ORIGINAL PAPER

Citizen science reveals current distribution, predicted habitat suitability and resource requirements of the introduced African Carder Bee *Pseudoanthidium* (*Immanthidium*) *repetitum* in Australia

Keeley Dart · Tanya Latty · Aaron Greenville

Abstract The introduction of non-native bee species is a major driver of ecosystem change resulting in the spread of non-native weeds, alterations to plant-pollinator interactions and competition with native species for food and nesting resources. Our lack of ecological information for many non-native organisms hinders our ability to understand the impacts of species introductions. This is often compounded by the Wallacean Shortfall—a lack of adequate knowledge of a species' distribution in geographic space. In Australia, the African carder bee (*Pseudoanthidium* (*Immanthidium*) *repetitum*) was first observed in 2000 and has since become one of the most common bees in some regions. Despite its rapid population increase and range expansion, little is known about the ecology or distribution of *P. repetitum*. In this study, we determine the flower preferences, current distribution and predicted areas at risk of future invasion of *P. repetitum* using opportunistic data collected from citizen science websites, social media and museum records. We found that the current distribution of *P. repetitum* in Australia encompasses approximately 332,000 km^2 concentrated along the eastern coast. We found considerable suitable habitat outside the current distribution including biodiversity hotspots and world heritage listed natural areas. *Pseudoanthidium repetitum* foraged on a wide range of plants from many families and can thus be classified as a generalist forager (polylectic). Our results suggest that *P. repetitum* is well suited for continued expansion in coastal Australia. Our results demonstrate the effective application of opportunistic data in overcoming knowledge gaps in species ecology and modelling of introduced species distribution.

Introduction

The introduction of pollinator species outside their native range is a major driver of ecosystem change (Goulson 2003; Ward 2007). Negative impacts caused by non-native bee species include competition with native species for floral and nesting resources, spread of non-native weeds and potential "sleeper" weeds (weeds that could be invasive, but are yet to find effective pollinators in the environment) and disruption of plant-pollinator interactions potentially affecting the seed set of native plants (Goulson 2003; Ashcroft et al. 2012; New 2016). The negative effects of pollinator introductions therefore have the potential to alter local ecosystem

K. Dart · T. Latty (✉) · A. Greenville
School of Life and Environmental Sciences, University of Sydney, Sydney, Australia
e-mail: tanya.latty@sydney.edu.au

K. Dart · T. Latty · A. Greenville
Sydney Institute of Agriculture, Sydney, Australia

function by changing dynamic interactions between flora and fauna (Goulson 2003; Stout and Morales 2009; New 2016).

The study of introduced bees suffers from the classic Wallacean shortfall (sensu Lomolino 2004): a lack of knowledge about species distribution that hinders our ability to monitor and predict ongoing spread. Monitoring introduced species on a large scale is often cost prohibitive and consequently lacks adequate funding and personnel. The absence of coordinated survey efforts can result in disjointed information about the introduced organism's distribution and spread in the environment (Pusceddu et al. 2019).

For newly introduced species, knowledge of basic ecology can help identify suitable habitat near the points of entry and aid modelling efforts for predicting range expansion (Ashcroft et al. 2012). Understanding an organisms' requirements for resources such as food and nesting habitat are another key component of the management of introduced species (Ward 2007; Ashcroft et al. 2012; Roy-Dufresne et al. 2019; Eckert et al. 2020). Introduced bee species largely display polylectic (aka 'generalist') foraging behaviour, allowing them to adapt to new environments and establish outside their native range (Goulson 2003). Knowledge about an invasive bees' degree of floral specialisation can provide information about habitat suitability and inform the likelihood of competition with native species (Cane and Sipes 2006; Fortel et al. 2016). Although knowledge of the ecological requirements of non-native species is an important prerequisite for effective management, the professional collection of such data is time consuming and expensive.

Citizen science, defined here as "the engagement of volunteers in data collection or scientific research", provides an exciting opportunity for collecting data about the habitat requirements and distribution of introduced species at large geographic scales; these data can then be incorporated into monitoring programs (Silvertown 2009). Indeed, a review by Encarnação et al. (2021) found 126 studies where citizen science had been used to study biological invasions including the spread of Hemlock Wholly adelgids (*Adelges tsugae*) in the USA (Fitzpatrick et al. 2009), monitoring of invasive plants in Texas (Gallo and Waitt 2011), and detecting and monitoring the spread of European green crabs (*Carcinus maenas*) on the Pacific coast of the USA (Grason et al. 2018).

Traditional contributory approaches to citizen science projects typically involve training a number of non-expert volunteers who then actively collect information with support from scientists (Encarnação et al. 2021; Danielsen et al. 2020). Such projects can be logistically challenging as they may require training for volunteers. More recently, the rise of free biodiversity-focused citizen science sites such as iNaturalist and Project Noah have allowed scientists to recruit a larger number of volunteers who may or may not have specific training. For example, researchers tracking the spread of the invasive ladybeetle *Harmonia axyridis* in Argentina, encouraged the public to report sightings via iNaturalist, Facebook, Whatsapp, email or via a Google form. Such an approach has the advantage of lower time investment as volunteer 'training' occurs asynchronously via a website and brochures. However, researchers still need to recruit sufficient numbers of volunteers; this can be a significant limitation, particularly if the focal organism is geographically widespread.

In contrast, opportunistic approaches to citizen science collect information from content posted on social media (reviewed in Encarnação et al. 2021) or on biodiversity focused citizen science sites such as iNaturalist or Project Noah. Opportunistic data collection does not require participants to sign up to particular projects, rather, data is extracted from posts made on publicly viewable platforms. Data derived from opportunistic collection has been effective in mapping species distributions and at identifying the first observations of a species beyond their current distribution (Ward 2007; Werner et al. 2016; Maistrello et al. 2016; Zapponi et al. 2017; Walther and Kampen 2017; Pusceddu et al. 2019; Moulin 2020; Prendergast 2020). Opportunistic data collection potentially offers a powerful and logistically feasible tool for overcoming the Wallacean shortfall.

In this study, we use opportunistic data from a range of citizen science, social media and professionally curated biodiversity databases to investigate the distribution and plant preferences of the introduced African Carder Bee, *Pseudoanthidium* (*Immanthidium*) *repetitum* (previously *Afranthidium repetitum*) in Australia. *Pseudoanthidium repetitum* is an afro-tropical Megachilidae bee that was first reported in Australia in 2000 (Burwell and King 2000). In some years, *Pseudoanthidium repetitum* has been reported as one of the most common bees in Sydney

community gardens (Makinson et al. 2017). Despite the rapid spread of *P. repetitum*, there is no information available about the ecology of this species in either its native or introduced ranges. Indeed, as of December 2021, there have only been four studies of *P. repetitum,* none of which focus exclusively on foraging ecology or habitat requirements (Baumann et al. 2016; Litman et al. 2016; Makinson et al. 2017; Prendergast 2020).

Specifically we aimed to:

1. Determine the current distribution of *P. repetitum* in Australia.
2. Determine which areas in Australia have a suitable climate for *P. repetitum* and thus may be at risk for future invasions.
3. Determine which host plants *P. repetitum* uses for foraging (pollen and nectar) and nest material collection.

There have been no dedicated professional surveys or monitoring programs targeted at *P. repetitum* (Baumann et al. 2016). Monitoring of *P. repetitum* in Australia is complicated by the country's sheer size (7,692,024 km^2) which makes country-scale professional surveys infeasible. Thus, a secondary goal of our study was to assess the feasibility of using opportunistic data to track the spread of an invasive insect across a continent. To maximise the number of sightings, we chose to use a wide range of opportunistic data sources including social media, biodiversity-focused citizen science webpages, solicited email, and professionally curated biodiversity databases.

Materials and methods

Study species

Pseudoanthidium repetitum is native to South Africa, although its full distribution in Africa is currently unknown. The first official sighting of *P. repetitum* in Australia was in Brisbane, Queensland (North-east Australia) in 2000 (Burwell and King 2000). Established populations were observed in Sydney, New South Wales (south-east Australia) in 2008. By 2015, *P. repetitum* was being reported much further south in Melbourne, Victoria (South-east Australia) and had also expanded its northward range up to

Rockhampton, Queensland (North-east Australia) (Baumann et al. 2016). In 2019 and 2020 sightings were recorded from Western Australia, on the western coast of Australia, representing a major range exansion (Prendergast 2020; pers. comm., Widmer 2020).

Although little is known about the foraging preferences of *P. repetitum* in Australia, data from its native range suggests it may have a preference for collecting pollen and nectar from species in the Lamiaceae (mint family) (Gess and Gess 2007).

Pseudoanthidium repetitum is an above ground cavity nesting species that uses plant fibres to build cocoon-like structures in gregarious nests (Gess and Gess 2007). Appropriate plant fibres are therefore an important resource for nesting *P. repetitum*. To date, there are no records of fibre preferences for *P. repetitum*.

Data collection and verification

We extracted presence and location metadata for *P. repetitum* in its Australian range from photographs posted on biodiversity-focused citizen science sites such as iNaturalist (72 sightings), Bowerbird (51 sightings), Project Noah (1) and Canberra Nature Map (18 sightings), professionally curated databases such as Atlas of Living Australia (61 sightings) and the Global Biodiversity Information Facility (6), and social media sites such as Flickr (9) and Facebook (2; Insects of Perth, Amateur Entomology Australia). To be included in the dataset, sightings had to include an associated image or physical specimen for verification. Search terms across the platforms included the current '*Pseudoanthidium repetitum*' and previous '*Afranthidium repetitum*' taxonomic classifications (Litman et al. 2016).

We accessed all databases between 14 December 2019 to 31 May 2020.

We also created and advertised a project webpage (https://www.tanyalatty.com/citizen-science) and accepted sightings via a dedicated email address (africancarderbeeproject@gmail.com). Sightings collected through direct email (14 sightings) were required to have exchangeable image file format (EXIF) metadata containing location coordinates or, location coordinates obtained through google earth or a compass app.

All sightings of *P. repetitum* were verified by Dart. *Pseudoanthidium repetitum* is easily distinguished

from other Megachilidae present in Australia due to distinct integumental markings on the abdomen identifiable in images; native Megachilids have bands composed of fur, rather than tegument (Fig. 1). Only verified sightings were included in the final analysis. All sighting records included the location in latitude and longitude using World Geodetic System 1984 (WGS84), the date of sighting, family of flower/plant being visited and the behaviour displayed at the flower/plant.

Contact with museums and state government agencies as part of data collection resulted in the identification of four new specimens in a donated collection at the CSIRO Australian National Insect Collection, from Boonah, Queensland, March 2000. Unfortunately, these particular samples did not contain accurate coordinate information and were therefore not included in modelling. The specimens have since been catalogued (accession numbers at the Australian National Insect Collection for two females: 32-153860, 32-153861; two males: 32-153862, 32-153863).

Floral preferences and specialisation

Floral associations were established in 79 of the 140 sightings of *P. repetitum*. Plants were identified to the level of family by Dart. Where possible, the behaviours of the bee were classified as foraging (for pollen/nectar) or collecting fibre (for nesting). We classified a bee as 'foraging' when images showed the bee's body or tongue in contact with floral reproductive organs. Classification as 'fibre collection' required the bee to be in contact with a plant stem or leaf and to be visibly removing fibre with its mandibles or rolling fibres with its abdomen. We used photos showing *P. repetitum* foraging on flowers to classify the degree of flower specialisation following the percentages outlined in (Müller 1996):

- Oligolectic: > 95% of plant images belong to one family;
- Polylectic with a strong preference for one plant family: 70–95% of plant images belong to one family;
- Polylectic: no plant family is represented by greater than 70% of images.

Determining the current distribution of P. repetitum in Australia

The current distribution of *P. repetitum* was determined using alpha hulls. Alpha hulls are an adaptation of the minimum convex polygon method. Minimum convex polygons (convex hulls) have been used extensively as the internationally accepted standard in Extent of Occurrence (EOO) estimation for threatened species (Rivers et al. 2010; IUCN 2012). Minimum convex polygons produce the smallest polygon containing all sightings in which no internal angle exceeds 180°, meaning concave surfaces are prohibited on the polygon edge (Phillips et al. 2008). This can result in an overestimation
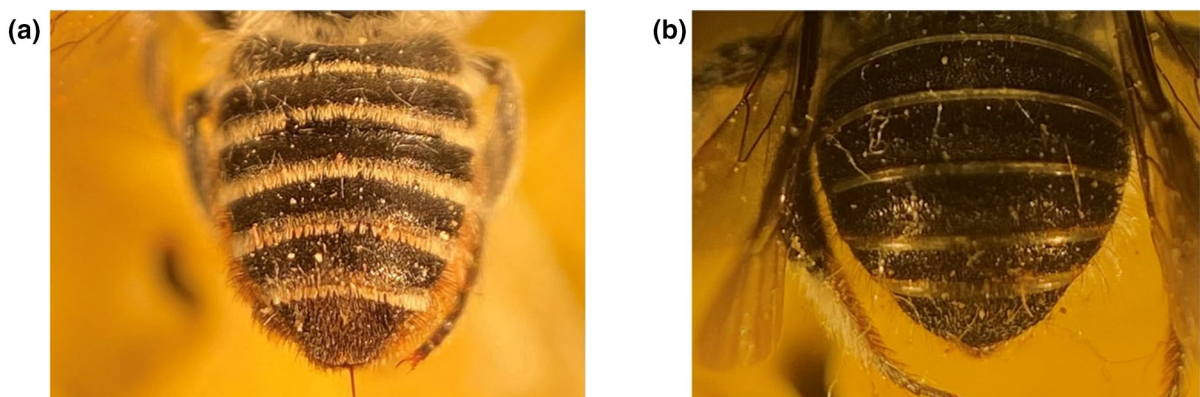
Fig. 1 **a** Shows the abdomen of a native Megachilid bee. Note that the light coloured bands consist of hairs, **b** the abdomen of a *P. repetitum*. Note that the light coloured bands are made from integument, rather than hair

of the EOO and is not robust enough for presence-only data (Burgman and Fox 2003; Phillips et al. 2008).

Alpha hulls, on the other hand, allow the production of disjoint polygons with concave edges (Burgman and Fox 2003; Phillips et al. 2008). The alpha hull is created using Delauney Triangulation between the points with no lines intersecting, maximising the minimum angle of all triangles in the triangulation. The average line length of each triangle is calculated and all lines greater than a multiple (alpha) of the average line length are deleted. This allows the production of disjoint polygons when there is a large area of uninhabitable space between, allowing for sub populations or multiple incursions when assessing non-native species introductions (Burgman and Fox 2003; Phillips et al. 2008; Rivers et al. 2010). By allowing concave edges and disjoint populations the method avoids over estimation of the distribution (Burgman and Fox 2003; Meyer et al. 2017). The use of alpha hulls is relatively unbiased (Phillips et al. 2008), and does not weight the boundary range by the density of points, thus avoiding sampling bias often present in citizen science datasets (Burgman and Fox 2003). The alpha hull is also more appropriate for handling small datasets when compared to Fixed and Adaptive Kernel methods, and is suitable for use in presence only datasets (Burgman and Fox 2003; Boyle et al. 2009).

To produce the current distribution polygons, we used the getDynamicAlphaHull function from the RangeBuilder package, version 1.5, in R, version 4.0.2 (Davis Rabosky et al. 2016; R Core Team 2020). The polygons were set to enclose a minimum of 95% of sightings. The alpha value was set to three (three times mean line length) (Burgman and Fox 2003). The CoordError function in the rangeBuilder package was used to determine the point accuracy of the sightings, and this was compared to any listed location accuracy buffer on the open source platforms from which the sightings were collected. The larger location accuracy error of the two sources was chosen as the buffer ($\pm$ 31 km) around the edge of the distribution polygon to ensure all points were included, despite location accuracy errors. The final area enclosed by the polygons was estimated using the area function in the raster package, version 3.1.5 (Hijmans 2020).

Predicting the climatic suitability for population spread

We assessed climatic suitability for *P. repetitum* using correlative bioclimatic species distribution models (SDMs). SDMs use the statistical correlation between climatic predictor variables and the focal species' current presence locations to predict areas of suitable habitat (Phillips et al. 2008). SDM techniques have been widely employed for invasive plants, mammals and insects, particularly when more detailed biological information, such as traits or floral preferences are lacking (Ward 2007; Ashcroft et al. 2012; Crall et al. 2015; Roy-Dufresne et al. 2019; Eckert et al. 2020; Silva et al. 2020).

We used all 140 verified sighting points to construct the SDM. To produce the model, we used three bioclimatic variables: maximum temperature in the warmest month (bio 5), minimum temperature in the coldest month (bio 6) and annual precipitation (bio 12). These variables were obtained from the World-Clim database 'current climate' at a 5-arc minute resolution, and clipped to the Australian extent (Fick and Hijmans 2017). Temperature is known to greatly affect ectothermic organisms such as insects, while precipitation can directly affect pollen and nectar resource distribution and availability (Ulrichs and Hopper 2008; Dixon et al. 2009; Lawson and Rands 2019). The model was produced using Maxent, version 3.4.0, through the Dismo package, version 1.1.4, in R (Hijmans et al. 2017). MaxEnt models have been used extensively in species distribution modelling, are compatible with presence-only datasets, and perform well with small dataset sizes (Phillips et al. 2006; Elith et al. 2011).

The point data were split into training (80%) and validation (20%) datasets. A bias file was created based on the sampling densities of *P. repetitum* and used to weight the selection of 10,000 pseudo-absence points (background points) to reduce the effect of geographic sampling bias that is common in citizen science data (Elith et al. 2006). Sampling densities represent changes in survey effort across the landscape and the probability of occurrence at a given location is assumed to be affected by survey effort (Elith et al. 2011; Merow et al. 2013). The model was evaluated using the validation dataset, producing the receiver operating characteristic (ROC) curve. The area under the ROC curve (AUC) was used as the test

statistic for model fit. The ROC is produced by plotting the true positive rate against the false positive rate, with high true positive at low false positive rates resulting in a better fit (Graham and Hijmans 2006). The AUC is a value between 0 and 1 where 0.5 indicates that the model is no better than random, < 0.5 the model is worse than random, and 1 indicating the model perfectly discriminates between presence and absence for the species (Graham and Hijmans 2006; Elith et al. 2011).

The Maxent model output is a continuous surface from 0 to 1 where zero is unsuitable and the closer to one the more likely the species occurs at that location. The model output was then classified into five equal categories 0–0.2, 0.2–0.4, 0.4–0.6, 0.6–0.8 and 0.8–1.0, corresponding to "unsuitable", "low", "moderate", "good" and "high" suitability respectively.

Results

Floral specialisation

Of our 140 verified sightings, 79 contained images of plants that could be identified to the Family level. Of the 79 floral associations, 70 were classified as foraging for pollen or nectar and the remaining 9 were associated with fibre collection (Fig. 2). Based on

these flower associations, we classified the level of floral specialisation displayed by *P. repetitum* as polylectic. The greatest proportion of visitation to any one plant family for pollen or nectar was Lamiaceae (60.0%) followed by Asteraceae (32.9%) (Fig. 2). In contrast, fibre collection was constrained to only two plant families, Lamiaceae and Asteraceae (22.2% and 77.8%, respectively; Fig. 2).
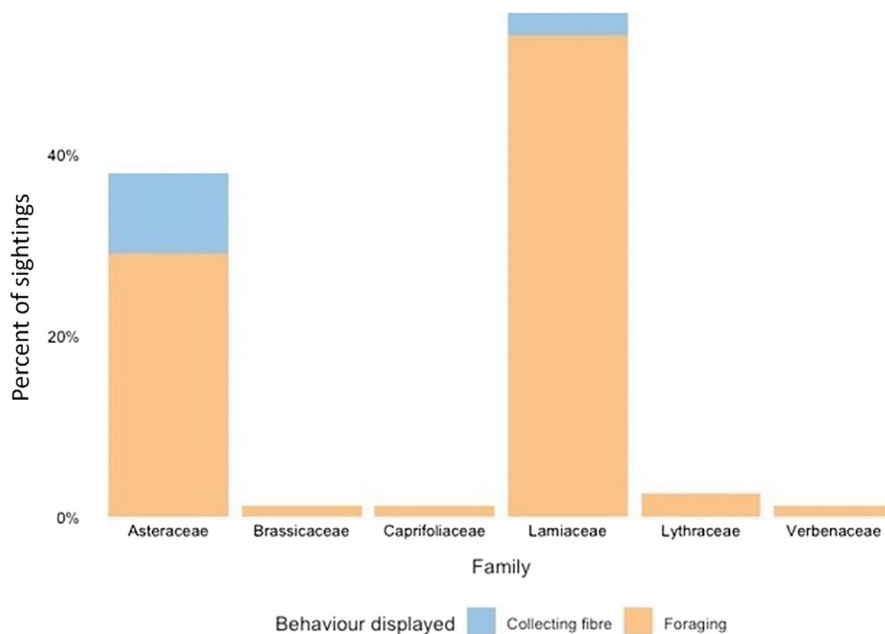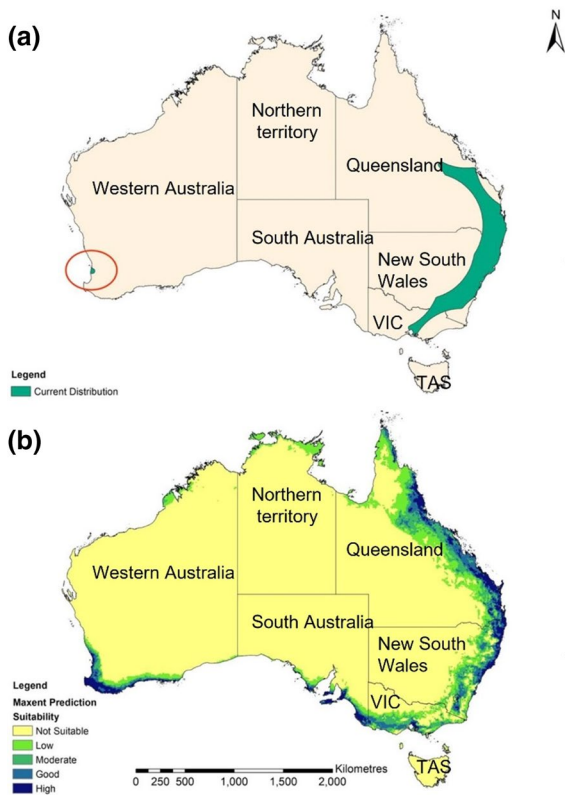
Current distribution

The current distribution of *P. repetitum* consisted of two established, disjunct populations capturing 139 of the 140 sightings. A single South Australian sighting fell outside the current distribution polygons. The two polygons enclose a total area of approximately 332,484 km^2. The distribution shows a population in the eastern states spanning Queensland, New South Wales, the Australian Capital Territory and Victoria and a second population in Western Australia established in 2019 (Fig. 3a).

Predicted climatic suitability

All three bioclimatic variables used to model environmental suitability had high levels of variable importance; the model had a high ability to differentiate between an areas relative importance (AUC = 0.942)

Fig. 2 Floral preferences of *Pseudoanthidium repetitum* in foraging and fibre collection. Proportion is given as a percentage of all plant visitation associated images from a total of 79 associated sightings

Fig. 3 Current distribution and predicted habitat suitability of *Pseudoanthidium repetitum* in Australia. **a** Current distribution of *P. repetitum* in Australia across two disjoint polygons. The red circle highlights the small polygon in Western Australia. State names are in black text; Victoria and Tasmania have been abbreviated to 'VIC' and 'TAS', respectively. **b** The predicted climatic suitability for the future range expansion of a *P. repetitum* in Australia based on maximum temperature in the warmest month, minimum temperature in the coldest month and annual precipitation

Table 1 Australian land area occupied by each defined climatic suitability category

Suitability classification	Approximate area (km^2)
Not suitable	6,640,105
Low	452,887
Moderate	247,622
Good	178,553
High	133,805

140 sightings of *P. repetitum* in Australia, including 79 images that contained identifiable plant species; these data allowed us to identify *P. repetitium*'s current range and to predict the location of suitable—but currently uninvaded-habitat. We were further able to infer that *P. repetitum* is a broadly generalist forager, collecting pollen and nectar from several plant families. We argue that opportunistic biodiversity data in the form of publicly shared photographs are a potentially rich source of information, particularly when professional survey efforts are logistically infeasible. For flower visiting organisms, photographs may also shed light on plant-animal interactions.

We found that *P. repetitum* is polylectic, using a broad range of flowers for nectar and pollen. The majority of sightings were observed on the plant family Lamiaceae (60.0%) followed by Asteraceae (32.9%). In South Africa, closely related bees in the tribe Anthidiini also show a high degree of polylecty with Asteraceae being a highly-used resource for pollen and nectar (Struck 1994; Gess and Gess 2007).

Since we do not have direct data on the abundance of Lamiaceae and Asteraceae in the sampled Australian environments, we cannot conclusively demonstrate that *P. repetitum* has a preference for either plant family, since the observed associations could be due to high abundance of both plant families. Anecdotally, both Asteraceae and Lamiaceae are commonly cultivated in urban Sydney, Australia (pers. comm., Latty 2020); our results could therefore reflect the high abundance of those two species. Further, the associations we observed may partially reflect the tendency of observers to photograph particularly attractive or conspicuous flower species. Targeted observations and experiments are thus needed to determine if *P. repetitum* truly has a preference for flowers of Asteraceae and Lamiaceae.

(Elith et al. 2006). The predicted climatic suitability model shows suitable climate present along most of the Australian east coast (Fig. 3b). The island-state of Tasmania is unlikely to be colonised as we did not identify any suitable habitat. The current distribution of *P. repetitum* on both sides of the continent has the potential to expand further along coastal Australia (Fig. 3b, Table 1).

Discussion

By combining opportunistic sightings from a range of publicly available platforms, we were able to collect

Nevertheless, we have shown that both families likely provide important foraging resources for *P. repetitum* in Australia.

Floral specialisation in bees is often classified based on the number of plant species bees collect pollen from (Cane and Sipes 2006; Ritchie et al. 2016). We were unable to distinguish pollen or nectar foraging from photos and so have inferred foraging specialisation based on general foraging rather than pollen-specific foraging. Nevertheless, the large range of flower visitation we observed strongly argues that *P. repetitum* is indeed a generalist species that visits a wide range of flower species across several plant families. In future, researchers could take note of whether foraging *P. repetitum* are harvesting nectar or pollen, or they could sample pollen directly from the bodies or brood cells of *P. repetitum*.

Members of the bee family Megachilidae, to which *P. repetitum* belongs, are heavily dependent on non-floral plant resources for nesting and protection (Requier and Leonhardt 2020). In our study, *P. repetitum* was only observed collecting nesting fibres from Asteraceae (77.8%) and Lamiaceae (22.2%)

suggesting that these two families are an important source of nesting resources. As with flowers used for foraging, we do not know whether the focus on these two plant families is due to their abundance in the sampled environments, sampling bias by citizen scientists or due to actual preference on the part of *P. repetitum*. However, it is striking that of the six plant families *P. repetitum* foraged on for nectar and/or pollen, only two families were used for fibre collection. Our results tentatively suggest that *P. repetitum* has a narrower host range when foraging for nesting materials.

Although *P. repetitum* appears to use a limited number of plant families for the collection of building materials, this specialisation is unlikely to limit the spread of *P. repetitum* as both plant families are common across Australia and include many widely cultivated species. It's possible that *P. repetitum*'s spread will be limited by access to appropriate nesting cavities; however, the frequent finding of *P. repetitum* nests inside metre boxes (Gess and Gess 2007; see Fig. 4) suggests it is a highly adaptable nester. Future field research is needed to develop a
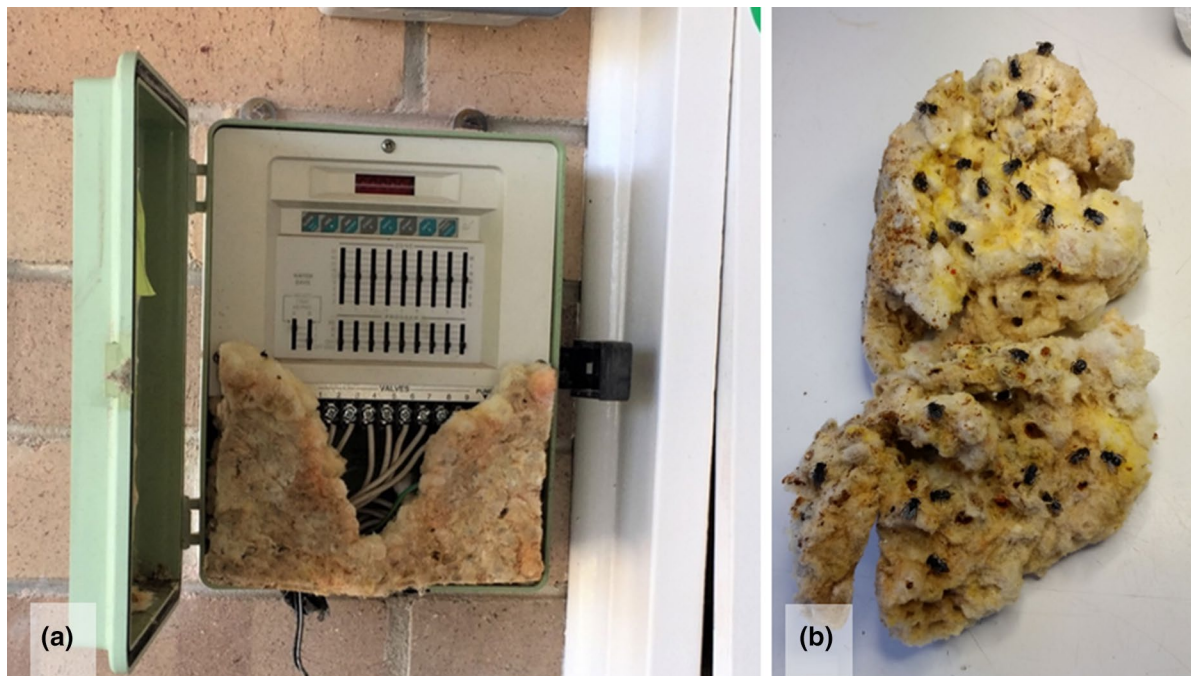
Fig. 4 a Nest of *Pseudoanthidium repetitum* found inside a metre-box in Halls Head, Western Australia, 2019. **b** Nest extracted from a metre-box in Mandurah, Western Australia, 2019. Photos obtained from Marc Widmer, Department of Primary Industries and Regional Development

clearer understanding of *P. repetitum*'s nesting biology with particular focus on its use of non-floral plant resources.

The current extent of *P. repetitum* in Australia is estimated at 332,484 km^2 divided into two disjunct subpopulations on either side of the continent. The current distribution includes 99.2% of sightings, excluding only one sighting in South Australia. The sighting in South Australia was made in 2015, with no further confirmed sightings of the species since (pers. comm., Parslow 2020). One point location does not infer establishment and could be the result of a population that failed to establish or a population density too low for detectability. Ground truthing the sighting locations used in this study would provide further information about the establishment of the species in areas highlighted by citizen science sightings.

The flight distance of a species limits its ability for range expansion such that shorter flight range decreases dispersal (Rasmussen and Cameron 2010). The buffer around the verified sightings was set to 31 km based on the location accuracy of included sightings. If dispersal is restricted by flight distance, then flight distance would potentially be a more ecologically appropriate buffer. However, empirical data on the flight range of *P. repetitum* is currently lacking. In addition, the act of range expansion itself may select for increased flight distance and thus may not remain temporally static (Therry et al. 2014; Merwin 2019). The observation that *P. repetitum* often nests in artificial cavities such as metre boxes (Fig. 4) and window frames (Latty pers obs) raises the possibility that *P. repetitum* could expand its range via unintended human-mediated dispersal, for example, if concealed nests are accidentally transported to new areas. Indeed, it has been suggested that the recent range expansion of *P. repetitum* into Western Australia (which would require movement across the large and inhospitable Nullaboor plain), could have been the result of inadvertent human-mediated dispersal (Prendergast 2020).

Using correlative bioclimatic modelling, we determined that there is significant suitable area for future range expansion within Australia. Of the total area classified as both good and high suitability, 222,105 km^2 is outside *P. repetitum's* current distribution. In contrast > 57,000 km^2 of the current distribution is classified by our model as unsuitable. Despite the predicted unsuitability of the area, some sightings correspond with these locations. These sightings may be from established populations, or alternatively, represent unsuccessful exploration into new habitat. Systematic surveys are needed to determine if *P. repetitum* has indeed colonised areas we have classified as unsuitable, and if so, if there are bioclimatic factors that have allowed establishment. Future survey efforts should also be targeted at the regions identified as suitable outside the current range to identify if the *P. repetitum* population extends beyond the range identified from opportunistic sightings.

Species distribution models assume that the species presence locations represent equilibrium with its environment (Elith et al. 2010). *Pseudoanthidium repetitum* was first reported in 2000, with first entry into Australia likely prior to March 1999, and the species has continued to increase its distribution. It is therefore unlikely *P. repetitum* has reached equilibrium with the Australian environment. Nevertheless, our modelling represents an important first step toward a better understanding of *P. repetitum*'s distribution in Australia. Mechanistic SDMs that address trait variation (either evolved or plastic) by incorporating functional species traits with climatic variation (Kolbe et al. 2010), may be used to further refine distribution models for *P. repetitum*. However, developing these mechanistic models will require a much better understanding of *P. repetitum*'s ecology.

Given its rapid spread and the fact that a large proportion of coastal Australia appears to contain suitable habitat, we strongly recommend research on the potential interactions of *P. repetitum* with native flora and fauna. It is important to note that only a small fraction of introduced species ultimately cause significant environmental or economic harm (Schlaepfer et al. 2011). There is currently no evidence that *P. repetitum* will have a negative impact on Australian biota. Interesting, Makinson et al. (2017) noted a positive relationship between *P. repetitum* abundance and native bees, suggesting that *P. repetitum* is not excluding other species from foraging resources. However, Makinson et al. (2017) was not specifically designed to test for competitive interactions between *P. repetitum* and native flower-visitors and so may have missed more subtle impacts.

Although we controlled for location survey bias in this study, since data collection was opportunistic, we necessarily could not entirely control for collection biases; this limited our ability to address

some types of questions. For example, the majority of sightings came from urban areas, which could reflect *P. repetitum*'s affinity for urban areas, or the simple fact that relatively few users posted photographs from natural areas. Similarly, the plant preferences we observed could be influenced by under sampling of inaccessible flowering tree canopies (a major source of nectar and pollen in the Australian context). Despite these drawbacks, we argue that the use of citizen science data provides a useful tool for preliminary investigations, particularly in cases where sampling is required over a large area. Our results will allow us to prioritise professional sampling efforts, for example, in climatically suitable areas at the edges of *P. repetitum*'s current range. Our study also uncovered interesting results that can be formally addressed via future research, for example, determining if the range of non-floral resources used as building materials is indeed constrained to a small number of families.

A major benefit of using citizen science data is the ability to sample a large geographic range. A professional survey across Australia would have been prohibitively expensive and time consuming. For this reason, citizen science has been used to determine the distribution of a number of species, including introduced species (Pusceddu et al. 2019; Encarnação et al. 2021). For small inconspicuous invertebrate species, opportunistic citizen science data has proven especially effective at identifying species distribution (Ward 2007; Zapponi et al. 2017; ElQadi et al. 2017; Pusceddu et al. 2019). The approach taken in this study further demonstrates that opportunistic sighting data from open source platforms can be a useful tool for collecting valuable ecological data on introduced insects.

Declarations

References

Ashcroft MB, Gollan JR, Batley M (2012) Combining citizen science, bioclimatic envelope models and observed habitat preferences to determine the distribution of an inconspicuous, recently detected introduced bee (*Halictus smaragdulus* Vachal Hymenoptera: Halictidae) in Australia. Biol Invasions 14:515–527. https://doi.org/10.1007/s10530-011-0092-x

Baumann JM, Walker K, Threlfall C, Williams NS (2016) African Carder bee, *Afranthidium* (*Immanthidium*) *repetitum* (Hymenoptera: Megachilidae): a new exotic species for Victoria. Vic Nat 133:21–25

Boyle SA, Lourenço WC, da Silva LR, Smith AT (2009) Home range estimates vary with sample size and methods. Folia Primatol 80:33–42. https://doi.org/10.1159/000201092

Burgman MA, Fox JC (2003) Bias in species range estimates from minimum convex polygons: implications for conservation and options for improved planning. Anim Conserv 6:19–28. https://doi.org/10.1017/S1367943003003044

Burwell C, King J (2000) News from the Queensland Museum. Entomol Soc Qld 28:25.

Cane JH, Sipes S (2006) Characterizing floral specialization by bees: analytical methods and a revised lexicon for oligolecty. In: Plant-pollinator interactions: from specialization to generalization. pp 99–122 University of Chicago Press, Chicago

Crall AW, Jarnevich CS, Young NE et al (2015) Citizen science contributes to our knowledge of invasive plant species

distributions. Biol Invasions 17:2415–2427. https://doi.org/10.1007/s10530-015-0885-4

Danielsen F, Adrian T, Jensen PM, Muñoz J, Shirk JL, Burgess ND (2020) THREE. Current approaches to citizen science. In: Handbook of citizen science in ecology and conservation. University of California Press, Berkeley, pp 25–30

Davis Rabosky AR, Cox CL, Rabosky DL et al (2016) Coral snakes predict the evolution of mimicry across New World snakes. Nat Commun 7:1–9. https://doi.org/10.1038/ncomms11484

Dixon AF, Honěk A, Keil P et al (2009) Relationship between the minimum and maximum temperature thresholds for development in insects. Funct Ecol 23:257–264

Eckert S, Hamad A, Kilawe CJ et al (2020) Niche change analysis as a tool to inform management of two invasive species in Eastern Africa. Ecosphere. https://doi.org/10.1002/ecs2.2987

Elith J, Graham CH, Anderson RP et al (2006) Novel methods improve prediction of species' distributions from occurrence data. Ecogr 29:129–151

Elith J, Kearney M, Phillips S (2010) The art of modelling range-shifting species. Methods Ecol Evol 1:330–342. https://doi.org/10.1111/j.2041-210x.2010.00036.x

Elith J, Phillips SJ, Hastie T et al (2011) A statistical explanation of MaxEnt for ecologists. Divers Distrib 17:43–57. https://doi.org/10.1111/j.1472-4642.2010.00725.x

ElQadi MM, Dorin A, Dyer A et al (2017) Mapping species distributions with social media geo-tagged images: case studies of bees and flowering plants in Australia. Eco Inform 39:23–31. https://doi.org/10.1016/j.ecoinf.2017.02.006

Encarnação J, Teodósio MA, Morais P (2021) Citizen science and biological invasions: a review. Front Environ Sci 8:303

Fick SE, Hijmans RJ (2017) WorldClim 2: new 1-km spatial resolution climate surfaces for global land areas. Int J Climatol 37:4302–4315. https://doi.org/10.1002/joc.5086

Fitzpatrick MC, Preisser EL, Ellison AM, Elkinton JS (2009) Observer bias and the detection of low-density populations. Ecol Appl 19(7):1673–1679

Fortel L, Henry M, Guilbaud L et al (2016) Use of human-made nesting structures by wild bees in an urban environment. J Insect Conserv 20:239–253. https://doi.org/10.1007/s10841-016-9857-y

Gallo T, Waitt D (2011) Creating a successful citizen science model to detect and report invasive species. BioScience 61(6):459–465

Gess SK, Gess FW (2007) Notes on nesting and flower visiting of some Anthidiine bees (Hymenoptera: Megachilidae: Megachilinae: Anthidiini) in Southern Africa. J Hymenopt Res 16:30–50

Goulson D (2003) Effects of introduced bees on native ecosystems. Annu Rev Ecol Evol Syst 34:1–26

Graham CH, Hijmans RJ (2006) A comparison of methods for mapping species ranges and species richness. Glob Ecol Biogeogr 15:578–587. https://doi.org/10.1111/j.1466-822x.2006.00257.x

Grason EW, McDonald PS, Adams J, Litle K, Apple JK, Pleus A (2018) Citizen science program detects range expansion of the globally invasive European green crab in Washington State (USA). Manag Biol Invasions 9(1):39

Hijmans RJ (2020) raster: geographic data analysis and modeling. R package version 3:1–5. http://cran.rproject.org/package=raster.

Hijmans RJ, Phillips S, Leathwick J, Maintainer JE (2017) Package "dismo" type package title species distribution modeling. R package version 1:1–4. https://CRAN.R-project.org/package=dismo

IUCN (2012) IUCN Red List categories and criteria: version 3.1, 2nd edn. IUCN, Gland, Cambridge

Kolbe JJ, Kearney M, Shine R (2010) Modeling the consequences of thermal trait variation for the cane toad invasion of Australia. Ecol Appl 20:2273–2285

Lawson DA, Rands SA (2019) The effects of rainfall on plant–pollinator interactions. Arthropod Plant Interact 13:561–569

Litman JR, Griswold T, Danforth BN (2016) Phylogenetic systematics and a revised generic classification of anthidiine bees (Hymenoptera: Megachilidae). Mol Phylogenet Evol 100:183–198. https://doi.org/10.1016/j.ympev.2016.03.018

Lomolino MV (2004) Conservation biogeography. In: Lomolino MV, Heaney LR (eds) Frontiers of biogeography: new directions in the geography of nature. Sinauer, Sunderland

Maistrello L, Dioli P, Bariselli M et al (2016) Citizen science and early detection of invasive species: phenology of first occurrences of *Halyomorpha halys* in Southern Europe. Biol Invasions 18:3109–3116. https://doi.org/10.1007/s10530-016-1217-z

Makinson JC, Threlfall CG, Latty T (2017) Bee-friendly community gardens: impact of environmental variables on the richness and abundance of exotic and native bees. Urban Ecosyst 20:463–476. https://doi.org/10.1007/s11252-016-0607-4

Merow C, Smith MJ, Silander Jr JA (2013) A practical guide to MaxEnt for modeling species' distributions: what it does, and why inputs and settings matter. Ecogr 36(10):1058–1069

Merwin AC (2019) Flight capacity increases then declines from the core to the margins of an invasive species' range. Biol Lett. https://doi.org/10.1098/rsbl.2019.0496

Meyer L, Diniz-Filho JAF, Lohmann LG (2017) A comparison of hull methods for estimating species ranges and richness maps. Plant Ecol Divers 10:389–401. https://doi.org/10.1080/17550874.2018.1425505

Moulin N (2020) When Citizen Science highlights alien invasive species in France: the case of Indochina mantis, *Hierodula patellifera* (Insecta, Mantodea, Mantidae). Biodivers Data J. https://doi.org/10.3897/bdj.8.e46989

Müller A (1996) Host-plant specialization in western Palearctic Anthidiine bees (Hymenoptera: Apoidea: Megachilidae). Ecol Monogr 66:235–257

New TR (2016) Alien species and insect conservation. Springer International Publishing, Cham

Phillips SJ, Anderson RP, Schapire RE (2006) Maximum entropy modeling of species geographic distributions. Ecol Model 190:231–259. https://doi.org/10.1016/j.ecolmodel.2005.03.026

Phillips BL, Chipperfield JD, Kearney MR (2008) The toad ahead: challenges of modelling the range and spread of an invasive species. Wildl Res 35:222–234

Prendergast KS (2020) First records of the introduced African carder bee, *Pseudoanthidium* (*Immanthidium*) *repetitum* (Hymenoptera: Megachilidae), in Western Australia. Pac Conserv Biol. https://doi.org/10.1071/PC20041

Pusceddu M, Floris I, Mannu R et al (2019) Using verified citizen science as a tool for monitoring the European hornet (*Vespa crabro*) in the Island of Sardinia (Italy). NeoBiota 50:97–108. https://doi.org/10.3897/neobiota.50.37587

R Core Team (2020) R: a language and environment for statistical computing. R Foundation for Statistical Computing, Vienna, Austria. URL https://www.R-project.org/.

Rasmussen C, Cameron SA (2010) Global stingless bee phylogeny supports ancient divergence, vicariance, and long distance dispersal. Biol J Lin Soc 99:206–232

Requier F, Leonhardt SD (2020) Beyond flowers: including non-floral resources in bee conservation schemes. J Insect Conserv 24:5–16

Ritchie AD, Ruppel R, Jha S (2016) Generalist behavior describes pollen foraging for perceived Oligolectic and Polylectic bees. Environ Entomol 45:909–919. https://doi.org/10.1093/ee/nvw032

Rivers MC, Bachman SP, Meagher TR et al (2010) Subpopulations, locations and fragmentation: applying IUCN Red List criteria to herbarium specimen data. Biodivers Conserv 19:2071–2085. https://doi.org/10.1007/s10531-010-9826-9

Roy-Dufresne E, Saltré F, Cooke BD et al (2019) Modeling the distribution of a wide-ranging invasive species using the sampling efforts of expert and citizen scientists. Ecol Evol 9:11053–11063. https://doi.org/10.1002/ece3.5609

Schlaepfer MA, Sax DF, Olden JD (2011) The potential conservation value of non-native species. Conserv Biol 25(3):428–437

Silva DP, Hall HG, Ascher JS (2020) Predicting the distribution range of a recently described, habitat specialist bee. J Insect Conserv 24:671–680. https://doi.org/10.1007/s10841-020-00241-3

Silvertown J (2009) A new dawn for citizen science. Trends Ecol Evol 24:467–471

Stout JC, Morales CL (2009) Ecological impacts of invasive alien species on bees. Apidologie 40:388–409

Struck M (1994) Flowers and their insect visitors in the arid winter rainfall region of southern Africa: observations on permanent plots. Insect visitation behaviour. J Arid Environ 28:51–74

Therry L, Gyulavári HA, Schillewaert S et al (2014) Integrating large-scale geographic patterns in flight morphology, flight characteristics and sexual selection in a range-expanding damselfly. Ecography 37:1012–1021. https://doi.org/10.1111/ecog.00630

Ulrichs C, Hopper KR (2008) Predicting insect distributions from climate and habitat data. BioControl 53:881–894. https://doi.org/10.1007/s10526-007-9143-8

Walther D, Kampen H (2017) The citizen science project "Mueckenatlas" helps monitor the distribution and spread of invasive mosquito species in Germany. J Med Entomol 54:1790–1794. https://doi.org/10.1093/jme/tjx166

Ward DF (2007) Modelling the potential geographic distribution of invasive ant species in New Zealand. Biol Invasions 9:723–735. https://doi.org/10.1007/s10530-006-9072-y

Werner D, Zielke DE, Kampen H (2016) First record of *Aedes koreicus* (Diptera: Culicidae) in Germany. Parasitol Res 115:1331–1334. https://doi.org/10.1007/s00436-015-4848-6

Zapponi L, Cini A, Bardiani M et al (2017) Citizen science data as an efficient tool for mapping protected saproxylic beetles. Biol Conserv 208:139–145. https://doi.org/10.1016/j.biocon.2016.04.035