26 February 2018

Laura Torrible
Team Leader
Biodiversity Reform and Sustainability Team
Resources and Industry Policy Branch
320 Pitt Street
Sydney NSW 2001

Dear Laura,

Re: Western Sydney Strategic Plan – Threatened Species Habitat Mapping

Project no. 26276

This paper outlines the key concepts, goals and methodology for conducting the threatened species mapping for the Priority Growth Areas (PGAs) as part of the Cumberland Plain Conservation Plan (CPCP).

Methodology

Biosis understand that the aim of the species habitat mapping is to develop habitat maps for Category 1 EPBC Act listed threatened species, species credit species and migratory species, identified within the PGAs.

The purpose of this mapping is to:

- Inform the assessment of the impacts and benefits of the CPCP in the context of threatened species habitat remaining across the Cumberland Plain.
- Support extrapolation of threatened species survey data across vegetation zones, including presence/absence data and identification of species polygons.

The mapping will be based on the following tools and datasets, including, but not limited to:

- Atlas of NSW Wildlife (BioNet).
- Other existing species records available.
- Refined mapping of PCTs and condition undertaken by Biosis in Summer 2018.
- Remnant Vegetation of the Western Cumberland Subregion 2013 Update VIS_ID 4207 (OEH 2013).
- The Native Vegetation of the Sydney Metropolitan Area VIS_ID 4489 (OEH 2016).
- Terrestrial Vertebrate Fauna of the Greater Southern Sydney Region Volume 2 (DECC 2007).
- Regional Soils Landscape layers.
- Hydrological modelling obtained from LiDAR data.
- Topographic information, including height and slope, obtained from LiDAR data.

- NSW/Commonwealth Threatened Species Profiles, conservation advices and recovery plans.
- OEH Threatened Biodiversity Data Collection.

The primary data for habitat mapping is species distribution data, typically available as occurrence records, where an observer has noted the location of an organism at a given point in time. However, incomplete or biased sampling and difficulties in detecting species can mean that these records often do not represent the entire habitat for a species. Therefore, understanding impacts and/or prioritising sites for conservation based solely on known occurrence data is likely to be biased towards those sites that are well sampled, or where common and easy to survey species are located, with the very real risk of missing important habitats.

One approach to overcome these issues is to use species distribution modelling, which, given sufficient data, can account for detectability and biased sampling to provide accurate predictions of where habitat for a given species is likely to occur. However, the accuracy of species distribution modelling predictions depends strongly on the occurrence data, where the number and age of the records available for each species have an influence. In addition, species distribution modelling also requires data on the factors (or 'covariates') that drive the species occurrence, its detectability and the biases in sampling (e.g., vegetation cover, soil and elevation data, population or road density etc.). The quality of the output of the species distribution modelling is dependent on the amount and quality of the data representing all these factors.

Where much of this data is missing, other approaches, such as the use of knowledge based information, as described below, will be used.

An initial desktop assessment has identified 18 threatened flora species and eight threatened fauna species known to occur within the PGAs. Of these species, seven flora species and six fauna species have less than 50 records and may not be appropriate for species distribution modelling approaches. Therefore, a simplified map will be prepared for these species as described below.

Considering this we propose to undertake the species habitat mapping using two methodologies:

1. Detailed mapping using species distribution modelling.
2. Simplified mapping using knowledge base information.

Detailed mapping using species distribution modelling

Species distribution modelling provides a tool that can predict the likely distribution of a species based on known occurrence data and environmental conditions at these localities. The technique has been used successfully on the Lower Hunter Valley Strategic Plan (Kujala. H et.al, 2015) and provides a robust and transparent method for predicting the distribution of species from available data. The maps will identify the suitable habitat for each of the species within the Cumberland Plain, either through expert interpretation of knowledge base information, or species distribution modelling approaches.

Where possible, habitat maps will be developed in the form of statistical models of predicted habitat distribution using specialised modelling software (e.g. MaxEnt and Generalised Additive Models using the S+ Statistical package in ESRI ArcGIS). It is noted that statistical modelling approaches rely on:

- There being an adequate number of species records available for use in the analysis.
- The availability of spatial environmental layers of appropriate resolution and accuracy (covariates).
- There being strong enough relationships between species locations and the environmental variables for which spatial data layers exist.
- For threatened species where detectability is an issue, data will be required on factors that influence detectability (e.g. vegetation height).

- For threatened species where the location data is highly biased, data will be required that describe these biases (e.g. population density, locations of popular bird watching spots, etc.)

Threatened species profiles will be prepared to define “suitable habitat”, for each species within the PGA, to inform the statistical models, for example, patch size, distance from waterways, PCT’s, geographic constraints, and height above sea level.

An additional approach that may be explored for species that have small number of occurrences within the PGA, is to expand the study area to a larger region where there are a greater number of records. A species distribution model can then be produced for this larger region, and then clipped back down to the original study area once the modelling is finished.

The species distribution modelling will be undertaken at RMIT University, under the supervision of Dr Ascelin Gordon with help from Vira Koshkina a PhD student in the RMIT School of Science undertaking her PhD in species distribution modelling. The analysis will proceed with the following steps:

1. Determining which flora and fauna species are appropriate for species distribution models obtaining information on the likely important covariate for each species, whether detectability is likely an issue and analysis of any spatial biases in the point records for each species.
2. Decisions on whether and how much data to hold aside for testing/validation of models (for small numbers of records, it may not be appropriate to hold out data).
3. Gathering data on variables, preparation and validation of this covariate data.
4. Exploratory modelling with Maxent, using different combinations of covariates to produce initial species distribution models.
5. Model selection procedure to decide which covariates to use for each species to generate candidate final species distribution models.
6. Gathering feedback via phone with species experts on the model outputs for each species regarding the realism of the model and where it may be predicting incorrectly.
7. Refinement of the model based on feedback to produce final species distribution models.
8. Validation of the models using testing data (if any) and generating estimates of uncertainty on the model outputs.

Simplified mapping using knowledge base information

Where statistical approaches are not appropriate due to lack of species data, or lack of data on appropriate covariates, we will develop broad indicative habitat maps using the best available information for each species, including known records and habitat preferences, based on the data sources listed below. These maps will identify known, likely and potential habitat for each species.

Data Sources required

The following spatial and non-spatial data is required to undertake the species habitat mapping in accordance with the methodology described above:

- EPBC Protected Matters Tool.
- Commonwealth SPRAT database.
- Atlas of NSW Wildlife (BioNet).
- OEH Threatened Biodiversity Data Collection.
- Remnant Vegetation of the Western Cumberland Subregion 2013 Update VIS_ID 4207 (OEH 2013).
- The Native Vegetation of the Sydney Metropolitan Area VIS_ID 4489 (OEH 2016).
- Cumberland Plain Recovery Plan (OEH 2011).

- Other recovery plans/conservation advices.
- NSW Threatened Species Profiles.
- BirdLife Australia shorebird data.
- Terrestrial Vertebrate Fauna of the Greater Southern Sydney Region Volume 2 (DECC 2007).
- Hydrological modelling derived from LiDAR data.
- Topographic information, including height and slope, obtained from LiDAR data.
- Temperature data (e.g. mean annual temperature, mean temperature of the coldest/hottest period (ANUCLIM).
- Rainfall and solar radiation data (e.g. mean annual rainfall and mean annual solar radiation (ANUCLIM).
- The altitude above sea level and slope.
- Soils data digital atlas of Australian soils.

For the threatened species distribution modelling this is often the sort of data that is also required.

Outputs

The completion of the species habitat mapping will result in the creation of GIS layers that provide species habitat maps based on the available data and stratifies NSW and Commonwealth listed species into:

- Detailed mapping using species distribution modelling.
- Simplified mapping using knowledge base information.

It is worth noting that it is expected that most or all of the species data available will be presence-only data, meaning it is data without any 'absence' records of species (i.e. where observers have not noted the absence of a species at surveyed sites). Undertaking species distribution modelling with presence-only data results in relative measures of habitat suitability. This means, for example, that a location with a predicted high habitat suitability score is only high relative to other locations in the study area, and not necessarily high quality habitat in absolute terms. Thus, care needs to be taken in interpreting the outputs of species distribution models and surveys, or other approaches should be used to validate the results.

Assumptions and limitations

The following assumptions have been made with regard to this methodology:

- The outputs from the analysis are only as good as the input data. Poor quality (i.e. inaccurate species data) or missing data may result in the identification of conservation priorities or potential development impacts that do not align with the Western Sydney environment.
- Ideally the conservation area vegetation mapping will be undertaken before the species mapping as this will add additional areas of higher confidence vegetation condition and extent mapping.
- Species distribution models may not be possible, even for species with large numbers of occurrences, if the interactions driving the current locations of a species are cryptic, or depend on factors for which there is not data (e.g. how specific landuses have changed over time at a location, or whether there has been firewood collection etc.).

Please contact me on (02) 4201 1054 if you would like to discuss further.

Yours sincerely

Rebecca Dwyer

Team Leader - Ecology

References

ANUCLIM Temperature data Online resource: http://fennerschool.anu.edu.au/research/products/anuclim-vrsn-61

Kujala H, Whitehead AL & Wintle BA (2015) Identifying conservation priorities and assessing impacts and trade‐offs of potential future development in the Lower Hunter Valley in New South Wales. The University of Melbourne, Melbourne, Victoria. Pp. 100

Soils data digital atlas of Australian soils Online Resource: https://data.gov.au/dataset/9e7d2f5b-ff51-4f0f-898a-a55be8837828