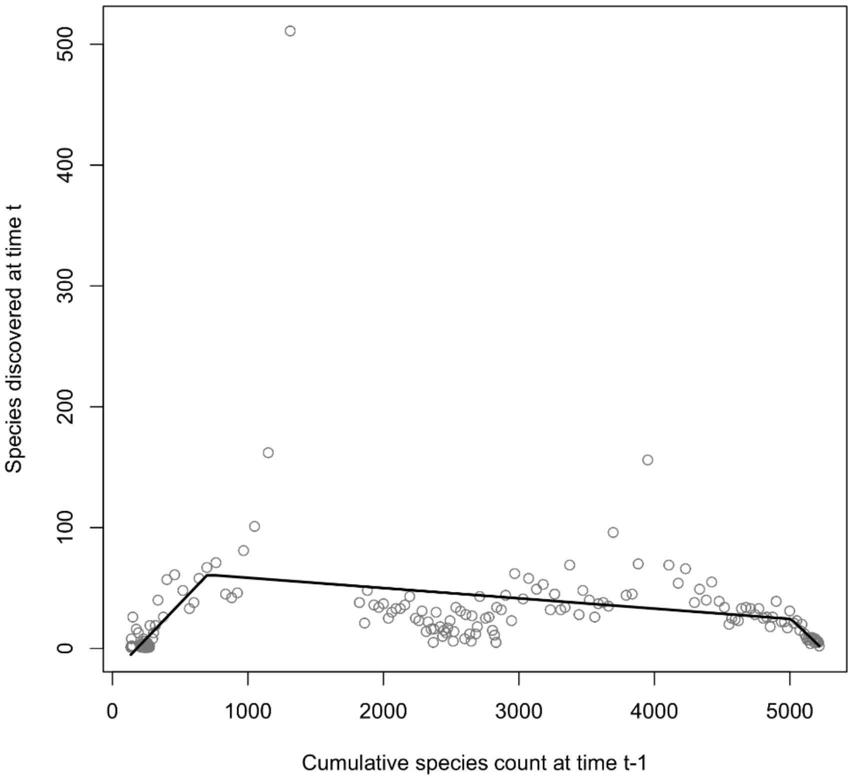


Figure 19 Segmented regression analysis of the slope of the discovery curve (Figure 18) for NSW vascular plants

Slope is broken into segments to isolate the final section of the curve (Figure 19) under the assumption that it is most informative of the asymptote. Variance explained (adjusted r^2) of the complete model is 0.15.

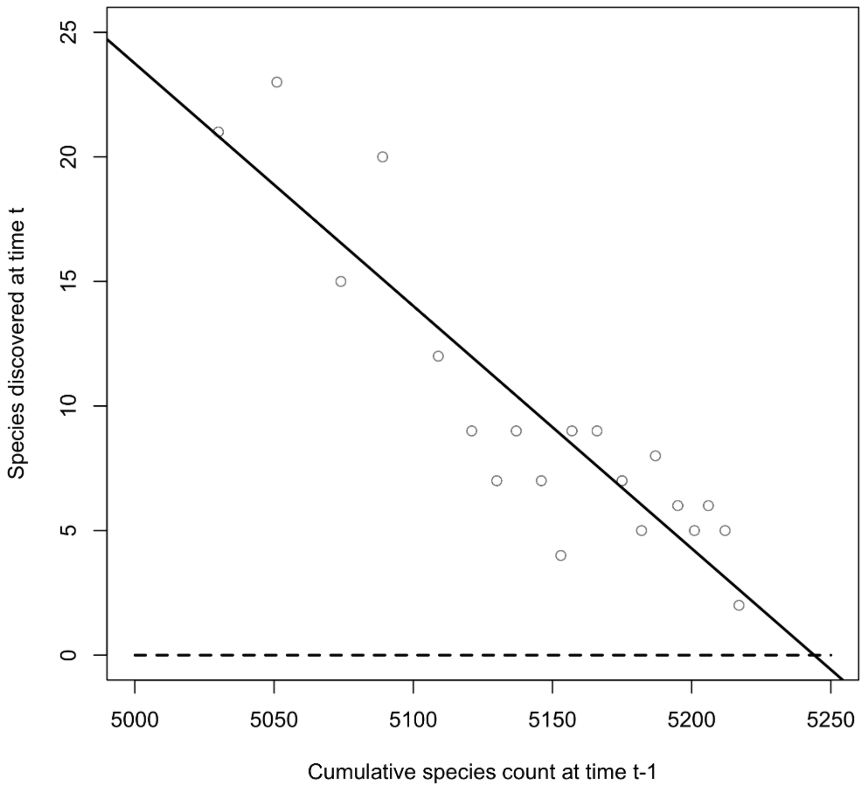


Figure 20 Linear regression analysis of the final segment of the slope of the discovery curve (Figure 14) for NSW vascular plants

The dashed line indicates the intercept of the model and is the estimate of the asymptote of the discovery curve. Variance explained (adjusted r^2) of this segment is 0.80.

The asymptote of the discovery curve for the matched species was estimated to be 5244 species, indicating that the NSW species inventory is 99.7% complete. Assuming that the total PlantNET native list of 6062 species is also 99.7% complete, then the total NSW inventory is 6091 species. Assuming that 72.8% of species (from the representative sets) are expected to survive in 100 years, then the total number of plant species expected to survive in 100 years (including unknown species) is 4427.

4.5 First assessment for indicator: Extant area occupied (for all known species)

The mean proportion of original AOO remaining up to 2013, after down-weighting by ecological condition or ecological carrying capacity, is given in Table 6. Mean proportion was calculated for all NSW plants in the dataset and for each representative species set.

On average, about 56% of NSW plant species' original habitat (i.e. based on occurrence observed since 1950) remains effective when taking ecological condition into account. For the representative species sets, an average of 57% of habitat remains effective. When considering ecological carrying capacity, an average of about 49% of a species' original habitat remains effective, both overall and for the representative species sets.

Table 6 First assessment results (up to 2013) for the indicator: Extant area occupied (for all known species)

Set	Risk category	Mean percentage	
		Ecological condition	Ecological carrying capacity
All NSW plants in dataset	Highest risk	14.6	16.1
	Higher risk	41.6	37.7
	Lower risk	59.4	59.4
	Lowest risk	74.6	73.8
	All categories	56.5	49.3
Representative species sets (mean ± SD)	Highest risk	17.3 ± 1.5	16.8 ± 0.4
	Higher risk	40.8 ± 0.3	37.8 ± 0.2
	Lower risk	59.7 ± 0.1	59.6 ± 0.1
	Lowest risk	75.3 ± 0.1	73.0 ± 0.3
	All categories	56.8 ± 0.3	49.1 ± 0.3

Notes: Indicator is the mean percentage of area of occupancy remaining of all species in the set.
Also given is the breakdown for species in each risk category from a provisional risk assessment (section 4.3).
For the 'Representative species set', results are the mean and standard deviation across nine different sets generated from nine random starts of the species selection algorithm.

4.6 First assessment for indicator: Within-species genetic diversity (for all known species)

The proportion of genetic diversity remaining for each species was calculated (equation 1) from the proportion of AOO remaining up to 2013, after down-weighting by ecological condition or ecological carrying capacity. The frequency distributions for all species and for a selected representative species set are given in Figure 21. Mean values for the first assessment of the indicator are given in Table 7.

Based on the proportion of remaining effective habitat in each species' original range, an average of 79–91% of original within-species genetic diversity is retained, both for all NSW vascular plant species and for the representative species sets. When considering ecological carrying capacity, an average of about 74–89% of original within-species genetic diversity remains.

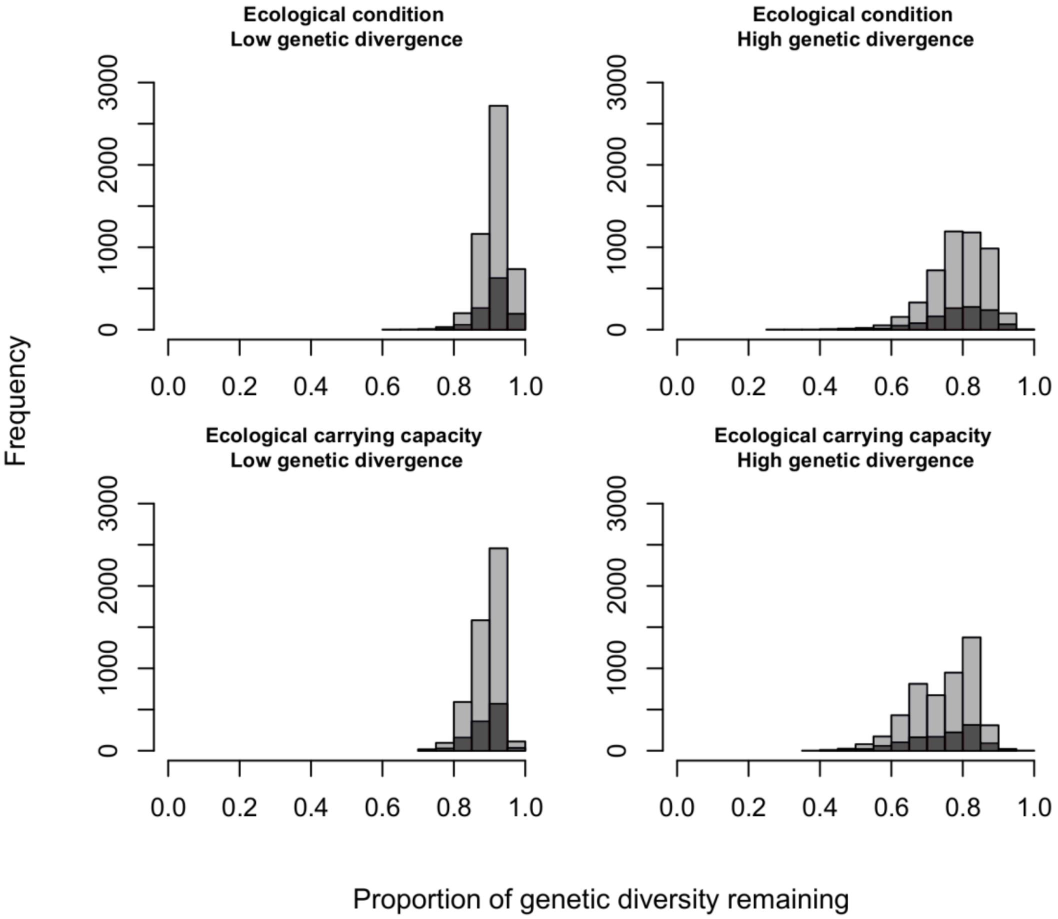


Figure 21 Frequency distribution of the proportion of within-species genetic diversity remaining, inferred from the proportion of area of occupancy remaining after down-weighting by ecological condition or ecological carrying capacity

The full dataset of plant species is shown by the heights of the upper columns and a selected representative set (the second random start) is shown by the lower columns. The graphs on the left assume relatively low genetic divergence between populations (F_{ST} ~ 0.3) while the graphs on the right assume relatively high divergence (F_{ST} ~ 0.9). The mean of the values corresponds to the within-species genetic diversity (for all known species) indicator.

Table 7 First assessment results (up to 2013) for the indicator: Within-species genetic diversity (for all known species)

Habitat quality indicator	Set	Risk category	Percent genetic diversity remaining	
			Low divergence	High divergence
Ecological condition	All NSW plants in dataset	Highest risk	74.5	45.8
		Higher risk	87.5	70.2
		Lower risk	92.4	81.1
		Lowest risk	95.7	88.9
		All categories	91.4	79.0
	Representative species sets (mean ± SD)	Highest risk	76.7 ± 1.1	49.4 ± 1.8
		Higher risk	87.2 ± 0.1	69.5 ± 0.2
		Lower risk	92.5 ± 0.1	81.3 ± 0.1
		Lowest risk	95.8 ± 0.1	89.2 ± 0.1
		All categories	91.5 ± 0.1	79.2 ± 0.2
Ecological carrying capacity	All NSW plants in dataset	Highest risk	75.9	48.0
		Higher risk	86.2	67.4
		Lower risk	92.4	81.1
		Lowest risk	95.5	88.5
		All categories	89.4	74.6
	Representative species sets (mean ± SD)	Highest risk	76.4 ± 0.3	48.8 ± 0.5
		Higher risk	86.2 ± 0.1	67.4 ± 0.1
		Lower risk	92.5 ± 0.1	81.2 ± 0.1
		Lowest risk	95.4 ± 0.1	88.2 ± 0.1
		All categories	89.4 ± 0.1	74.5 ± 0.2

Notes: Indicator is the mean genetic diversity of all species in the set.
Percent genetic diversity remaining given for low and high genetic divergence between populations.
For the ‘Representative species sets’, results are the mean and standard deviation across nine different sets generated from nine random starts of the species selection algorithm.

5. Discussion

5.1 Interpretation

Most vascular plant species in New South Wales have experienced a significant proportion of effective habitat loss within their original range (Figure 9). On average up to 2013, about 56% of NSW plant species' original habitat (i.e. based on occurrence observed since 1950) remains effective based on ecological condition, or about 49% when considering ecological carrying capacity (Table 6). For representative species sets, the results are virtually the same, being about 57% for ecological condition, and about 49% for ecological carrying capacity. This is an outcome of the relatively long history (since European settlement) of land transformation in New South Wales, especially in forested habitats (Bradshaw 2012), and is concerning because habitat loss, alteration and fragmentation are well understood to be dominant factors in predicting extinction risk (Crooks et al. 2017). Note that it is not possible to separate total loss of habitat from habitat quality. The ecological condition and carrying capacity indicators include the extreme case of total loss of effective habitat (an index of 0).

The estimate of 56% effective habitat on average for NSW vascular plants, based on the ecological condition of terrestrial habitats (Table 6), corresponds to an average retention of 79–91% of original within-species genetic diversity (Table 7). This estimate of genetic diversity remaining is very similar to that for the representative species sets (Table 7). Each plant species will respond to loss of range size differently, depending on factors like dispersal capacity and degree of adaptation to local environmental conditions (Alsos et al. 2012). Allowing for a range of values due to genetic divergence does capture this variance, however, future analyses that measure geographic separation and within-range environmental diversity may more accurately model differences in within-species genetic diversity (see section 5.2.2).

The provisional risk assessment suggests that only a relatively small proportion (12%) of NSW vascular plant species are potentially at the lowest level of extinction risk (Table 5). While much caution needs to be applied when translating between the putative risk categories and formal threat categories (see section 5.2.3), this result implies that many more NSW vascular plant species (up to 88%) are threatened than have been assessed and formally listed as such. This proportion represents a significant 'knowledge shortfall' (Hortal et al. 2015) regarding the true risk status of biodiversity in New South Wales.

Based on the provisional risk assessment, about 73% of NSW vascular plant species are expected to be still surviving in 100 years (Table 6). This proportion is the same for the mean of the representative species sets. This precautionary *lower* limit is an indicator, not a prediction. This is because our putative risk categories do not have known extinction probabilities (see section 5.2.3). Such uncertainty does not impact on the information value of the indicator as long as genuine increases in extinction risk lead to lower indicator values (and vice versa), and that species' range size and habitat loss are reliable predictors of extinction risk (Gaston & Fuller 2009; Crooks et al. 2017).

Indicators for all vascular plant species and for the mean of representative species sets are very similar. Despite some variation, expected diversity values for the nine random starts do all fall within the 95% confidence interval for random subsets (Figure 12). This might imply that representative sets drawn from a nearly complete inventory of species within a biological group, such as vascular plants, are no different than random subsets (see section 5.2.5 for discussion).

Extrapolation of the asymptote of the discovery curve for NSW vascular plants indicates that most species have been described and the inventory is very nearly (99.7%) complete (Figures 13 and 15). As a first test of representative species sets, application to a well-known group is useful as it allows comparison between indicators derived from a representative species set to those derived from a large proportion of all species (known and unknown). If indicators derived from representative species sets are representative of broader biodiversity (beyond vascular plants), then estimation of the overall rate of species loss is possible (Mimura et al. 2017).

5.2 Caveats

5.2.1 Uncertainty regarding the true range size of species

The observed range size of a species is a function of the density of occurrence records. When a species is known from only a handful of records, range size (either AOO or EOO) is likely to be under-estimated (Gaston & Fuller 2009). If only the confirmed range is used (from reliable occurrence records), a poorly known species is more likely to be considered threatened due to having a small observed range size (Mace et al. 2008). An overly stringent data cleaning process would exacerbate such problems. While species distribution models (SDMs) could potentially be used to fill the gaps in range size that are attributable to under-sampling, the resulting area is likely to be over-estimated. This is because SDMs predict suitable habitat rather than occupancy per se (Elith & Leathwick 2009). Not all suitable habitat is necessarily occupied due to a range of factors including limits to dispersal and species interactions (Nipperess & Beattie 2004).

Our analyses used all reliable occurrence records held in natural history collections since 1950 as the basis of our calculations of ‘original’ range area size. Using a relatively large temporal range helps ensure enough spatially reliable records for a reasonable estimate of range area but necessarily obscures any temporal patterns. This study opted to treat a species’ range as effectively static to provide a measure of its ‘original range extent’, but allowed that observed area to be modified by a general model of ecological condition and carrying capacity designed to be broadly applicable to terrestrial biodiversity (Love et al. 2020). Because that model uses remote-sensing data, there is potential to track changes in the ecological condition and carrying capacity of the area potentially occupied by a species at a fine temporal resolution.

This study has assumed a very simple model of the relationship between ecological condition or carrying capacity and the effective area occupied by a species. The AOO was multiplied by the mean index (scaled 0–1) within the AOO and the resulting area interpreted as the proportion of AOO that is intact (i.e. with a condition score of 1). In effect, we assume that the area of suitable and occupied habitat within a species’ original (since 1950) AOO is a linear function (where slope = 1) of the proportion of the AOO still intact. While it is certainly safe to assume a monotonic relationship between habitat alteration and effective habitat area, species are known to differ markedly in the form of that relationship. For example, some bird species are known to exhibit a threshold effect in response to habitat alteration, only occurring above a certain threshold of tree density (Radford et al. 2005). Our general model (linear response to ecological condition) may not necessarily be the most appropriate for a given species. While it is not feasible to fit separate models for each species, empirical studies of samples of species within taxonomic or functional groups could provide guidance on whether a linear relationship is a reasonable approximation for a specified grouping.

5.2.2 Uncertainty regarding the relationship between range size and within-species genetic diversity

The estimate of remaining genetic diversity is based on AOO, for which suitable scaling parameters are known (Mimura et al. 2017), but the geographic dispersion of occupied habitat might be a better predictor. It is well understood that there is a relationship between the geographic separation of populations and their genetic divergence (Slatkin 1993) and thus EOO, which does capture information on geographic separation, might be a more appropriate surrogate for genetic diversity. However, the extent to which geographic separation leads to greater genetic diversity is largely determined by the dispersal capacity of the species in question (Alsos et al. 2012). Differing dispersal capacities may well be adequately modelled by allowing for a range of possible values for genetic divergence because, all else being equal, species with poor dispersal abilities can be expected to have higher genetic divergence among populations.

Occupancy of environmental space might be a better predictor of within-species genetic diversity than geographic range size. In addition to genetic isolation by distance, populations diverge due to local adaptation to differing environmental conditions in different parts of a species' geographic range (Leimu & Fischer 2008). It follows that, all else being equal, species with ranges that cover a larger diversity of environmental conditions would have more genetic diversity due to local adaptation (Mimura et al. 2017). By quantifying the environmental coverage of a species' geographic range, by applying the environmental diversity framework (Faith & Walker 1996) for example, it would be possible to develop suitable scaling parameters that convert environmental coverage to genetic diversity. Further work is needed to adapt the method and data to this end.

5.2.3 Uncertainty regarding risk assessment and the true probability of survival

If one were to translate the provisional risk categories to corresponding threatened species categories (i.e. vulnerable, endangered and critically endangered), our analysis suggests that the majority (~88%) of NSW vascular plants would be classified as threatened if a comprehensive formal risk assessment of all known species were to be undertaken. When compared to global comprehensive assessments by the IUCN (e.g. ~41% of amphibian species are threatened; Hoffman et al. 2010), this figure seems unusually high. Conversely, of those species used in this analysis that were also listed under the BC Act, 24% were assigned to the lowest risk category by the provisional risk assessment (Table 4) when they should plausibly be assigned a higher extinction risk.

Because of the need to include as many species as possible, the provisional risk assessment is, in at least one sense, less conservative in assigning threat categories to species than a formal process. Two criteria are used: AOO (adjusted by ecological condition) and the proportional loss of ecological condition of AOO. While our assessment is fundamentally based on criterion B2 of the *IUCN Red List Categories*, the IUCN categorisation requires small range size and at least two of three additional subcriteria:

1. severe fragmentation or restricted number of locations
2. continuing decline in abundance, range area, habitat quality or number of locations
3. extreme fluctuations in range area, abundance or number of locations.

The provisional risk assessment is based on only one of the three subcriteria, that is, decline in range area (subcriterion 2 above). The risk assessment was, however, conservative about what level of decline was considered necessary to assign a risk category. Because all species have some loss of ecological condition across their range since the pre-industrial era, a literal adoption of the *IUCN Red List Guidelines* would have resulted in all species in the analysis being considered threatened (see Figure 16). For this reason, additional critical thresholds were used by adapting criterion A of the *Red List of Ecosystems* (Bland et al. 2017). Future assessments may be able to add an additional dimension to the provisional risk assessment via the ecological carrying capacity of terrestrial habitat indicator. This dimension would correspond to the IUCN B2 subcriterion of ‘severe fragmentation’ although critical thresholds would need to be determined.

Our analysis assumes that reduction in range size is the key indicator of extinction risk for species. While there is a well-known relationship between range size (as AOO) and abundance (Gaston & Fuller 2009), the density of individuals per unit area varies across species and also within a single species’ range (due to spatial variation in habitat suitability, for example). Further, a species could potentially decline in abundance without a consequent decline in range area if the density of individuals were to become reduced. Importantly, habitat loss, alteration and fragmentation while significant, are not the only processes threatening species with extinction. A species could decline in abundance, due to threatening processes like introduced species, disease or climate change, for example, without a change in ecological condition or ecological carrying capacity across its range.

As the survival (or otherwise) of a species is a future event, the true probability cannot be known. Population viability analysis provides a proven means for exploring the range of possible population trajectories (Brook et al. 2000) but can only be done on the relatively small number of species for which there are sufficient demographic data. However, IUCN Red List categories can be assigned explicit probabilities via criterion E (Kindvall & Gärdenfors 2003) and are based on a rigorous evidence-based assessment process (Mace et al. 2008). Although the risk categories used for this indicator are modelled on IUCN categories, they are not equivalent to the detailed risk assessments undertaken by a panel of scientific experts such as the NSW Threatened Species Scientific Committee. While the putative risk categories have been assigned probabilities based on their corresponding IUCN threat category, this is likely to be an under-estimate of the true survival probability due to the less conservative and incomplete risk assessment process. Despite uncertainty about true survival probability, the derived risk categories should be viewed as reliable indicators of the *relative* extinction risk of species. Thus, while the true amount of expected diversity may be unknown, changes in the indicator attributed to change in habitat loss, alteration and fragmentation will reliably track changes in relative extinction risk and thus, changes in the rate of biodiversity loss.

5.2.4 Uncertainty regarding the true number of species in New South Wales

Inevitably, the more complete an inventory, the more confident we can be about the completeness of the inventory. Extrapolating from a limited sample necessarily results in a corresponding high degree of uncertainty (Bebber et al. 2007). In the case of vascular plants, the discovery curve is very close to an asymptote (Figure 18 and Figure 20), indicating a near-complete inventory for New South Wales and therefore, high confidence in the estimate. Some caution must be taken because this fit was based on data from a subset of known NSW vascular plant species. Assuming no bias in the subset, the estimate of completeness should be robust. A key assumption of the analysis is that the current decline in the rate of discovery will persist and thus the asymptote will soon be reached. The historical discovery curve (Figure 18) shows that the rate of discovery has varied markedly over the past two centuries and could, conceivably, rise again from its current pattern of decline, perhaps because of the discovery through genetic studies of many cryptic species.

5.2.5 Uncertainty regarding the true representativeness of species sets

If the purpose of an indicator is to assess the rate of loss of biodiversity, then the pattern discerned should be *representative* of all species in the group in question. Given the robust assumption that species are distributed along environmental gradients (Whittaker 1972), it follows that the diversity of environments is a good surrogate of biodiversity (Faith et al. 1987; Faith & Walker 1996). A set of species specifically selected to represent environmental diversity in their habitats should therefore be representative of all the species within their group (Mimura et al. 2017). Given their trophic position at the base of NSW terrestrial ecosystems, vascular plants in particular should be a good basis for a representative species set that may extend more generally as a surrogate for other biological groups that depend on them. Spatial patterns in plant species composition, for example, have been shown to be a useful surrogate for similar patterns in other biological groups (Su et al. 2004; Ware et al. 2018).

For all indicators, the mean of the representative species sets is very similar to the result for all plant species in the dataset. The larger the size of the set, the more likely this will be true. Choosing an optimal size for the representative species set involved balancing the competing requirements of minimising environmental redundancy and avoiding bias towards small-ranged species (which are more closely associated with distinct areas in environmental space). The chosen threshold of 25% of the 4595 species considered seems to balance these requirements adequately. The results are very similar to that for the full dataset which suggest the set is truly representative. Even though random subsets of the same size performed similarly, the fact that the species were specifically chosen to be environmentally representative increases our confidence in the surrogate value of the set. Further, given that, by our estimate, vascular plant diversity is well known for New South Wales, the similarity of results between representative sets and the full dataset provides support for the capacity of representative species sets to represent the full diversity of a group, and supports the application of the method to less well-known groups.

5.2.6 Sensitivity to change over five years

Sensitivity to change of the biodiversity indicators reported here relies on the sensitivity of the ecological condition and ecological carrying capacity indicators, because all changes in range size are inferred from changes in habitat quality (loss, alteration and fragmentation). Structural changes in habitat quality are detectable at fine-spatial resolutions from remote-sensing data, although processing and public release of these data is currently lagged (2013 reported here). Nevertheless, even if reporting is offset by up to five years, sensitivity to change on a more regular basis is likely to be high in semi-natural areas undergoing significant alteration of habitat for which the implications in terms of rates of potential biodiversity loss can be reported at a relatively fine temporal scale.

These biodiversity indicators are not sensitive to apparent changes in range size due to temporal variation in the spatial pattern of occurrence records. This is because all available records (after cleaning) since 1950 are used to determine range size. Occurrence records are simply too sparse for many species to allow the reliable tracking of changes in range size over time. It is certainly possible that a less stringent cleaning process or new records could change the known original range size of a species; and this information would be used to adjust any previous assessments of the indicator as well as the current reporting period. Trends in effective range size would be inferred solely from changes in habitat quality.

A more direct measure of species-level trends is provided by population monitoring such as reported through the site monitoring of species and ecosystems indicator (OEH & CSIRO 2019). Such an approach has the advantage of providing direct evidence of change, and, given frequent monitoring, can be reported on a regular (e.g. annual) basis. Population monitoring data can also be combined in a composite index such as the Living Planet index (Collen et al. 2009) or the National Environmental Science Program threatened species index (Threatened Species Recovery Hub, undated). However, the values of these indices are entirely dependent on the sample of species being used as input, which is itself determined by the availability and reliability of time series of abundance data. Such samples are taxonomically biased and not necessarily representative of broader biodiversity trends. There is, therefore, a need for complementary sets of indices that provide both direct and indirect measures of the rate of biodiversity loss.

5.3 Data gaps and considerations

5.3.1 Lists of NSW species for taxonomic groups and their earliest known record

The National Herbarium of New South Wales maintains an online flora database, PlantNET (Royal Botanic Gardens and Domain Trust 2018) that provides a list of accepted names for NSW plants. Discounting introduced species, PlantNET lists 6062 species occurring naturally in New South Wales. While this is a useful resource for the state, taxonomy is not entirely concordant with the national Australian Plant Census (Council of Heads of Australasian Herbaria 2018). Of the 6062 species in the PlantNET native species list, 5304 have a matching name in the Australian Plant Census. When retrieving earliest NSW records from the Atlas of Living Australia (which uses the Australian Plant Census taxonomy), 5219 species from the PlantNET list could be matched to a reliable NSW record. Going forward, it would be useful to either maintain a list of NSW plant species, either standardised or mapped to the Australian Plant Census.

5.3.2 Parameters for estimating within-species genetic diversity for NSW plants

The scaling parameters used in the within-species genetic diversity indicator were taken from an analysis by Mimura et al. (2017) that used data on northern hemisphere plants from Europe and North America (Alsos et al. 2012). We are currently lacking information on the specific range of responses (modelled by the z parameter) to habitat loss for NSW vascular plants and, potentially, for other biological groups. Acquiring this information would require a similar analysis to Alsos et al. (2012) for populations of a set of species across New South Wales where the relationship to habitat loss is estimated by randomly removing populations and recalculating genetic diversity. Following Mimura et al. (2017), a power curve is then fitted to the relationship for each species and the z parameter estimated. Genetic divergence (as the fixation index, F_{ST}) would also be calculated for each species to determine the appropriate range of z values for NSW vascular plants. Such a study would also provide a means to test the assumptions of the model. There is also the opportunity to develop appropriate models for EOO and environmental coverage to provide a measure of adaptive genetic diversity.

6. Guidelines for running the next assessment

The data and scripts provided in the data package (Nipperess et al. 2020) for these indicators will allow the first analyses to be re-run. The indicators and data are designed such that future assessments can be run on an annual basis if desired. Updating the underlying species occurrence data could be the basis for generating new representative species sets but the current sets could be used for the near future. Changing the sets would necessarily require the recalculation of previous assessments of the indicators. Calculation of range size for each species will need to be repeated when there are significant updates to the underlying occurrence records, the GDM-based environmental space model or the condition assessment indicators. Provided below are guidelines on how to re-run the analyses and to adapt the data package for future assessments of the indicators.

6.1 Software requirements

Generation of the representative species set requires custom software: the ‘representative species set selector’ tool (Windows executable, Manion 2018). Future assessments of the indicators will use the same representative species sets and so this tool will only be needed if it is decided to generate new sets. This would be the case if there were substantial gains in new knowledge about species distributions making it worthwhile to re-run the selection a few years in the future.

Grid mapping of occurrence records and calculation of range size was achieved using tools available in ArcGIS^{TM} 10.5 (ESRI 2016). See section 3.3 for more details.

For each indicator, a custom script was written in *R: A language and environment for statistical computing* (R Core Team 2017, referred to in this report as ‘*R*’). Scripts are provided as part of the data package and require the base installation of *R*. The script for the generation of the discovery curve (part of the script for the expected survival of all known species indicator) also uses the ‘segmented’ library (Muggeo 2008) to estimate the asymptote of the discovery curve.

6.2 Workflow

This section refers to the workflow diagrams for each indicator and the representative species sets calculation (included as appendices to the report; see grey ‘Process’ boxes). Object identifiers listed below relate directly to parts of the workflow.

6.2.1 Clean and harmonise species data (workflow object P00001)

This process was implemented by an *R* script written by Stuart Allen and is not part of the data package. The cleaned data is provided as part of the data package. As the representative species sets are expected to be retained over multiple assessments of the indicators, this process should not need to be repeated unless it is decided to generate a new representative species set, or substantial gains in new knowledge about species distributions make it worthwhile to re-run the selection in the future, or a different suite of decisions are applicable to the process of cleaning and filtering the data. The process could be implemented by following the description of the species data sourcing, cleaning and manipulation given in section 3.1.

6.2.2 Attribute species records with GDM-scaled environmental variables (workflow object P00002)

This process was implemented using the *.NET Survey Gap Analysis Tool* (Manion 2018), and specifically the ‘build full predictor table from (“spp_code, x, y”)’ function from the drop-down tools menu. This step also requires a GDM for the biological group being assessed. This process should not need to be repeated unless it is decided to generate a new or refined model for the biological group.

6.2.3 GDM sample selection (workflow object P00003)

This process was implemented using the *.NET Survey Gap Analysis Tool* (Manion 2018), and specifically the ‘create uniformly distributed sample point file (_id, X, Y, …)’ function from the drop-down tools menu. This is a step towards demand point generation and would only need to be repeated if the grid extent or size changed or the sample size changed (e.g. increased).

6.2.4 Demand point generation (workflow object P00004)

This process was implemented using the *.NET Survey Gap Analysis Tool* (Manion 2018), and specifically the ‘autoselect user defined number of demand points from a sample table (_id, X, Y, …)’ function from the drop-down tools menu. This step also requires a GDM for the biological group being assessed. This process should not need to be repeated unless it is decided to generate a new or refined model for the biological group.

6.2.5 Representative species selector (workflow object P00005)

This process was implemented using the *.NET Survey Gap Analysis Tool* (Manion 2018), and specifically the new ‘derive a representative sample of species from demand point set and site data via ED’ function from the drop-down tools menu. This step uses the demand point (workflow object P00004) and the species records (workflow object P00002) both attributed with the GDM-scaled environmental variable values, and would need to be repeated if any precursor inputs have changed or been updated.

6.2.6 Test demand point sample sizes for optimal species representativeness (P00006)

This process was implemented by an *R* script (included in the data package) entitled ‘representative species sets’. This script generates diagnostics to determine an appropriate size for the representative species set. As the representative species set is expected to be retained over multiple assessments of the indicators, this process should not need to be repeated unless it is decided to generate a new representative species set. Required inputs are:

1. A set of files giving the identifiers for the species selected for representative species sets for every size of set within a range (all set sizes from 1 to 4595 species were tested) and for every random start (nine random starts were tested).
2. The value of the diagnostic (mean distance of unselected species to their nearest demand points) for every size of set and every random start.
3. Range sizes (as AOO and EOO) for every species.

Inputs 1 and 2 are generated by the representative species selector tool (Manion 2018).

6.2.7 Adapted IUCN process for grid mapping (2 km) of habitat extent (workflow object P00007)

This process is implemented in ArcGIS™ 10.5 (ESRI 2016) (see section 3.3). Required data inputs are a cleaned set of occurrence records (DD00001) and a GIS mask of New South Wales.

6.2.8 Species' AOO calculator (workflow object P04501)

This process is implemented in ArcGIS™ 10.5 (ESRI 2016) (see section 3.3). Required data inputs are the output of workflow process P00007 (see above) and the condition model.

6.2.9 Species assignment to extinction risk categorisation (workflow object P04502)

This process is implemented as part of an *R* script for the expected survival of all known species indicator. Required input is a spreadsheet of range sizes (AOO) adjusted and unadjusted by the condition model (as a .csv file) generated as output from the species area of occupancy calculator (P04501).

6.2.10 Expected diversity calculator (workflow object P04503)

This process is implemented as part of an *R* script for the expected survival of all known species indicator. It follows directly from the risk categorisation process (P04502) and requires no external inputs.

6.2.11 Summary analysis of fractional AOO (workflow object P06001) and genetic diversity calculator (workflow object P05501

This process is implemented as part of a combined *R* script for the within-species genetic diversity (for all known species) indicator, and the extant area occupied (for all known species) indicator. Required input is a spreadsheet of range sizes (AOO) adjusted and unadjusted by the condition model (as a .csv file) generated as output from the species area of occupancy calculator (P04501).

6.3 Report card

The report card provides a single result for the representative species sets for the assessment up to 2013 for each indicator. Assessment results are given as a percentage as either a mean across representative sets (expected survival indicator) or as a mean across species, averaged across representative sets (within-species genetic diversity indicator and extant area occupied indicator). Within-species genetic diversity reports two amounts: one assuming low within-species genetic divergence and the other assuming high divergence. Indicators are reported only for the representative sets because the indicators are intended for all species within the biological group, including unknown species. Expected survival is given as a percentage rather than number of species because the number of species is simply a function of the size of the representative species set and thus less informative than a percentage. Also, because the indicator is intended to be representative of the biological group, percentages (or proportions) are more appropriate to be reported.

6.4 Triggers for running next assessments

New assessments of the indicator can be re-run any time new data become available. The most significant trigger would be an update to the ecological condition of terrestrial habitat indicator. Condition-adjusted range sizes of the species in the representative species sets could then be updated and the indicators calculated for a new time step. Range size calculations could use updated species occurrence data, which would necessitate a recalculation of previous assessment results and any subsequent assessments, in light of an improved understanding of original range size of the representative species. The adoption of a revised GDM model would necessitate the derivation of new sets of representative species. Future assessments of the indicators can continue to use the representative species sets from this first assessment until there is a significant update to the underlying GDM used to select them. Generation of new representative species sets would require recalculation of all previous assessments.

6.5 Potential methodological changes to the next assessment

A critical element in the methodology of the indicators reported here is the calculation of range area. Improving the method of calculation will decrease uncertainty and improve confidence. As discussed in section 5.2.1, occurrence records are likely to under-estimate range size (as either AOO or EOO).

Species distribution models (SDMs) are often used to infer the occupied range of a species by mapping suitable habitat (based on climatic and other factors). While it would be a large task to model all species in a biological group, modelling a subset (such as a representative species set) may be tractable. Once a SDM is generated for a species, producing a map of habitat suitability per grid cell (usually scaled 0–1), range size can be estimated by converting to a binary map via selection of a threshold (Elith & Leathwick 2009). However, as previously discussed (section 5.2.1), SDMs will tend to over-estimate range size because not all suitable habitat (predicted from a SDM) will be occupied. Further, the relationship between probability of occurrence and habitat suitability will vary between species, complicating the choice of an appropriate threshold. A SDM could be adjusted by down-weighting those grid cells that are less likely to be occupied based on spatial (and possibly temporal) distance from a known occurrence record (Bush et al. 2014). This would potentially allow for a more accurate estimation of range size while avoiding over-estimation due to suitable but unoccupied habitat.

Given that the observed range size is a function of sampling effort, modelling the relationship between number of occurrence records and observed range size (as either AOO or EOO) provides a means of assessing confidence in estimates of range size. By randomly subsampling the occurrence records for a species, a rarefaction curve could be generated for each species, showing how the accumulation of occurrence records increases the estimate of range size. As a rarefaction curve approaches an asymptote, we can be increasingly confident that the true range size has been determined. For AOO, the exact analytical solution for rarefaction, which is much more efficient than repeated resampling, is identical to that for species richness (Hurlbert 1971). Further, Chao (1984) showed that the asymptote (the true range size) could be estimated from the slope of the rarefaction curve and therefore determine how completely the true range size of a species has been sampled by known occurrence records. Thus, in the case of AOO, it is possible to predict by rarefaction how many 2 kilometre x 2 kilometre grid cells are occupied and not sampled by occurrence records, but not where these grid cells are (for which an SDM would be needed). Nevertheless, rarefaction would be a useful technique for assessing confidence in range size estimates and could be used to determine which species have sufficient data available for inclusion in representative species sets.

Related to the issue of sampling effort is the data cleaning process adopted by this study, which is inevitably a compromise between ensuring data quality versus sampling sufficiency. If too many occurrence records are discarded, range size will be under-estimated. In future assessments, data cleaning could be relaxed and informed by a per-species assessment of the completeness of range size estimates (by rarefaction analysis).

As discussed in section 5.2.1, a simple linear relationship was assumed between ecological condition and carrying capacity and area of effective habitat. Because fitting separate response to disturbance models for each species would not be feasible, a general habitat condition model for all biological groups has been adopted. Empirical data for selected representative species could suggest a different form of this relationship that might be broadly appropriate for a particular biological group, such as vascular plants. Importantly, an empirical study should include a test for the relationship of occupancy with ecological condition and ecological carrying capacity indices (as used in the respective indicators) in addition to individual habitat attributes (such as tree cover). In the absence of such studies, a linear response model is arguably the most generally applicable.

AOO was used instead of EOO in the indicators reported here because of the relatively simple translation between grids of ecological condition and carrying capacity and grids of effective habitat. However, EOO does capture information on spatial dispersal that is not captured by AOO, which might be especially useful for estimation of genetic diversity. Alternatives to all indicators using EOO, for comparison with the results derived using AOO, could be adopted in a future assessment and would require an additional spatial process for adjusting EOO by intersecting ecological condition or carrying capacity. Moreover, in the case of genetic diversity, parameters for translating proportional habitat area to proportional genetic diversity have not yet been specifically developed for EOO.

The analysis of within-species genetic diversity applied here using the AOO is mainly relevant to the neutral component of genetic diversity. However, two types of genetic diversity are commonly considered in ecological studies: neutral and adaptive. Whitlock (2014) provides a summary of the mechanisms connecting genetic diversity and ecological structure. Neutral genetic diversity arises through location-specific processes related to genetic drift and migration, and are selectively neutral; whereas adaptive variants of genetic diversity influence the phenotype and fitness of the organisms that carry them. Future applications of within-species genetic diversity should then appropriately also consider adaptive variants by using the AOO to also measure the environmental diversity across the range of a species and develop a general proxy measure for adaptive diversity.

The risk assessment approach could be modified in a future assessment of the indicators. This first assessment may be too conservative because small-ranged species will not be assigned to higher risk categories unless they have also experienced significant loss of effective habitat. Given a sufficiently small range size, a species can be considered at risk even in the absence of habitat loss (Mace et al. 2008). Such species can be accounted for by adopting criterion D2 of the vulnerable category of the *IUCN Red List Categories*, which allows for species to be listed as vulnerable if they have a very restricted AOO and do not otherwise qualify for other categories (IUCN 2012). This would equate to species being assigned to the lower risk category if adjusted AOO is less than 20 km^2 and proportional loss of condition is less than 30%. As mentioned earlier (section 5.2.3), an additional subcriterion based on ‘severe fragmentation’ could be added if suitable thresholds could be determined.

Range area as used for the risk assessment in this report was range in New South Wales only. This is because we were interested in extinction risk specifically in New South Wales which matches the current approach to listing threatened species in New South Wales. Under the *Common Assessment Method* (Department of the Environment and Energy Australia 2018), national, state and territory governments have agreed to a common approach to assessing and listing threatened species based on the *IUCN Red List Categories* (IUNC 2012), where species will be assessed as threatened at a national scale.

Only those species not listed as threatened nationally could be separately listed as threatened within a specific jurisdiction (i.e. state or territory). This procedural change could affect future assessments of this indicator if it was decided to follow the *Common Assessment Method*. For example, nationally listed NSW species also found in other jurisdictions could be assessed as having larger range areas and possibly assigned a lower risk of extinction. However, by taking this approach, estimated biodiversity loss would not necessarily be specifically for New South Wales, that is, extinction probabilities would be loss nationally rather than loss from New South Wales (and possible survival elsewhere).

Finally, a much broader risk assessment strategy could be adopted in a future assessment that attempts to draw on multiple lines of evidence to assign a probability of extinction. Evidence could include population monitoring data, formal listing (at a global, national, or state level) and inferred habitat loss (as done here). Such an approach would need to be based on a clearly defined decision framework (modelled on *the IUCN Red List of Threatened Species*) and at least partially automated using a script, ensuring feasibility, transparency and repeatability.

7. Data products

The data used and derived as a product of this analysis will be publicly available through the CSIRO Data Access Portal (data.csiro.au). The following data package is available for download:

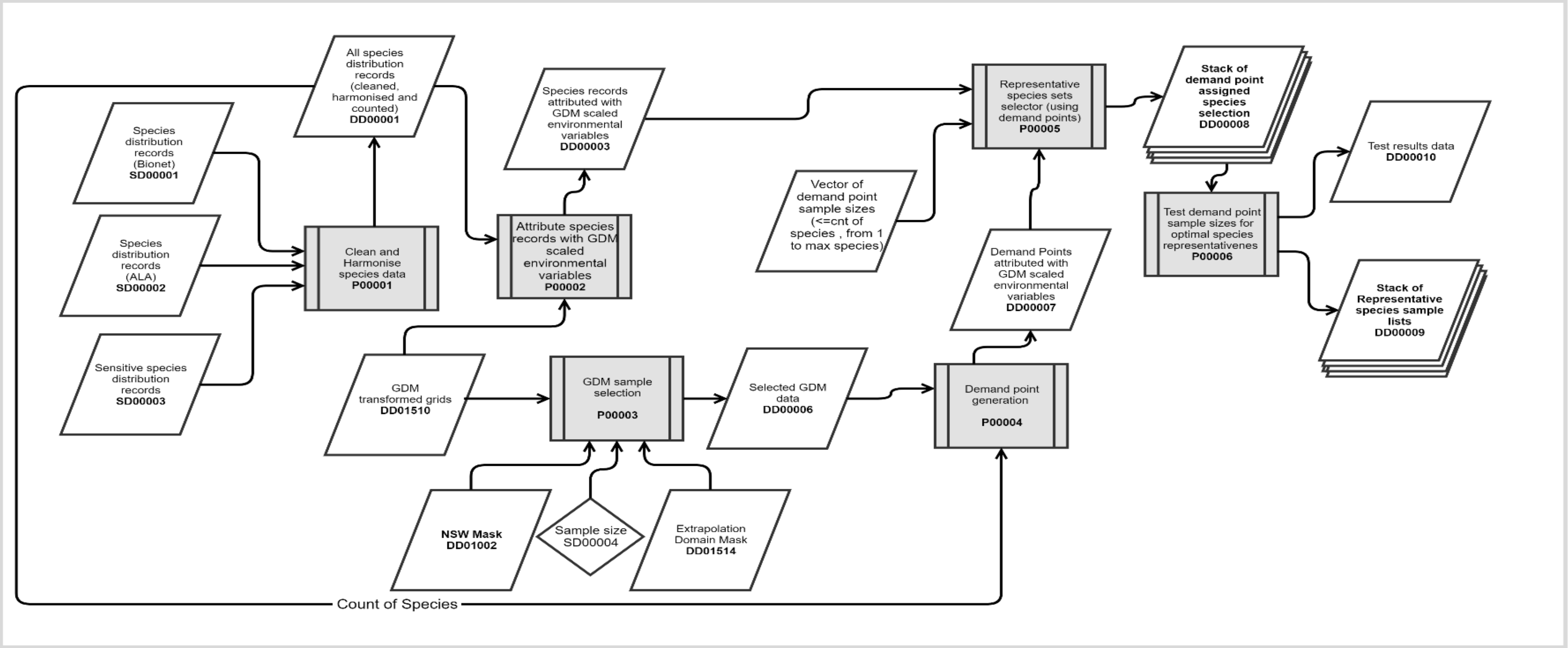
> Nipperess DA, Faith DP, Williams KJ, King D, Manion G, Ware C, Schmidt R, Love J, Drielsma M, Allen S & Gallagher R 2020, Expected survival and state of all known species: Data packages for the Biodiversity Indicator Program, first assessment, *SEED Portal*, datasets.seed.nsw.gov.au/dataset/biodiversity-indicator-program-data-packages.

The data package is part of a collection hosted on the Sharing and Enabling Environmental Data (SEED) Portal (seed.nsw.gov.au). The collection includes links to all available data packages for the first assessment of the Biodiversity Indicator Program:

> Department of Planning, Industry and Environment 2020, Data packages for the Biodiversity Indicator Program: First assessment, SEED Portal, Sydney, Australia, datasets.seed.nsw.gov.au/dataset/biodiversity-indicator-program-data-packages.

Appendices

Workflow representative species selection



Legend

Shapes

▥ **Process** (with unique identifier starting with 'P')

▭ **Document** (with unique identifier starting with 'RC' for report cards)

▱ **Dataset** (with unique identifier starting with 'DD' for derived dataset and 'SD' for source datasets)

◇ **Decision variable** (treated like a source data set but separately defined in the workflow)

Colours

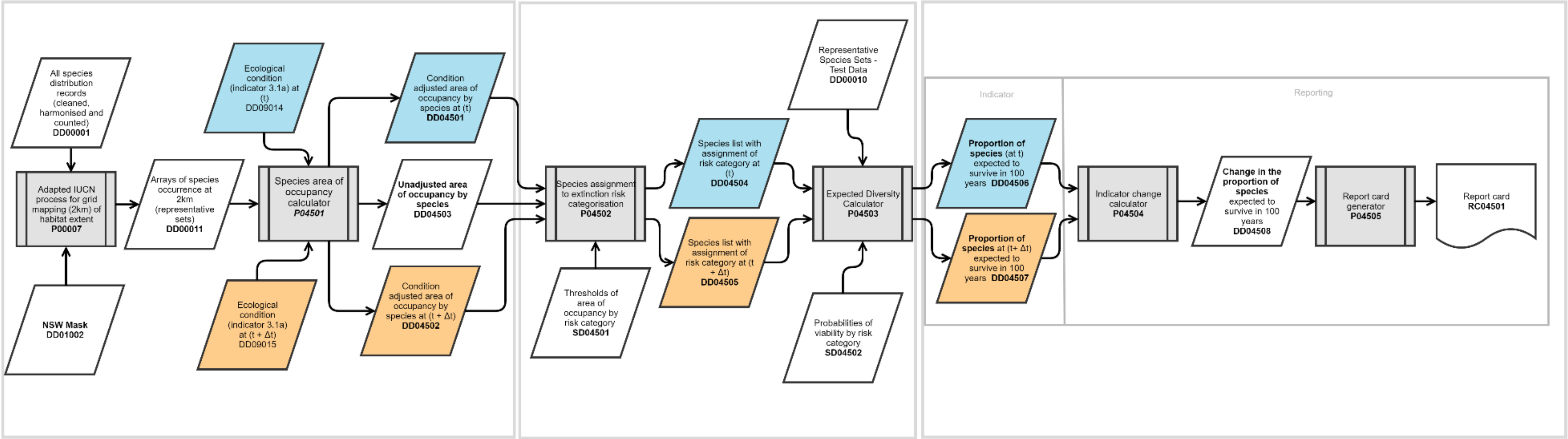
🟦 **Time t**

🟧 **user defined number of years after time t**

Object id range

00001-00500

Workflow for indicator: Expected survival of all known species



Legend

Shapes

▥ **Process** (with unique identifier starting with 'P')

▱ **Dataset** (with unique identifier starting with 'DD' for derived dataset and 'SD' for source datasets)

▭ **Document** (with unique identifier starting with 'RC' for report cards)

◇ **Decision variable** (treated like a source data set but separately defined in the workflow)

Colours

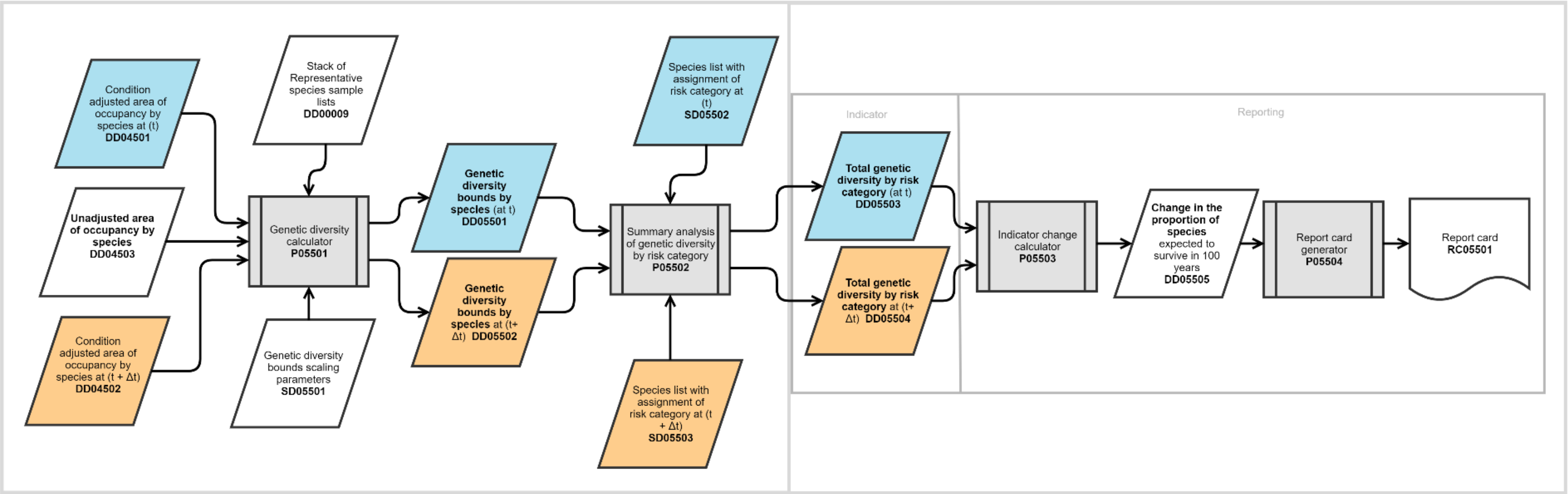
🟦 **Time t**

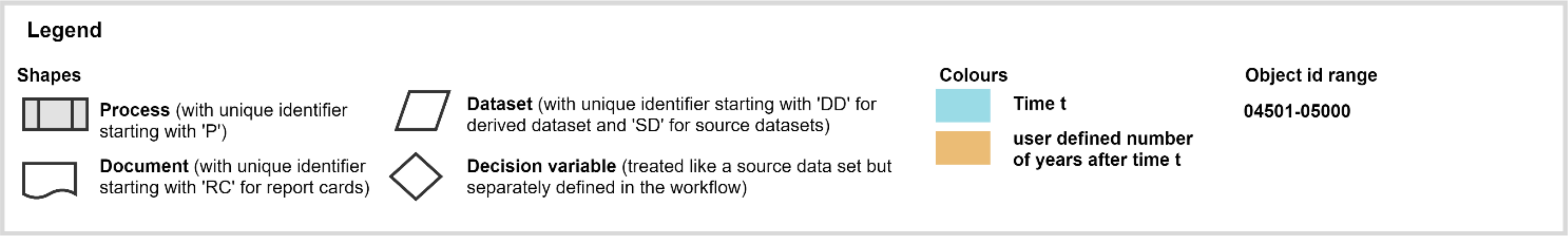
🟧 **user defined number of years after time t**

Object id range

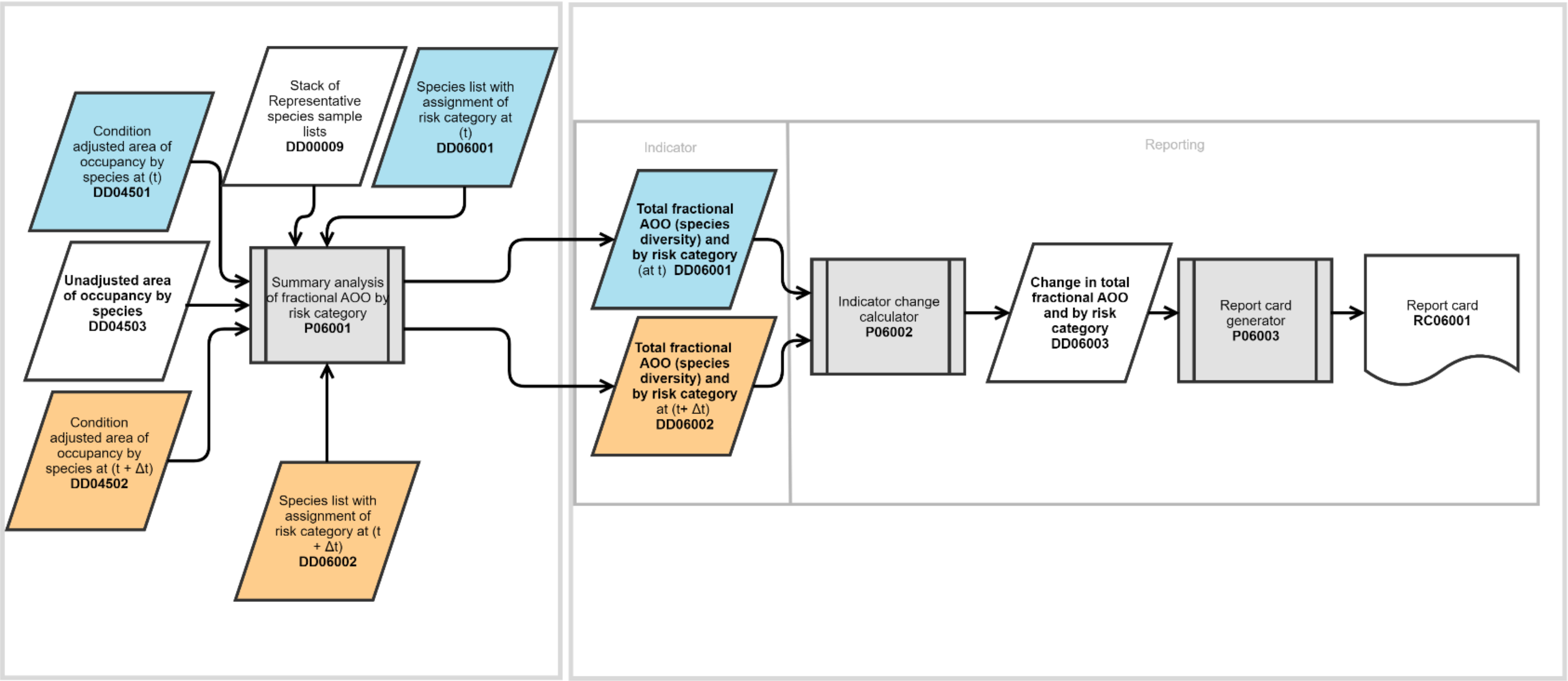
04501-05000

Workflow for indicator: Within-species genetic diversity (for all known species)





Workflow for indicator: Extant area occupied (for all known species) indicator workflow



Legend

Shapes

Process (with unique identifier starting with 'P')

Dataset (with unique identifier starting with 'DD' for derived dataset and 'SD' for source datasets)

Document (with unique identifier starting with 'RC' for report cards)

◇ **Decision variable** (treated like a source data set but separately defined in the workflow)

Colours

🟦 **Time t**

🟧 **user defined number of years after time t**

Object id range

06001-06500

Glossary

Area of occupancy (AOO): as defined by the NSW Scientific Committee, is the area within the total range (and hence within the extent of occurrence, or EOO) that is currently occupied by the species. It excludes unsuitable and unoccupied habitat. In some cases, (e.g. irreplaceable colonial nesting sites, crucial feeding sites for migratory taxa) the area of occupancy is the smallest area essential at any stage to the survival of existing populations of a taxon.

Assessment: using biophysical data collected through monitoring, combined with other inputs such as benchmarks, to make judgements about environmental condition and trends.

Biodiversity (biological diversity): variability among living organisms from all sources (including terrestrial, freshwater, coastal, marine and other ecosystems and ecological complexes of which they are part), which includes genetic diversity, species diversity and ecosystem diversity.

Bioregion: relatively large land area characterised by broad, landscape-scale natural features and environmental processes that influence the functions of entire ecosystems and capture large-scale biophysical patterns. These patterns in the landscape are linked to fauna and flora assemblages and processes at the ecosystem scale. There are 18 bioregions represented in New South Wales.

Complementarity: is the marginal gain in biodiversity provided by a location, place or region relative to other locations, places or regions. For example, the complementarity value of an area (for example, National Park A) is given by the number of so-far-unrepresented features of biodiversity that it contributes, relative to other areas (for example, National Parks B to Z). The principle of (biodiversity) complementarity, applied to conservation planning, ensures that places chosen for inclusion in a network of conservation areas complement those already selected in order to increase the overall representation of different features of biodiversity within the network.

Confidence interval: a range of possible values for some parameter with upper and lower limits. The interval is associated with a level of confidence (usually 95%) which is interpreted broadly as the probability that the true value of the parameter lies within the interval.

Connectivity: the degree to which the landscape facilitates animal or plant movement or spread and ecological flows.

Conservation: in relation to biodiversity, conservation is the protection, maintenance, management, sustainable use, restoration and improvement of the natural environment. In relation to natural and cultural heritage, conservation generally refers to the safekeeping or preservation of the existing state of a heritage resource from destruction or change.

Demand points: used in location-allocation analysis, to be the uniformly distributed set of locations that most efficiently service demands (like the location of fire stations). In ecology, demand points are uniformly distributed locations in continuous environmental space used in the environmental diversity (ED) method to calculate biodiversity complementarity.

Dispersal: the spread of animals and plants into new areas.

Disturbance: (Ecology) any process or event which disrupts ecosystem structure and resource availability.

Ecological connectivity: Biodiversity Indicator Program indicator [3.1b] ecological connectivity of terrestrial native habitat. This indicator is a measure of the value that habitat at each location contributes to habitat connectivity and estimates each location’s contribution to habitat connectivity and the ecological carrying capacity of surrounding locations by way of its habitat quality and its relative position in the landscape (e.g. as part of a habitat corridor, or a stepping stone) measured across multiple ecological scales. It can be reported for individual locations.

Ecological carrying capacity: Biodiversity Indicator Program indicator [3.1c] ecological carrying capacity of terrestrial native habitat. This indicator is a combined measure of a location's habitat quality and how well it's connected with surrounding habitat, and can be reported for individual locations and regions including the whole of New South Wales. It estimates the intactness and naturalness of terrestrial habitat for supporting biodiversity and considers how connectivity with surrounding habitat enables biological movement such as foraging, dispersal and migration across multiple ecological scales. It is used to account for the carrying capacity of locations and regions to support their original complement of biodiversity and ecosystems.

Ecological community: an assemblage of species occupying a particular area.

Ecological condition: Biodiveristy Indicator Program indicator [3.1a] ecological condition of terrestrial native habitat. This indicator is a measure of the quality of habitat at each location, and can be reported for individual locations and regions including the whole of New South Wales. It estimates the intactness and naturalness of terrestrial habitat without considering the indirect effects of surrounding habitat loss and fragmentation.

Ecological integrity: maintaining the diversity and quality of ecosystems and enhancing their capacity to adapt to change and provide for the needs of future generations.

Ecosystem: a dynamic complex of vegetable, animal and microorganism communities and their non-living environment that interact as a functional unit. Ecosystems may be small and simple, like an isolated pond, or large and complex, like a specific tropical rainforest or a coral reef in tropical seas.

Ecosystems classification: a delineation of areas, characterised by distinct species composition and where, ideally, each area is equally spaced in environmental space.

Environmental diversity (ED): is a specific surrogates-based approach to measuring biodiversity complementarity, where distance between locations in an environmental space is proportional to their complementarity in species composition.

Environmental space: a multidimensional mathematical space defined by environmental variables.

Effective habitat area: the proportion of residual habitat quality at a site following the impacts of clearing, degradation and fragmentation at the site and in its neighbourhood.

Expected diversity (expected survival): the number of features (species, genes, ecosystems) that are expected to be extant in 100 years' time, estimated from survival probabilities. Can be reported as a proportion of either original (ie. pre-industrial) or current diversity. The remainder is expected loss.

Extant: still in existence, surviving.

Extent: the area covered by something.

Extent of occurrence (EOO): as defined by the NSW Scientific Committee, the area of the total geographic range that includes all extant populations of the species.

Extinct: no individuals are remaining, either within a region or globally. A species or subspecific taxon is regionally extinct if no individuals remain within a region but are present elsewhere. A species (or subspecific taxon) with individuals in captivity but no individuals living independently in the wild would be 'Extinct in the Wild'. A species or subspecific taxon without viable populations (i.e. functionally extinct) in the long term is not extinct until the last known individual has died. Due to the difficulty detecting rare species, species or subspecific taxa that are presumed to be extinct are assumed to have a small but non-zero probability of survival.

Extrapolation: in statistics, the prediction of the value of a variable beyond its observed range based on its known relationship with another variable.

Fragmentation: the division of continuous habitat by vegetation clearance for human land-use activities, which isolates the remnant patches of vegetation and the species within them, and limits genetic flow between populations.

Gene flow: the transfer of genetic variation between populations by the dispersal of individuals.

Generalised dissimilarity modelling (GDM): a statistical technique for analysing and predicting spatial patterns of change in species composition across large regions.

Genetic divergence: the process in which two or more populations of an ancestral species accumulate independent **genetic** changes (mutations) through time, often after the populations have become reproductively isolated for some period of time.

Genetic diversity: the range of intrinsic differences in genes among individual organisms within a species, or among different species within a taxonomic group. There are several hypotheses to account for the emergence of genetic diversity. The two considered here are neutral and adaptive genetic diversity. Neutral genetic diversity results from the accumulation of neutral substitutions (i.e. processes such as gene flow, migration or dispersal which are influence by the spatial configuration of habitats and their relative connectivity across whole landscapes). Adaptive genetic diversity results from subpopulations of a species living in different environments that select for different alleles at a particular locus.

Geographic range: the area within which a species occurs.

Grid: a georeferenced spatial raster dataset consisting of a two-dimensional array of pixels each with cells containing numerical values representing spatial features as a continuous range of values or discrete categories with additional attributes optionally stored in an associated raster attribute table.

Grid cell: a single location of a specified size represented by an individual pixel in a raster dataset with a categorical or continuous value representing some measure or characteristic of that location.

Habitat: an area or areas occupied, or periodically or occasionally occupied, by a species, population or ecological community, including any biotic or abiotic component.

Index (plural indices): a metric used to quantify the information represented by an indicator.

Indicator: provides information on the condition of the environment. It may also include information on pressures on the environment, environmental conditions and societal responses.

IUCN: International Union for the Conservation of Nature, a union of government and non-government organisations that provides public, private and non-governmental organisations with the knowledge and tools that enable human progress, economic development and nature conservation to take place together.

Landscape: a heterogeneous area of local ecosystems and land uses that is of sufficient size to achieve long-term outcomes in the maintenance and recovery of species or ecological communities, or in the protection and enhancement of ecological and evolutionary processes.

Modelling: simulation of a process, concept, or the operation of a system, commonly with the aid of a computer.

Model: an abstract, usually mathematical, representation of a system, which is studied to gain understanding of the real system.

Monitoring: in this context, activities to collect new biophysical data.

Null Model: a distribution of values generated by a random process to determine if an observed pattern is discernable from a random process.

Population: (Ecology) a group of individuals of the same species occurring together in a particular area at a particular time.

Power curve: a non-linear relationship between two variables where one varies as the mathematical power of the other.

Pre-industrial era: A baseline epoch at which a region's habitat is expected to be in an intact state with no significant anthropogenic degradation or resultant loss of biodiversity. The Biodiversity Indicator Program estimates this to be c. 1750.

Rarefaction: (Ecology) technique to assess the relationship between an observed derived value (like the count of species or geographic range size) and the number of samples (occurrence records, counted individuals or sites). Rarefaction allows the prediction of the derived value for a given number of individual samples, based on the construction of so-called rarefaction curves.

Raster: Spatial categorical or continuously valued data classifying or measuring some aspect of a region as a square grid of equally sized grid cells at a specified resolution.

Red List of Ecosystems: a global standard for how to assess the risk of extinction status of ecosystems, applicable at local, national, regional and global levels. Under the auspices of the IUCN.

Red List of Threatened Species: a global standard for how to assess the risk of extinction status of species, applicable at local, national, regional and global levels. Under the auspices of the IUCN.

Representative species: a species (or subset of species) that represents or is typical of that group of species.

Species: a taxon comprising one or more populations of individuals capable of interbreeding to produce fertile offspring.

Species composition: the particular set of species at a particular location at a particular time.

Status: the condition or 'health' of a species, population, community, habitat or ecosystem, scaled relative to predetermined minimum and maximum potentials.

Suitable habitat: suitable habitat is predicted by identifying where each species lived originally and its associated environment.

Surrogate, biodiversity: a species, group of species or ecosystem that can be used as a substitute for wider biological groups.

Threatening process: a process that threatens, or that may threaten, the survival or evolutionary development of species or ecological communities.

Trends: directions of significant change in the environment, as shown by the changing values of measures (like essential variables, indicators or indices).

Vascular plant: plants containing vascular tissue (tissue specialised for the conduction of fluids); the more highly evolved plants above mosses and liverworts. Also known as tracheophytes or higher plants. Includes ferns, gymnosperms and flowering plants.

References

Alsos IG, Ehrich D, Thuiller W, Eidesen PB, Tribsch A, Schönswetter P, Lagaye C, Taberlet P & Brochmann C 2012, Genetic consequences of climate change for northern plants, in *Proceedings of the Royal Society of London, B*, 279(1735), pp.1–10.

Atlas of Living Australia 2018, *Atlas of Living Australia*, available at www.ala.org.au.

Bebber DP, Marriott FHC, Gaston KJ, Harris SA & Scotland RW 2007, Predicting unknown species numbers using discovery curves, in *Proceedings of the Royal Society of London, B*, 274(1618), pp.1651–1658.

Beier P & de Albuquerque FS 2015, Environmental diversity is a reliable surrogate for species representation, *Conservation Biology*, 29, pp.1401–1410.

Bell KL, Heard TA, Manion G, Ferrier S & van Klinken RD 2014, Characterising the phytophagous arthropod fauna of a single host plant species: Assessing survey completeness at continental and local scales, *Biodiversity and Conservation*, 23(12), pp.2985–3003.

Bland LM, Keith DA, Miller RM & Murray NJ (eds.) 2017, *Guidelines for the application of IUCN Red List of Ecosystems categories and criteria*, International Union for the Conservation of Nature, Gland, Switzerland, pp.1–110.

Boyce MS 1992, Population viability analysis, *Annual Review of Ecology and Systematics*, 23, pp.481–506.

Bradshaw CJA 2012, Little left to lose: Deforestation and forest degradation in Australia since European colonization, *Journal of Plant Ecology*, 5(1), pp.109–120.

Brook BW, O’Grady JJ, Chapman AP & Burgman MA 2000, Predictive accuracy of population viability analysis in conservation biology, *Nature*, 404(6776), pp.385–387.

Bruford MW, Davies N, Dulloo ME, Faith DP & Walters M 2017, Monitoring changes in genetic diversity, in M Walters & RJ Scholes (eds), *The GEO handbook on biodiversity observation networks*, Springer, Berlin, Germany, pp.107–128.

Bush AA, Nipperess DA, Duursma DE, Theischinger G, Turak E & Hughes L 2014, Continental-scale risk to the Australian Odonata from climate change, *PLoS One*, 9(2), e88958.

Butchart SHM, Akçakaya HR, Chanson J, Baillie JEM, Collen B, Quader S, Turner WR, Amin R, Stuart SN & Hilton-Taylor C 2007, Improvements to the Red List index, *PLoS One*, 2(1), e140.

Butchart SHM, Stattersfield AJ, Bennun LA, Shutes SM, Akçakaya HR, Baillie JEM, Stuart SN, Hilton-Taylor C & Mace GM 2004, Measuring global trends in the status of biodiversity: Red List indices for birds, *PLoS Biology*, 2(12), e383.

Cardillo M, Mace GM, Gittleman JL, Jones KE, Bielby J & Purvis A 2008, The predictability of extinction: Biological and external correlates of decline in mammals, *Proceedings of the Royal Society of London, B*, 275(1641), pp.1441–1448.

Ceballos G & Ehrlich PR 2009, Discoveries of new mammal species and their implications for conservation and ecosystem services, *Proceedings of the National Academy of Sciences*, 106(10), pp.3841–3846.

Chao A 1984, Non-parametric estimation of the number of classes in a population, *Scandinavian Journal of Statistics*, 11, pp.265–270.

Chapman AD 2009, *Numbers of living species in Australia and the world*, 2nd edition, Australian Biological Resources Study, Canberra, Australia, pp.1–80.

Collen B, Loh J, Whitmee S, McRae L, Amin R & Baillie JEM 2009, Monitoring change in vertebrate abundance: The Living Planet Index, *Conservation Biology*, 23(2), pp.317–327.

Council of Heads of Australasian Herbaria 2018, *Australian Plant Census*, Council of Heads of Australasian Herbaria, available at www.anbg.gov.au/chah/apc/.

Crooks KR, Burdett CL, Theobald DM, King SRB, Di Marco M, Rondinini C & Boitani L 2017, Quantification of habitat fragmentation reveals extinction risk in terrestrial mammals, *Proceedings of the National Academy of Sciences,* 114(29), pp.7635–7640.

Department of the Environment and Energy (Australia) 2018, *Common assessment method,* Department of the Environment and Energy (Australia), Canberra, Australia, available at www.environment.gov.au/biodiversity/threatened/cam.

DPIE 2020, *NSW Biodiversity Outlook Report: Results from the Biodiversity Indicator Program, first assessment*, Department of Planning, Industry and Environment NSW, Sydney, Australia.

Drielsma MJ, Howling G & Love J 2012, *NSW native vegetation management benefits analyses: Technical report*, NSW Office of Environment and Heritage, Sydney, Australia.

Drielsma MJ, Love J, Williams KJ, Manion G, Saremi H, Harwood T & Robb J 2017, Bridging the gap between climate science and regional-scale biodiversity conservation in south-eastern Australia, *Ecological Modelling*, 360, pp.343–362.

Drielsma MJ, Love J, Williams KJ, White M & Ferrier S 2020, *A model-based indicator of capacity for biodiversity persistence using vascular plant records and habitat condition,* Biodiversity Indicator Program Implementation Report, Department of Planning, Industry and Environment NSW, Sydney, Australia.

Elith J & Leathwick JR 2009, Species distribution models: Ecological explanation and prediction across space and time, *Annual Review of Ecology, Evolution, and Systematics*, 40(1), pp.677–697.

ESRI 2016, *ArcGIS desktop: release 10.5* (software), Environmental Systems Research Institute, Redlands, California, USA.

Evans JP, Ji F, Lee C, Smith P, Argüeso D & Fita L 2014, Design of a regional climate modelling projection ensemble experiment – NARCliM, *Geoscientific Model Development*, 7, pp.621–629.

Faith DP 2008, Threatened species and the potential loss of phylogenetic diversity: Conservation scenarios based on estimated extinction probabilities and phylogenetic risk analysis, *Conservation Biology*, 22(6), pp.1461–1470.

Faith DP & Walker P 1996, Environmental diversity: On the best-possible use of surrogate data for assessing the relative biodiversity of sets of areas, *Biodiversity and Conservation*, 5(4), pp.399–415.

Faith DP, Ferrier S & Walker P 2004, The ED strategy: How species-level surrogates indicate general biodiversity patterns through an ‘environmental diversity’ perspective, *Journal of Biogeography*, 31, pp.1207–1217.

Faith DP, Minchin P & Belbin L 1987, Compositional dissimilarity as a robust measure of ecological distance, *Vegetatio*, 69, pp.57–68.

Ferrier S 2002, Mapping spatial pattern in biodiversity for regional conservation planning: Where to from here?, *Systematic Biology*, 51(2), pp.331–363.

Ferrier S, Manion G, Elith J & Richardson K 2007, Using generalized dissimilarity modelling to analyse and predict patterns of beta diversity in regional biodiversity assessment, *Diversity and Distributions*, 13(3), pp.252–264.

Funk VA, Richardson KS & Ferrier S 2005, Survey-gap analysis in expeditionary research: Where do we go from here?, *Biological Journal of the Linnean Society*, 85, pp.549–567.

Gaston KJ & Fuller RA 2009, The sizes of species' geographic ranges, *Journal of Applied Ecology*, 46(1), pp.1–9.

Gaston KJ, Blackburn TM, Greenwood JJD, Gregory RD, Quinn RM & Lawton JH 2000, Abundance–occupancy relationships, *Journal of Applied Ecology*, 37(s1), pp.39–59.

Haque MM, Nipperess DA, Gallagher RV & Beaumont LJ 2017, How well documented is Australia's flora? Understanding spatial bias in vouchered plant specimens, *Austral Ecology*, 42, pp.690–699.

Harwood TD, Donohue RJ, Williams KJ, Ferrier S, McVicar TR, Newell G & White M 2016, Habitat condition assessment system: A new way to assess the condition of natural habitats for terrestrial biodiversity across whole regions using remote sensing data, *Methods in Ecology and Evolution*, 7(9), pp.1050–1059.

Hoffman JI, Simpson F, Patrice D, Rijks JM, Kuiken T, Thorne MAS, Lacy RC & Dasmahapatra KK 2014, High-throughput sequencing reveals inbreeding depression in a natural population, *Proceedings of the National Academy of Sciences*, 111(10), pp.3775–3780.

Hoffman M, Hilton-Taylor C, Angulo A, Böhm M et al. 2010, The impact of conservation on the status of the world's vertebrates, *Science*, 330, pp.1503–1509.

Holsinger KE & Weir BS 2009, Genetics in geographically structured populations: Defining, estimating and interpreting F_{ST}, *Nature Reviews Genetics*, 10(9), pp.639–650.

Hortal J, de Bello F, Diniz-Filho JAF, Lewinsohn TM, Lobo JM & Ladle RJ 2015, Seven shortfalls that beset large-scale knowledge of biodiversity, *Annual Review in Ecology, Evolution and Systematics*, 46(1), pp.523–549.

Hurlbert S 1971, The nonconcept of species diversity: A critique and alternative parameters, *Ecology*, 52, pp.577–586.

IUCN 2012, *IUCN Red List categories and criteria: version 3.1*, 2nd edition, International Union for the Conservation of Nature, Gland, Switzerland, pp.1–38.

IUCN 2017, *Guidelines for using the IUCN Red List categories and criteria: version 13*, Standards and Petitions Subcommittee, International Union for the Conservation of Nature, http://cmsdocs.s3.amazonaws.com/RedListGuidelines.pdf.

IUCN 2018, *The IUCN Red List of Threatened Species, version 2018-1*, http://www.iucnredlist.org.

Kindvall O & Gärdenfors U 2003, Temporal extrapolation of PVA results in relation to the IUCN Red List criterion E, *Conservation Biology*, 17(1), pp.316–321.

Kissling WD, Ahumada JA, Bowser A, Fernandez M et al. 2018, Building essential biodiversity variables (EBVs) of species distribution and abundance at a global scale, *Biological Reviews*, 93(1), pp.600–625.

Leimu R & Fischer M 2008, A meta-analysis of local adaptation in plants, *PLoS One*, 3(12), e4010.

Love J, Drielsma MJ, Williams KJ & Thapa R 2020, *Integrated model-data fusion approach to measuring habitat condition for ecological integrity reporting: Implementation for habitat condition indicators*, Biodiversity Indicator Program Implementation Report, Department of Planning, Industry and Environment NSW, Sydney, Australia.

Mace GM, Collar NJ, Gaston KJ, Hilton-Taylor C, Akçakaya HR, Leader-Williams N, Milner-Gulland EJ & Stuart SN 2008, Quantification of extinction risk: IUCN's system for classifying threatened species, *Conservation Biology*, 22(6), pp.1424–1442.

MacKenzie DI, Nichols JD, Lachman GB, Droege S, Royle JA & Langtimm CA 2002, Estimating site occupancy rates when detection probabilities are less than one, *Ecology*, 83(8), pp.2248–2255.

Manion G & Ridges M 2009, An optimisation of the survey gap analysis technique to minimise computational complexity and memory resources in order to accommodate fine grain environmental and site data, in RS Anderssen, RD Braddock and LTH Newham (eds), *18th World IMACS Congress and MODSIM09 International Congress on Modelling and Simulation*, Modelling and Simulation Society of Australia and New Zealand and International Association for Mathematics and Computers in Simulation, 2514–2520.

Manion G 2018, *.NET survey gap analysis tool version 7.0* (software), Cryzan (with NSW Office of Environment and Heritage and CSIRO), Armidale, Australia.

Mimura M, Yahara T, Faith DP, Vázquez Domínguez E, Colautii RI, Araki H, Javadi F, Núñez-Farfán J, Mori AS, Zhou S, Hollingsworth PM, Neaves LE, Fukano Y, Smith GF, Sato Y-I, Tachida H & Hendry AP 2017, Understanding and monitoring the consequences of human impacts on intraspecific variation, *Evolutionary Applications*, 10, pp.121–139.

Mooers AO, Faith DP & Maddison WP 2008, Converting endangered species categories to probabilities of extinction for phylogenetic conservation prioritization, *PLoS One*, 3(11), e3700.

Mora C, Tittensor DP, Adl S, Simpson AGB & Worm B 2011, How many species are there on Earth and in the ocean?, *PLoS Biology*, 9(8), e1001127.

Muggeo VMR 2008, segmented: An R package to fit regression models with broken-line relationships, *R News*, 8(1), pp.20–25.

Nipperess DA 2015, A separate creation: Diversity, distinctiveness and conservation of Australian wildlife, in AJ Stow, N Maclean & GI Holwell (eds), *Austral ark: The state of wildlife in Australia and New Zealand*, Cambridge, UK, Cambridge University Press, pp.1–23.

Nipperess DA & Beattie AJ 2004, Morphological dispersion of *Rhytidoponera* assemblages: The importance of spatial scale and null model, *Ecology*, 85, pp.2728–2736.

Nipperess DA, Faith DP, Williams KJ, King D, Manion G, Ware C, Schmidt R, Love J, Drielsma M, Allen S & Gallagher R 2020, Expected survival and state of all known species: Data packages for the Biodiversity Indicator Program, first assessment, *SEED Portal*, datasets.seed.nsw.gov.au/dataset/biodiversity-indicator-program-data-packages.

Noël S, Ouellet M, Galois P & Lapointe F-J 2006, Impact of urban fragmentation on the genetic structure of the eastern red-backed salamander, *Conservation Genetics*, 8(3), pp.599–606.

Noss R 1990, Indicators for monitoring biodiversity: A hierarchical approach, *Conservation Biology*, 4, pp.355–364.

OEH 2013, *Saving our Species technical report*, Office of Environment and Heritage NSW, Sydney, Australia, 25 pp.

OEH 2016, *Biodiversity impacts and adaptation project (stage II) NSW and ACT regional climate modelling project*, Office of Environment and Heritage NSW, Sydney, Australia.

OEH & CSIRO 2019, *Measuring biodiversity and ecological integrity in New South Wales: Method for the Biodiversity Indicator Program*, Office of Environment and Heritage NSW and Commonwealth Scientific and Industrial Research Organisation, Sydney, Australia.

Pereira HM & Cooper HD 2006, Towards the global monitoring of biodiversity change, *Trends in Ecology & Evolution*, 21(3), pp.123–129.

R Core Team 2017, *R: A language and environment for statistical computing* (software), R Foundation for Statistical Computing, Vienna, Austria, www.R-project.org/.

Radford J, Bennett A & Cheers G 2005, Landscape-level thresholds of habitat cover for woodland-dependent birds, *Biological Conservation*, 124(3), pp.317–337.

Royal Botanic Gardens and Domain Trust 2018, *PlantNET: The NSW Plant Information Network System,* Royal Botanic Gardens and Domain Trust, available at http://plantnet.rbgsyd.nsw.gov.au.

Secretariat of the CBD 2011, *Strategic plan for biodiversity 2011–2020 and the Aichi targets "living in harmony with nature"*, Secretariat of the Convention on Biological Diversity, IUCN and UNEP, http://www.cbd.int/sp/.

Slatkin M 1993, Isolation by distance in equilibrium and non-equilibrium populations, *Evolution*, 47(1), pp.264–279.

Su J, Debinski D, Jakabauskas M & Kindscher K 2004, Beyond species richness: Community similarity as a measure of cross-taxon congruence for coarse-filter conservation, *Conservation Biology*, 18, pp.167–173.

Threatened Species Recovery Hub (undated), *Developing a threatened species index* (NESP Project 3.1), Threatened Species Recovery Hub, National Environmental Science Programme, Brisbane, Australia, www.nespthreatenedspecies.edu.au/projects/threatened-species-index.

Ware C, Williams KJ, Harding J, Hawkins B, Harwood T, Manion G, Perkins GC & Ferrier S 2018, Improving biodiversity surrogates for conservation assessment: A test of methods and the value of targeted biological surveys, *Diversity and Distributions* (in press).

Weitzman M 1992, On diversity, *The Quarterly Journal of Economics*, 107(2), pp.363–405.

Whitlock R 2014, Relationships between adaptive and neutral genetic diversity and ecological structure and functioning: A meta-analysis, *Journal of Ecology*, 102(4), pp.857–872.

Whittaker RH 1972, Evolution and measurement of species diversity, *Taxon*, 21(2/3), pp.213–251.

Legislation

Biodiversity Conservation Act 2016 (NSW Government), available at www.legislation.nsw.gov.au/#/view/act/2016/63, version as of 1 July 2018.