Data Article

Data representing climate-induced changes in the spatial distribution of key bee forage species for southwest Western Australia

Vidushi Patel\textsuperscript{a,b,*}, Bryan Boruff\textsuperscript{a,b,c}, Eloise Biggs\textsuperscript{c}, Natasha Pauli\textsuperscript{a,c}

\textsuperscript{a} UWA School of Agriculture and Environment, The University of Western Australia, 35 Stirling Highway, Crawley 6009, Western Australia, Australia
\textsuperscript{b} Cooperative Research Center for Honey Bee Products, The University of Western Australia, 35 Stirling Highway, Agriculture North MO85, Crawley 6009, Western Australia, Australia
\textsuperscript{c} Department of Archaeology, Geography and Anthropology, School of Social Sciences, The University of Western Australia, 35 Stirling Highway, Crawley 6009, Western Australia, Australia

\textbf{A R T I C L E   I N F O}

Article history:
Received 27 September 2022
Revised 17 November 2022
Accepted 21 November 2022
Available online 25 November 2022

Dataset link: Climate induced change in spatial distribution of bee forage species in Southwest Western Australia (Original data)

Keywords:
Migratory beekeeping
Bee forage availability
Species distribution models
Climate change
Range shift

\textbf{A B S T R A C T}

The dataset includes (i) species occurrence points, and (ii) Species Distribution Model (SDM) outputs under current conditions and a moderate emission (RCP 6.0) climate scenario, for 30 key bee forage species in southwest Western Australia (WA). Occurrence data were obtained from open data sources and through stakeholder engagement processes. SDM outputs were predicted using the Maxent algorithm with the change in species range analysed using QGIS software. The model outputs provide insight into the potential implications of climate change on important bee forage species in southwest WA, including dominant melliferous tree and shrub species. Changes in these species are likely to have repercussions to the ecological and social systems where a facilitatory relationship exists. This dataset is important for informing conservation efforts within the southwest Australian biodiversity hotspot.

\* Corresponding author.
E-mail address: vidushi.patel@uwa.edu.au (V. Patel).

Social media: @vidushi_patel (V. Patel), @hazographer (B. Boruff), @EllieMBiggs (E. Biggs), @natasha_pauli (N. Pauli)

https://doi.org/10.1016/j.dib.2022.108783
2352-3409/© 2022 The Authors. Published by Elsevier Inc. This is an open access article under the CC BY-NC-ND license (http://creativecommons.org/licenses/by-nc-nd/4.0/)
Specifications Table

| Subject | Ecological modeling |
|---------|---------------------|
| Specific subject area | Geography |
| Type of data | Table; Maps; Figure |
| How the data were acquired | Key bee forage species were identified through semi-structured interviews with WA commercial beekeepers [1], and from verified citizen science and WA herbarium records collated through the national open-source biodiversity data platform, the Atlas of Living Australia (ALA) [2]. These data were then used to identify species occurrence points and model spatial changes in species distribution for 30 melliferous species across southwest WA. Environmental data used in the SDM for current (1976 – 2005) and future (2055) time periods were bioclimatic variables prepared for Australia (sources listed in data accessibility section). |
| Data format | Filtered and analysed. |
| Description of data collection | The point locations of key forage species were identified through semi-structured interviews and participatory GIS mapping exercises conducted with 14 commercial beekeepers owning 50 or more beehives. Additional species occurrence data were collected from the ALA for 1950 to 2021. |
| Data source location | Region: Southwest Western Australia |
| Data accessibility | All filtered and analysed data are available at Patel et al [2]. Repository name: Mendeley Data Data identification number: 10.17632/9vnztvcr9c.3 Direct URL to data: https://data.mendeley.com/datasets/9vnztvcr9c.3 Other data sources: 1. Species occurrence data: https://spatial.ala.org.au [3] 2. Climate data: (i) Baseline scenario: http://www.bom.gov.au/jsp/awap/ [4] (ii) Future scenario: Vanderwal [5] Further information on the thematic methodology for the processed semi-structured interview data is available in Patel et al [1]. |
| Related research article | V. Patel, E. M. Biggs, N. Paul, and B. Boruff, Using a social-ecological system approach to enhance understanding of structural interconnectivities within the beekeeping industry for sustainable decision making, Ecology and Society. 2020 25(2):24. https://doi.org/10.5751/ES-11639-250224. |

Value of the Data

- The data present current and future occurrence maps for 30 species native to the southwest Australian biodiversity hotspot and provides estimates of potential shifts in distribution ranges based on a moderate emission (RCP 6.0) climate scenario.
- The data are beneficial to researchers examining the impact of climate change on the ecosystem service provision of melliferous flora in Western Australia.
- The data will inform conservation efforts within the southwest Australian biodiversity hotspot by providing an assessment of climate-induced change in the geographic distribution of key melliferous species.

---

1 The Atlas of Living Australia (ALA) is an Australian node of the Global Biodiversity Information Facility (GBIF).
1. Objective

The research this publication supports [1], which describes the structural interconnectivities between bees, beekeeping, and forage landscapes, and explains how natural and anthropogenic pressures acting upon these landscapes will affect the beekeeping system of Western Australia (WA). The data presented here was generated to support an enhanced understanding of the impacts of climate change on bee forage found within the South West Australia Ecoregion (SWAE), Australia’s only biodiversity hotspot. The 30 species included in this study represent the key forage species used for honey production in Western Australia. This data illustrates how key bee forage species distribution will change relative to a future climate scenario (Representative Concentration Pathway 6.0, Global Climate Model (GCM) CSIRO Mk3 - 2055), understanding how access to these species will change, and how such changes will alter beehive migration patterns across the state was the catalyst for this analysis and therefore defined the scope.

2. Data Description

The data provided here include (i) species occurrence points, (ii) current and future distribution range maps, and the magnitude and direction of shifts in distribution ranges, of 30 species targeted for honey production in Western Australia. The target species used in this study are listed in Table 1. Species occurrence data were compiled using records obtained through the Atlas of Living Australia (https://spatial.alanda.org.au), cleaned by removing incomplete information or incorrect coordinates, and merged with occurrence records collected through participatory GIS mapping exercises conducted with individual commercial beekeepers.

Table 1
List of thirty species targeted for honey production by beekeepers in Western Australia.

| Species Name | Common Name | Type |
|--------------|-------------|------|
| Banksia attenuata | Candle banksia/Yellow banksia/Slender banksia | Tree |
| Banksia menziesii | Firewood banksia/Red banksia/Menzies banksia | Shrub |
| Banksia sessilis | Parrot bush | Shrub |
| Banksia sphaerocarpa | Fox Banksia | Shrub |
| Calothamnthus quadrifidus | One-sided bottlebrush | Shrub |
| Corymbia calophylla | Marri/Red gum/Port Menziesy gum | Tree |
| Eucalyptus accedens | Powderbark wandoo | Tree |
| Eucalyptus annulata | Open-fruiting mallee | Tree |
| Eucalyptus burracoppinensis | Burracoppin mallee | Tree |
| Eucalyptus cornuta | Yate | Tree |
| Eucalyptus diversicolor | Karri | Tree |
| Eucalyptus dundasii | Dundas blackbutt | Tree |
| Eucalyptus focktoniae | Merrit | Tree |
| Eucalyptus incrassata | Lerp mallee/Yellow mallee | Tree |
| Eucalyptus lesouefii | Goldfields blackbutt | Tree |
| Eucalyptus longicornis | Red morrel/Morrel | Tree |
| Eucalyptus luxophleba | York gum | Tree |
| Eucalyptus marginata | Jarrah | Tree |
| Eucalyptus melanoxylon | Black morrel | Tree |
| Eucalyptus occidentalis | Flat-topped yate/Swamp yate | Tree |
| Eucalyptus platypus | Moort | Tree |
| Eucalyptus ravida | Bronze & silver gimlet | Tree |
| Eucalyptus reducna | Black marlock/Mallee form of Wandoon | Tree |
| Eucalyptus salubris | Gimlet | Tree |
| Eucalyptus stricklandii | Strickland’s gum | Tree |
| Eucalyptus transcontinentalis | Redwood | Tree |
| Eucalyptus wandoo | Wandoon/White gum | Tree |
| Hakea trifurcata | Two-leaf hakea/White bush/Kangaroo | Shrub |
| Leucopogon conosephioides | May flower/White bell | Shrub |
| Leucopogon oldfieldii | Oldfields beard-heath | Shrub |
SDM outputs were generated using MaxEnt. SDM performance was based on the Area Under the Curve (AUC) and True Skills Statistics (TSS). The AUC values approaching 1.0 for all model outputs indicated good predictive performance. Maximum training sensitivity plus specificity logistic thresholding was used to convert each species distribution model to a binary presence-absence grid. SDM outputs mapped for the 30 individual species are presented in Patel et al. [2], which indicate current species distribution and change in distribution relative to the moderate emission (RCP 6.0) climate scenario. The magnitude and direction of species distribution range shift is provided in Patel et al. [2] as Table 1.

3. Experimental Design, Materials and Methods

3.1. Identifying key forage species

To identify the forage species targeted by commercial apiarists, semi-structured interviews were conducted with 29 beekeepers (operating more than 50 hives). Human ethics approval was attained to undertake this research. The participants were selected under the guidance of the Bee Industry Council of Western Australia (BICWA), and the snowballing method. From the interview data, important bee forage species targeted by beekeepers were shortlisted (n = 30; Table 1) as the most mentioned species by the participants. The coordinates for each apiary permit owned by a participant beekeeper was collected from the WA state government Department of Biodiversity, Conservation and Attractions (DBCA). A Participatory Geographic Information System (PGIS) mapping approach was then used with each interviewee to identify specific target species for each apiary permit.

Additionally, species occurrence data for the 30 shortlisted species were extracted from the Atlas of Living Australia (https://spatial.alaska.ala.org.au) spatial portal. Occurrence records were clipped using the study area boundary, prepared from spatial data for the Interim Biogeographic Regionalisation for Australia (IBRA) (https://www.environment.gov.au/). PGIS mapping data and ALA records were combined, providing comprehensive representation of species occurrence samples, as an aim to overcome spatial bias in ALA records [6,7]. GIS vector files (shapefiles) from ALA and PGIS data were merged in QGIS 3.10 to compile occurrence points for each species-

3.2. Species spatial extent

Modern SDM in Australia began with the development of the BIOCLIM package in 1984 [8], and continuously progressed to more complex machine learning algorithms including MaxEnt [9]. In this study, easy-to-use presence-only method Maxent was used to obtain the geographic distributions of key bee forage species due to its high prediction performance when compared to other known methods [10]. MaxEnt uses presence-only data and background (pseudo-random) points randomly distributed across the study extent to estimate the closest to uniform (maximum entropy) distribution for a range of independent environmental variables [11]. Species occurrence data were randomly allocated as 70% training and 30% test data for species distribution modelling in MaxEnt version 3.4.1. SDMs for each species were calculated using 10,000 pseudo-

---

2 SDM outputs in Patel et al. [2] include baseline distribution, future distribution, change in distribution, and species range shift maps.
3 Interview themes used in semi-structured interviews are available in related research article Patel et al. [1].
4 Atlas of Living Australia (ALA) is a platform for providing open source biodiversity data covering over 85 million records of more than 111,000 species, aggregated from multiple sources and citizen science across Australia (ALA, 2020). Bias in ALA data has been recognised in the literature with recommendations for approaches such as additional sampling and digitising to overcome data quality gaps [7].
list of 19 bioclimatic variables obtained to represent baseline and future scenarios [4,5,13]. The variables selected for use in species distribution modelling (SDM) using MaxEnt software are highlighted with bold letters.

| Code | Variable |
|------|----------|
| Bio1 | Mean annual temperature |
| Bio2 | Mean diurnal range |
| Bio3 | Isothermality |
| Bio4 | Temperature seasonality |
| Bio5 | Max temperature of warmest month |
| Bio6 | Min temperature of coldest month |
| Bio7 | Temperature annual range |
| Bio8 | Mean temperature of wettest quarter |
| Bio9 | Mean temperature of driest quarter |
| Bio10 | Mean temperature of warmest quarter |
| Bio11 | Mean temperature of coldest quarter |
| Bio12 | Annual precipitation |
| Bio13 | Precipitation of wettest month |
| Bio14 | Precipitation of driest month |
| Bio15 | Precipitation seasonality |
| Bio16 | Precipitation of wettest quarter |
| Bio17 | Precipitation of driest quarter |
| Bio18 | Precipitation of warmest quarter |
| Bio19 | Precipitation of coldest quarter |

random points and six bioclimatic variables to obtain the logistic outputs for each species using a five-fold cross-validation. To increase model performance, only 'hinge features' were used [12].

3.3. Climate scenarios

Species distributions were obtained for two climate scenarios, baseline and future. The baseline scenario represents Bureau of Meteorology climate datasets (1976–2005) prepared for Australia and used for climate projects [4]. The future scenario uses data from the moderate emission Representative Concentration Pathway (RCP) 6.0 scenario for the Global Climate Model (GCM) CSIRO Mk3 for the year 2055, sourced from Vanderwal [5]. Total 19 bioclimatic variables (Table 2) were obtained for each climate scenario.

To minimize multicollinearity, six of these predictors including, Isothermality (Bio3), Maximum temperature of the warmest month (Bio5), Mean temperature of coldest quarter (Bio11), Annual precipitation (Bio12), Precipitation of wettest quarter (Bio16) and Precipitation of driest quarter (Bio17) were selected for use in MaxEnt modelling (presented with bold in Table 2). The variable selection was based on the Pearson correlation coefficient ($r < 0.7$) and prior SDM studies involving the study species [14,15].

3.4. Assessing the change in species geographic distribution ranges

MaxEnt outputs were converted from ASCII to Tiff file format using open-source QGIS [16] for further analysis. To quantify the change in species range between the baseline and future scenario, change in area and the magnitude and direction of shift were calculated using the QGIS. A complete GIS workflow is provided in Fig. 1.

Both baseline and future outputs for each species were reprojected to WGS 84/UTM Zone 50 S to facilitate area level calculations. The raster cell size was set at 3000m $\times$ 3000 m. The cell

---

5 The set of variables presented here is not generated using BIOCLIM. The dataset is a collection of bioclimatic variables provided as spatial layers downscaled to 0.05 degrees (~5km resolution).
size was selected to represent 3 km apiary separation regulation as discussed in [1]. Reprojected raster layers were then converted to a binary presence-absence raster using **maximum training sensitivity plus specificity** logistic threshold [14,15,17]. Using the threshold values (Table 1 in [2]), each species output was reclassified where pixel values less than the determined threshold were reclassed to 0, and pixel values greater than the threshold were reclassed to 1 (for baseline scenario) or 2 (for the future scenario). The spatial overlap between baseline and future scenarios for each species was calculated as the mathematical addition of the reclassified grids. The total number of presence pixels for each class including ‘baseline only’ (pixel value = 1), ‘future only’ (pixel value = 2), and ‘baseline and future’ (pixel value = 3) were then multiplied by the cell size (9 km²) to calculate species distribution areas by class (Fig. 1).

The percentage change in the area of each species’ geographic range in the future was calculated as: [(RF/RB) – 1] *100. Where RF represents the distribution area in the future (sum of the area of pixel values 2 and 3) and RB is the distribution area at baseline (sum of the area of pixel values 1 and 3). To assess the shift in distributions, the mean center (latitude and longitude) of

Fig. 1. Workflow highlighting steps for range change analysis using QGIS
presence cells for each species was calculated for the baseline and future scenarios. The distance and compass directional shift between mean centers for each scenario was then calculated and together, the two metrics provided an indication of the magnitude and direction of changes to each species distribution following climate change.

Ethics Statements

This work involved human subjects for data collection. We confirm that the relevant informed consent was obtained from the participants and the research was carried out upon the University of Western Australia human ethics approval (RA/4/1/9247).

CRediT Author Statement

**Vidushi Patel**: Conceptualization, Methodology, Programming, Formal analysis, Data curation, Writing – original draft preparation; **Bryan Boruff**: Conceptualization, Methodology, Writing – review & Editing, Resources, Supervision, Funding acquisition; **Eloise Biggs**: Writing – review & editing, Visualization, Supervision; **Natasha Pauli**: Writing – review & editing, Supervision.

Declaration of Competing Interest

The authors declare that they have no known competing financial interests or personal relationships that could have appeared to influence the work reported in this paper.

Data Availability

Climate induced change in spatial distribution of bee forage species in Southwest Western Australia (Original data) (Mendeley Data).

Acknowledgments

The authors acknowledge financial and in-kind support from the Cooperative Research Centre for Honey Bee Products (CRC20160042) and Department of Industry, Science, Energy and Resources, and Bee Industry Council of Western Australia (BICWA). We also acknowledge all participant beekeepers for contributing their time and knowledge to this research. We further acknowledge Manita Narongsirikul for support during the data collection process. Some visual components of the graphical abstract were sourced from the Integration and Application Network (ian.umces.edu/media-library) provided for use through an Attribution-ShareAlike 4.0 International (CC BY-SA 4.0) licence.

References

[1] V. Patel, E. Biggs, N. Pauli, B. Boruff, Using a social-ecological system approach to enhance understanding of structural interconnectivities within the beekeeping industry for sustainable decision making, Ecol. Soc. 25 (2) (2020), doi:10.5751/ES-11639-250224.
[2] V. Patel, E. Biggs, N. Pauli, B. Boruff, Climate induced change in spatial distribution of bee forage species in Southwest Western Australia, Mendeley Data V3 (2022), doi:10.17632/vnztvrcrp.3.
[3] Atlas of Living Australia, 2021. https://spatialala.org.au. Accessed December 10, 2021.
[4] Bureau of Meteorology climate datasets, 2021. http://www.bom.gov.au/jsp/awap/. Accessed March 25, 2021.
[5] J. Vanderwal, All future climate layers for Australia - 5km resolution, James Cook University, 2012 doi.org/10.25903/ky81-xc32.
[6] W. Fithian, J. Elith, T. Hastie, D.A. Keith, Bias correction in species distribution models: pooling survey and collection data for multiple species, Method. Ecol. Evol. 6 (4) (2015) 424–438, doi:10.1111/2041-210X.12242.

[7] S.A. James, P.S. Soltis, L. Belbin, A.D. Chapman, G. Nelson, D.L. Paul, et al., Herbarium data: Global biodiversity and societal botanical needs for novel research, Appli. Plant Sci. 6 (2) (2018) e1024, doi:10.1002/aps3.1024.

[8] T.H. Booth, H.A. Nix, J.R. Busby, M.F. Hutchinson, bioclim: the first species distribution modelling package, its early applications and relevance to most current MaxEnt studies, Diver. Distribut. 20 (1) (2014) 1–9, doi:10.1111/ddi.12144.

[9] S.J. Phillips, R.P. Anderson, R.E. Schapire, Maximum entropy modeling of species geographic distributions, Ecolog. Model. 190 (3) (2006) 231–259, doi:10.1016/j.ecolmodel.2005.03.026.

[10] J. Elith, C.H. Graham*, R.P. Anderson, M. Dudík, S. Ferrier, A. Guisan, et al., Novel methods improve prediction of species’ distributions from occurrence data, Ecography 29 (2) (2006) 129–151.

[11] J. Elith, C.H. Graham, Do they? How do they? WHY do they differ? On finding reasons for differing performances of species distribution models, Ecography 32 (1) (2009) 66–77, doi:10.1111/j.1600-0587.2008.05505.x.

[12] S.J. Phillips, M. Dudík, Modeling of species distributions with Maxent: new extensions and a comprehensive evaluation, Ecography 31 (2) (2008) 161–175, doi:10.1111/j.0906-7590.2008.5203.x.

[13] T.H. Booth, Checking bioclimatic variables that combine temperature and precipitation data before their use in species distribution models, Austral Ecol. 47 (7) (2022) 1506–1514.

[14] C.E. Gonzalez-Orozco, L.J. Pollock, A.H. Thornhill, B.D. Mishler, N. Knerr, S. Laffan, et al., Phylogenetic approaches reveal biodiversity threats under climate change, Nat. Clim. Change 6 (12) (2016) 1110–+, doi:10.1038/Nclimate3126.

[15] J.J. Hamer, E.J. Veneklaas, P. Poot, K. Mokany, M. Renton, Shallow environmental gradients put inland species at risk: Insights and implications from predicting future distributions of Eucalyptus species in South Western Australia, Austral Ecol. 40 (8) (2015) 923–932, doi:10.1111/aec.12274.

[16] QGIS Development Team, (Version 3.10, 2019). QGIS Geographic Information System, Open Source Geospatial Foundation Project. http://qgis.osgeo.org.

[17] C. Liu, P.M. Berry, T.P. Dawson, R.G. Pearson, Selecting thresholds of occurrence in the prediction of species distributions, Ecography 28 (3) (2005) 385–393, doi:10.1111/j.0906-7590.2005.03957.x.