The influence of data source and species distribution modelling method on spatial conservation priorities

William La Marca¹ | Jane Elith¹ | Ronald S. C. Firth²,³ | Brett P. Murphy⁴ | Tracey J. Regan¹,⁵ | John C. Z. Woinarski⁴ | Emily Nicholson⁶

Abstract

Aim: Species distribution models are an important conservation tool; however, performance can vary with factors including data inputs and modelling method. Model outputs are often under-evaluated for explanatory and predictive capacity. Our aim was to evaluate the capacity of existing data for seven small mammal species to provide useful inferences for management planning.

Location: Bathurst and Melville (collectively the Tiwi) Islands, Northern Territory, Australia.

Methods: We developed species distribution models (SDMs) with generalized linear models (GLMs) and boosted regression trees (BRTs) using survey data (351 sites) of small mammals, with two sets of environmental predictors: (a) field-study measurements and (b) available remotely sensed rasters. Predictive capacity of models was evaluated using percentage of deviance explained (%DE) and area under the receiver operating characteristic curve (AUC). We used Marxan to evaluate the influence of different model and data types as input for identifying spatial priorities.

Results: Field-informed SDMs performed well across both modelling methods, with relatively high test AUC values (mean = 0.82, range = 0.64–0.97) and test %DE (mean = 22.5%, range = 3.5%–65.8%). Remotely sensed models performed relatively poorly, with lower test AUC values (mean = 0.7, range = 0.56–0.86) and lower test %DE (mean = 8.9%, range = 0.03%–24.9%). A notable exception was remotely sensed models for Melomys burtoni (AUC = 0.85 & 0.86, %DE = 23.3% & 24.9%, Bathurst and Melville respectively). Marxan site irreplaceability rankings demonstrated low to marginal agreement using field-informed and remotely sensed inputs (Pearson correlation coefficient = 0.3), and similarly, using GLM and BRT model inputs (0.29).

Main conclusions: The occurrence of small mammals on the Tiwi Islands can be reasonably explained with field-informed variables, but not with remotely sensed alternatives. Different models lead to different conservation priorities. Our work emphasizes the importance of thoroughly testing SDMs prior to decision-making.
1 | INTRODUCTION

Species distribution models (SDMs) are developed by combining records of species occurrence or abundance with relevant environmental variables (Franklin, 2010). Information on species distributions can inform many ecological and conservation management questions, including assessments of biodiversity values and identification of priority areas (Guisan et al., 2013), by providing continuous estimates of suitability for a given species across geographic space (Guisan & Zimmerman, 2000). The accuracy of SDM predictions and their ability to reliably inform management decisions depend heavily on the quality of data for both species occurrences and environmental variables (Barry & Elith, 2006; Guillaumette et al., 2015), and on the predictability of the target system (Morán-Ordóñez, Lahoz-Monfort, Elith, & Wintle, 2017).

Several components of the construction of SDMs can affect the quality and certainty of model predictions (Barry & Elith, 2006; Elith, Burgman, & Regan, 2002; Ferrier, 2002; Heikkinnen et al., 2006; Macnally, 2000; Stockwell & Peterson, 2002; Wilson, Westphal, Possingham, & Elith, 2005). These include the following (a) the number and accuracy of occurrence records; (b) the accuracy of environmental data and their ability to accurately capture environmental processes which are functionally relevant to the target species; (c) modelling method, model parameterisation and model selection; and (d) model evaluation. Further uncertainties also exist when considering the spatial ecology and dynamism of the study species, as some species have relatively stable distributions which may be amenable to modelling, while others may be more fluid and temporally unstable, making modelling more challenging. Due to these potentially large uncertainties surrounding SDM predictions, it is imperative that proper consideration is given to the practical utility of model predictions and their ability to inform management decisions with reliability and confidence. The predictor variables most commonly used in SDMs for management applications are remotely sensed or modelled, as these provide continuous surfaces across the landscape capable of generating mapped predictions. These may be distal predictors which are correlated, though often only weakly, with proximal drivers (Austin, 2002). While it is desirable to predict the distribution of biotic entities using causal ecological parameters (i.e., proximal variables), distal variables are more commonly available (Guisan & Zimmerman, 2000). Field surveys can provide the opportunity to identify and measure predictors thought to be directly associated with a species' habitat requirements (e.g., the presence or absence of hollow logs which act as nesting hollows for small mammals). However, they can only be measured at the visited sites and cannot readily be projected beyond sites that have been surveyed.

One of the most widespread applications of SDMs is in conservation planning. Conservation planning typically considers multiple species (Kremen et al., 2008; Kujala, Lahoz-Monfort, Elith, & Moilanen, 2018; Leathwick et al., 2008; Linke, Pressey, Bailey, & Norris, 2007), and the influence of uncertainty on spatial priorities when combining multiple SDM predictions is not always clear and can be difficult to predict a priori. When SDMs are used in conservation planning to identify priority areas for protection, planning outputs can be highly sensitive to the input model predictions (Wilson et al., 2005).

In this study, we ask how data type and modelling method will affect the predictive performance of species distribution models, and how these different input types will influence spatial conservation priorities for multiple species. We explore these questions through a case study on the Tiwi Islands, the largest islands off the northern Australian coast (spatial extent ~8,000 km²). We evaluate the capacity of an existing survey dataset of species occurrence, in conjunction with two sets of predictor variables (one predominantly field-informed, the other completely remotely sensed) to predict distributions of small mammals, and to provide reliable information for conservation planning. The Tiwi Islands make a useful and relevant case study because of the availability of biodiversity data and both remotely sensed and field-informed covariates for a number of small mammal species. Similar to other islands (Burbidge & Manly, 2002; Burbidge, Williams, & Abbott, 1997), the Tiwi Islands appear to provide a safe haven for many species—including the small mammals modelled here—that have declined or disappeared on the mainland (Woinarski, Burbidge, & Harrison, 2015). However, recent evidence suggests that Tiwi populations of some species may be declining (Davies et al., 2018, 2017), and conservation efforts could benefit greatly from reliable predictions of species distributions to inform ongoing management planning. The Tiwi Islands are a good example of a typical scenario where modelling output could be useful, given recent silvicultural developments on the islands, and we explore through careful model building and evaluation how well the available data allow us to model the distribution of species of conservation concern and identify priority areas for conservation.

2 | METHODS

2.1 | Study area

The Tiwi Islands are located 80 km north of Darwin (11°36’S, 130°49’E) in the Northern Territory, Australia (Figure 1). The two largest islands, Bathurst Island (2,600 km²) and Melville Island
(5,786 km²), are separated by the Aspley Strait, which is on average 1.5 km wide. At the time of sampling for the mammal data described here, the islands supported relatively abundant and widespread populations of small- to medium-sized (<5 kg) native mammals in contrast to mainland Australia, where many mammal species are substantially reduced in both distribution and abundance (Firth et al., 2006). The islands are home to a tropical monsoonal climate with a significant rainfall gradient, from 1,400 mm in the east to 2,000 mm in the north-west. Eucalypt open forests are the dominant vegetation type, covering ~77% of the total land mass. These forests are composed largely of Darwin stringybark (Eucalyptus tetrodonta), Darwin woollybutt (Eucalyptus miniata) and Melville Island bloodwood (Corymbia nesophila). Perennial and annual grass species typically dominate the understorey; however, their composition is highly variable, shaped by factors such as the quality and type of soils, the impacts of fire, and relative landscape position. Coastal areas offer a diverse mosaic of vegetation types, the most extensive being mangrove forests, which extend inland along tidal watercourses (Messel, Wells, & Green, 1979).

2.2 | Species-occurrence data

We used species occurrence records derived from a systematic survey conducted for mammal abundance across Bathurst (115 survey sites) and Melville Islands (236 survey sites) from 2000 to 2002 by the Northern Territory Government (Table 1), with sites sampling the geographic and environmental gradients of the islands for three nights at each site. Described by Firth et al. (2006), the dataset used for modeling comprised presence-absence records for twelve small- to medium-sized non-flying mammal species across 351 survey sites. From the dataset, seven target species were chosen based on the availability and extent of the species occurrence data (>20 presence records; Table 1): the black-footed tree-rat (Mesembriomys gouldii), brush-tailed rabbit-rat (Conilurus penicillatus), common brushtail possum (Trichosurus vulpecula), delicate mouse (Pseudomys delicatulus), grassland melomys (Melomys burtoni), northern brown bandicoot (Isonodon macrourus) and pale field-rat (Rattus tunneyi). This set includes species with contrasting ecologies (e.g., arboreal/non-arboreal, granivorous/insectivorous, hollow-dependent/not) and representative of different taxonomic groups (marsupials/rodents).

| Species                  | Survey site occurrences | Home range size (ha) | Source                          |
|-------------------------|-------------------------|----------------------|---------------------------------|
|                         | Bathurst Is. | Melville Is. | Total     |                             |                                  |
| Conilurus penicillatus  | 12 (103)      | 40 (196)       | 52 (299)  | 0.8                          | Firth et al. (2006)              |
| Isonodon macrourus      | 70 (45)       | 102 (134)      | 172 (179) | 3.1                          | Gordon (1974)                    |
| Melomys burtoni         | 23 (92)       | 31 (205)       | 54 (297)  | 0.3                          | Begg, Walsh, Woerle, and King (1983) |
| Mesembriomys gouldii    | 0 (115)       | 74 (162)       | 74 (277)  | 12.3                         | Griffiths, Koenig, Carrol, and Price (2002) |
| Pseudomys delicatulus  | 2 (113)       | 22 (214)       | 24 (327)  | 0.7                          | Lock and Wilson (1999)¹          |
| Rattus tunneyi          | 10 (105)      | 26 (210)       | 36 (315)  | 0.7                          | Hooker and Innes (1995)¹         |
| Trichosurus vulpecula   | 57 (58)       | 53 (183)       | 110 (241) | 1                            | Kerle (1998)                     |

Notes: Presence and absence points are given in the form of #presences(#absences), that is C. penicillatus on Bathurst Island = 12 presences and 103 absences = 12(103).

¹Home range was inferred from a congener of comparable size.
2.3 Environmental predictor variables

For each species, a conceptual model of how the species interacts with environmental factors was developed based on the literature and expert knowledge to select field-informed and remotely sensed proxy variables for inclusion in the SDMs (see Appendix S1: Figure 1: for an example). Pairwise correlations between variables were also examined to determine inclusion into the modelling process (Appendix S1 Table 1). The predictor variables included measures of vegetation structure and composition, fire impact, water availability and geomorphology.

We developed two sets of predictor variables. The first set is referred to as “field-informed” and comprise the field-study data used in the modelling study of Firth et al. (2006). These included variables measured at each site, and four remotely sensed variables; annual rainfall, elevation, a measure of elevation range and a wetness index, for a total of 25 candidate variables (Table 2) that were later reduced to 15 per species (as detailed in modelling methods). The second set is referred to as “remotely-sensed” and comprises available raster layers of GIS data for the Tiwi Islands, representative of typical data used for modelling exercises (Table 2). After substantial efforts to source relevant predictors, these nine variables were selected from all available rasters on the basis of suitability for the target species.

For the set of remotely sensed models, we used an existing vegetation class map for the Tiwi Islands as a proxy for vegetation structure and composition measures (Richards, Liedloff, & Schatz, 2015). Vegetation classes were assessed for suitability for each of the target species based on the expert knowledge of two authors (JW and RF) and simplified into an initial binary variable (suitable vegetation class or not; Appendix S1: Table 2), as the species data were insufficient to support the many parameters needed with multiple classes (Merow et al., 2014). Rather than using only those vegetation values recorded at sites at which species were observed, measures of vegetation suitability, density and height (Table 2) were extended to account for each species’ home range size, drawn from the literature (Table 1). Specifically, we used focal neighbourhood calculations to depict average values for realistically accessible vegetation surrounding any given survey point, for each species. Landsat fire history records (1997–2002) were used to derive two measures of fire activity relevant to the time of survey; time since the last fire at any given location; and the fire frequency at any given location. Layers recording the presence of watercourses, namely those recording rivers and swamps/waterholes (“rivers” and “sw”; see Appendix S1: Table 3) were collated into a layer which reflected the distance to the nearest water.

All GIS layers were prepared using a combination of ArcGIS (ESRI, 2011; release 10) and the statistical software package “R” (R Development Core Team, 2013; version 3.4.2). Predictors were pre-processed to match the smallest available cell size, 25m, so that resolution of the finer variables were retained (Fournier, Barbet-Massin, Rome, & Courchamp, 2017), and transformed into the projection system Universal Transverse Mercator (GDA 94, Zone 52).

2.4 Species modelling methods

We explored two methods for modelling the dataset: generalized linear models (GLMs: McCullagh & Nelder, 1989) and boosted regression trees (BRTs: Elith, Leathwick, & Hastie, 2008). GLMs and BRTs are two different methods with different strengths and limitations, for example GLMs are a classical statistical method useful for fitting smoother models, while BRTs are machine learning and can help for automatic variable selection and identification of interactions, though with the likely cost of a more complex model (Merow et al., 2014). They provide contrasting examples of common methods utilized in distribution modelling exercises. This dataset was previously modelled by Firth et al. (2006) using GLMs, with their models including site-based information that cannot be projected over the entire Tiwi landscape. Here, we reproduced the Firth et al. (2006) models but using both modelling methods, and also fitted models solely using remotely sensed predictors that allow for projection across the entire study area.

Six of the seven target species occurred on both islands, with the exception of M. gouldii, which is not recorded on Bathurst Island. As we did not possess information on species detectability, true occurrence values are not known, and what we refer to as probabilities of occurrence may more accurately be described as probabilities of trapping. Due to potential but unmeasured differences across islands in populations, environmental or disturbance attributes and biotic interactions, we explored different methods of modelling occurrences, including (a) modelling both islands together with no differentiation, (b) modelling both islands together with island as a predictor, and (c) modelling both islands separately. To handle the small amount of data, yet robustly evaluate the models, and to understand model performance across islands, we developed a structured 10-fold cross-validation for methods (a) and (b), above. Cross-validation is a well-known method for creating train and test datasets in modelling (Hastie, Tibshirani, & Friedman, 2009); adding structure is useful for testing specific questions (Buston & Elith, 2011; Roberts et al., 2017). A training dataset refers to data to which the models are initially fitted. The fitted model is then used to predict the responses for the observations in the test dataset. Based on the number of surveyed sites on each island, we allocated threefold to Bathurst Island (which contained ~30% of surveyed locations; Appendix S1: Figure 2) and seven to Melville (~70%; Appendix S1: Figure 2). This process was suitable for the set of models we created as it allowed for the evaluation of predictive capacity to each island to be comparative across models.

We also sought to evaluate the relationship between field-informed and remotely sensed predictor variables over the 351 survey sites, for those pairs of field/remote variables that were most alike. In most cases, we did not have access to direct counterparts to the field-informed variables; however, we identified two pairs that could be expected to be somewhat related; field-measured canopy cover and remotely sensed vegetation height, as well as field-measured canopy cover and remotely sensed persistent
| Process       | Field variable | Field description (sourced from Firth et al. 2006)                                      | Remote variable | Remote description                                      | Remote source                                                                 |
|---------------|----------------|-----------------------------------------------------------------------------------------|-----------------|---------------------------------------------------------|--------------------------------------------------------------------------------|
| Fire          | fireimpact     | A 5-point scale measuring the apparent severity of fire impact, from 0 (no sign of fire) to 5 (conspicuous evidence of severe crown fire) | firefreq        | Fire frequency (total number 1997–2002)                 | Created from "lsfs97 – lsfs02" layers, Richards et al. (2012)                  |
|               | lastfire       | Estimated from fire scars and regeneration whether the quadrat had been burnt during the year of sampling, the year before, 2+ years before or long unburnt | tsafire         | Time since the last fire (years)                        | Created from "lsfs97 – lsfs02" layers, Richards et al. (2012)                  |
| Geomorphology | elev           | Altitude (m) derived from a digital elevation model                                      | slope           | Slope                                                    | Hollingsworth (2003)                                                           |
|               | elevr1k        | Maximum altitude minus minimum altitude within 1 km of quadrat                          |                 |                                                         |                                                                                 |
| Island        | island         | Island that survey site is located on                                                  | island          | Island that survey site is located on                   | N/A                                                                            |
| Habitat; vegetation structure and composition | agrass         | Percentage ground cover of annual grasses                                               | height_spcode   | Vegetation height (metres)                              | Focal alteration of "height" layer, TERN AusCover                               |
|               | bagt50         | Basal area (m²/ha) of trees with d.b.h. >50 cm, derived from two sweeps of a Bitterlich gauge | pg_spcode       | Persistent green cover                                  | Focal alteration of "pgreen" layer, TERN AusCover (2015)                      |
|               | bare           | Percentage of the ground cover that was bare                                           | veg_spcode      | Vegetation suitability, i.e. Expert-derived preferences to vegetation classes for each species | Alteration of vegetation map, Richards et al. (2015)                          |
|               | cancov         | Percentage foliage cover of canopy (estimated)                                         |                 |                                                         |                                                                                 |
|               | canht          | Height of tallest woody plants (m)                                                     |                 |                                                         |                                                                                 |
|               | cov1.3         | Percentage foliage cover in layer 1–3 m above ground (i.e. low shrubs)                  |                 |                                                         |                                                                                 |
|               | cov3.5         | Percentage foliage cover in layer 3–5 m above ground (i.e. shrubs)                      |                 |                                                         |                                                                                 |
|               | cov5.10        | Percentage foliage cover in layer 5–10 m above ground (i.e. tall shrubs and low trees) |                 |                                                         |                                                                                 |
|               | dead           | Basal area (m²/ha) of standing dead trees, derived from two sweeps of a Bitterlich gauge |                 |                                                         |                                                                                 |
|               | eryth          | Basal area (m²/ha) of Erythrophleum chlorostachys, derived from two sweeps of a Bitterlich gauge |                 |                                                         |                                                                                 |
|               | fruit          | Total basal area (m²/ha) of all woody plants that produce fleshy fruits, derived from two sweeps of a Bitterlich gauge |                 |                                                         |                                                                                 |
|               | logs           | The number of logs (fallen trunks or branches with diameter >5 cm) intercepted by a 100-m transect around half of the quadrat perimeter |                 |                                                         |                                                                                 |
|               | pandanus       | Basal area (m²/ha) of Pandanus spiralis, derived from two sweeps of a Bitterlich gauge |                 |                                                         |                                                                                 |
|               | pgrass         | Percentage ground cover of perennial grasses, measured by categorising cover at 100 1-m intervals around the quadrat |                 |                                                         |                                                                                 |
|               | rockcov        | Percentage of the ground cover that was rock                                           |                 |                                                         |                                                                                 |
|               | tgrass         | The sum of annual and perennial grass covers                                           |                 |                                                         |                                                                                 |
|               | totalba        | Basal area (m²/ha) of all woody plants, derived from two sweeps of a Bitterlich gauge   |                 |                                                         |                                                                                 |

(Continues)
green cover. These variable pairs were examined to determine their similarity.

Several steps in our model fitting methods aimed to make the most efficient use of the available data. We considered that the species data were insufficient to support large numbers of variables due to relatively few (<100) presence records being available for five of the seven species considered (Table 1; Merow et al., 2014). For field-informed SDMs, the 15 most functionally relevant predictors (from the set of 25) for each species—based on the expert knowledge of two authors (JW and RF)—were included in the stepwise selection process (see Appendix S1: Table 4). For remotely sensed SDMs, all nine predictors were included.

All GLMs were fit in R using the “stats” package (version 3.4.2). GLMs with the lowest Akaike’s information criterion (AIC: Akaike, 1973) were selected using the “step.Gam” function, available in the “gam” package (version 1.15), using both-directions stepwise selection. This process was undertaken for both field-informed and remotely sensed models and was consistent across all model types (single- or both-island models) and species. The maximum complexity of responses (linear, quadratic or cubic) was set by first using univariate generalized additive models (GAMs; Hastie & Tibshirani, 1990) to visualize response shapes across variables and species (see Appendix S1: Figure 3). This process was repeated independently for each species across both- and single-island models, as responses may differ between populations.

All BRTs were fit in R using the “dismo” package (version 1.0-12) with settings of: family = bernoulli, tree complexity = 1 and bag fraction = 0.75. A tree complexity of 1 fits “stumps” (i.e., trees with only one split point), meaning that the model is additive and interactions are not fitted. This is appropriate in small datasets, since with few presence records there is insufficient information to support complex models (Merow et al., 2014). The learning rate was initially set at 0.005 (enabling slow learning) and adjusted up or down, so a minimum of 1,000 trees were generated for each species/predictor data set (Elith et al., 2008).

A total of 80 models were developed across combinations of species, data type, modelling method and island location. Models were evaluated quantitatively using two statistics which measure different aspects of predictive performance (Pearce & Ferrier, 2000): (a) the percentage deviance explained and (b) the area under the receiver operating characteristic curve (AUC; Hanley & McNeil, 1982), on both training and cross-validated test data. These model evaluations were used to select sets of “best models” for each species, with two model sets created for use in spatial prioritisation scenarios: (a) models developed using different data types (field-informed and remotely sensed data), with modelling method kept constant (all GLMs) and (b) models developed using different modelling methods (GLMs and BRTs), with data type kept constant (all remotely sensed).

Due to the survey sites appearing agglomerated when viewed from the perspective of the entire landscape, we tested the potential for spatial autocorrelation calculating Moran’s I and generating a correlogram for one of the species and model combinations.
penicillatus, GLM; see Appendix S1: Figure 4). As no significant spatial autocorrelation was detected, this issue is not considered further in this paper.

2.5 Spatial prioritisation

As the field-informed models are unable to project across the entire landscape, we used the spatial prioritisation software Marxan (Ball, Possingham, & Watts, 2009) to examine the extent to which inputs based on differing data types (field-informed and remotely sensed data) and modelling methods (GLMs and BRTs) would impact the determination of site irreplaceability for the 351 survey sites considered. Marxan determines site irreplaceability (selection frequency in the best reserve system) based on the input biodiversity features, in this case predicted species probability of occurrence across all seven species at the 351 survey sites. While not typical of reserve design scenarios that consider continuous predictions of biodiversity over an entire landscape, this site-based approach allowed for the comparison of rankings generated using differing model inputs, as the field-informed models were limited to site-specific values. Thus, for the spatial prioritisation, the spatial extent for the models developed using the remotely sensed data were constrained to the same extent as the models developed for the field-informed data. To simulate a very basic spatial

### TABLE 3 The “best” models for each species and their associated model type and evaluation statistics (AUC values and the % of deviance explained) for both field-informed and remotely sensed SDMs across Bathurst and Melville Islands

| Island  | Species       | Field-informed SDMs | Remotely-sensed SDMs |
|---------|---------------|---------------------|----------------------|
|         | Model Type    | % DE  | AUC   | Model Type    | % DE  | AUC   |
| Bathurst| C. penicillatus| GLM (both islands) | 16  | 0.79 | BRT (single island) | 1  | 0.56 |
|         | I. macrourus  | BRT (single island) | 4   | 0.64 | GLM (single island) | 3  | 0.61 |
|         | M. burtoni    | GLM (both islands) | 66  | 0.97 | GLM (single island) | 23 | 0.85 |
|         | M. gouldii    | NA       | NA   | NA   | NA       | NA  | NA   |
|         | P. delicatulus| GLM (both islands) | 18  | 0.89 | GLM (both islands) | 3  | 0.79 |
|         | R. tunneyi    | BRT (single island) | 21  | 0.83 | Null Model | –  | –    |
|         | T. vulpecula  | GLM (both islands) | 28  | 0.87 | GLM (single island) | 0  | 0.59 |
| Melville| C. penicillatus| BRT (single island) | 28  | 0.87 | GLM (single island) | 12 | 0.76 |
|         | I. macrourus  | GLM (single island) | 6   | 0.64 | BRT (both islands) | 5  | 0.63 |
|         | M. burtoni    | BRT (single island) | 42  | 0.92 | GLM (single island) | 25 | 0.86 |
|         | M. gouldii    | GLM (single island) | 9   | 0.71 | GLM (single island) | 7  | 0.66 |
|         | P. delicatulus| BRT (single island) | 19  | 0.84 | GLM (both islands) | 3  | 0.70 |
|         | R. tunneyi    | BRT (both islands) | 20  | 0.82 | GLM (single island) | 7  | 0.72 |
|         | T. vulpecula  | GLM (single island) | 18  | 0.81 | GLM (single island) | 18 | 0.67 |

Notes: Reported values for evaluation statistics are cross-validated. Values for percentage of deviance explained are rounded to the nearest per cent. “NA” refers to a result being “Not Applicable” for that species/island combination (i.e., M. gouldii being absent on Bathurst Island). “Null Model” refers to an inability to generate a model for that species/island combination that performed better than a null model due to limited data points.
prioritisation scenario, Marxan was run with settings of: repeat runs = 1,000, cost = 1, proportion of planning units in initial reserve system = 0.2, and species penalty factor = 10, with all species weighted equally across all runs. Site irreplaceability values generated using different model sets were compared using Pearson's correlation coefficient to examine the influence of input data on spatial prioritisation scenarios.

3 | RESULTS

Here, we focus on results based on held-out test data. These results are island-specific and comparable across model runs. Estimates of performance based on training data are in Appendix S1: Tables 5 and 6 and not analysed further, since more complex models such as BRTs can look better when evaluated on training data simply because they are overfitted (Elith et al., 2008).

Field-informed models typically outperformed the remotely sensed models in both %DE and AUC, for all species and across both islands, and across both modelling methods (Table 3; see Appendix S1: Tables 7 and 8 for all cross-validated results). Here, we report on and refer only to the best-ranked models for each species-data-model combination, where "best" refers to the models with the highest predictive performance, as measured by the %DE and AUC. Models built with field-informed data showed generally good discrimination on test data, with AUC values > 0.80 across one or both islands for all species, with the exception of L. macrocirus and M. gouldii (AUC values 0.64 and 0.71, respectively, mean value across all species = 0.82; Table 3). The deviance explained at test sites by each model ranged from a low of 4% (L. macrocirus, Bathurst Island) to a high of 66% (M. burtoni, Bathurst Island), with a mean of 23%. There was no clear distinction between the performance of GLMs and BRTs built using field-informed data, and single- and both-island models were both well represented in the selection of best models.

Remotely sensed models performed relatively poorly across all species; it was not uncommon for models to explain <5%–10% of the total deviance at test sites (mean = 9%). Similarly, AUC values were low, ranging from 0.56 to 0.86 (mean = 0.7) and below 0.8 for all species except M. burtoni (%DE = 23.3% and 24.9%, AUC = 0.85 and 0.86, for Bathurst and Melville, respectively; Table 3). GLMs, particularly those built using single-island data, were the most commonly selected model type in the best set of models. Mapped occurrence predictions for M. burtoni are depicted in Figure 2a. It is important to note that probabilities of occurrence generated by field-informed and remotely sensed models for this species at the 351 survey sites were often discordant (Figure 2b).

As anticipated by the results of conceptual modelling, measures of vegetation composition (vegetation suitability, canopy height, canopy cover, the presence of fleshy fruits, annual and/or perennial grass cover, etc.) and measures of water access (annual rainfall, distance to the nearest watercourse, wetness index) were the most commonly selected variables across all model types and species (Table 4). Measures of fire impact, though present in some models, were not selected as frequently, consistent with the conceptual modelling.

The relationships between remotely sensed predictors and field-informed measurements were weak. Low correlations were depicted between canopy cover (field-informed) and persistent green cover (remotely sensed; Figure 3a; Pearson correlation coefficient = 0.32). The correlations between proposed counterparts canopy height (field-informed) and vegetation height...
(remotely sensed) were even lower (Figure 3b; Pearson correlation coefficient = 0.22).

Site priority, measured by Marxan site irreplaceability rankings in which a site that is selected more frequently across the repeat runs is considered more irreplaceable than a site selected less frequently, was sensitive to differences in model input types based on both data type and modelling method. Low to marginal agreement (Pearson correlation coefficient = 0.3; Figure 4a) was found between sites identified as highly irreplaceable using field-informed and remotely sensed model inputs. The site selection generated by the field-informed and remotely sensed model inputs had some agreement over which sites were of low priority, that is those selected <200 times in a hypothetical reserve design across 1,000 runs (Figure 4a). Low to marginal agreement (Pearson correlation coefficient = 0.29; Figure 4b) was found between sites identified as highly irreplaceable using GLM and BRT model inputs. In comparison with BRT inputs,

| Variables selected | C. penicillatus | I. macrourus | M. burtoni | M. gouldii | P. delicatulus | R. tunneyi | T. vulpecula |
|--------------------|----------------|--------------|------------|------------|----------------|------------|-------------|
|                     | B   | M   | B   | M   | B   | M   | B   | M   | B   | M   | B   | M   | B   | M   |
| distwater          | -   | +   | +   | +   |       |       |       |       |       |       |       |       |       |       |
| firefreq           | -   | -   |       |       |       |       |       |       |       |       |       |       |       |       |
| height_spcode      | +   | +   | +   | +   |       |       |       |       |       |       |       |       |       |       |
| island             |       |       |       |       | +   | +   |       |       |       |       |       |       |       |       |
| pg_spcode          | +   | +   | +   | +   |       |       |       |       |       |       |       |       |       |       |
| slope              | +   | -   | +   | +   |       |       |       |       |       |       |       |       |       |       |
| tsafire            |       |       |       |       | +   |       |       |       |       |       |       |       |       |       |
| veg_spcode         | +   | +   | +   | +   |       |       |       |       |       |       |       |       |       |       |
| wetness            | +   | +   | +   | +   |       |       |       |       |       |       |       |       |       |       |
| agrass             |       |       |       |       |       |       | +   |       |       |       |       |       |       |       |
| annrain            | +   | +   | +   | +   | -   | -   | -   | -   | -   | +   | +   | -   |       |       |
| bare               |       |       |       |       | +   | +   |       |       |       |       |       |       |       |       |
| bagt50             | +   | +   | +   | +   |       |       |       |       |       |       |       |       |       |       |
| cancov             |       |       |       |       | +   | +   | -   | +   | +   |       |       |       |       |       |
| canht              | +   | -   | -   | -   |       |       | +   | +   | +   |       |       |       |       |       |
| cov1.3             |       |       |       |       |       |       | +   | +   | +   |       |       |       |       |       |
| cov3.5             |       |       |       |       |       |       |       |       |       |       |       |       |       |       |
| cov5.10            |       |       |       |       |       |       |       |       |       |       |       |       |       |       |
| dead               |       |       |       |       |       |       | -   | +   |       |       |       |       |       |       |
| dist_stream        | +   | +   | +   | +   |       |       |       |       |       |       |       |       |       |       |
| elev               |       |       |       |       |       |       | -   | +   | -   | -   | -   | -   | -   | -   |
| elevr1k            | -   |       |       |       |       |       | +   | +   | +   | +   |       |       |       |       |
| eryth              |       |       |       |       |       |       |       |       |       |       |       |       |       |       |
| fireimpact         |       |       |       |       |       |       | +   | +   |       |       |       |       |       |       |
| fruit              | +   |       |       |       |       |       |       |       |       |       |       |       |       |       |
| island             |       |       |       |       |       |       |       |       |       |       |       |       |       |       |
| lastfire           |       |       |       |       |       |       | +   | -   | -   | -   | -   | -   | -   | -   |
| logs               |       |       |       |       |       |       | +   | -   |       |       |       |       |       |       |
| pandanus           |       |       |       |       |       |       |       |       |       |       |       |       | +   | +   |
| pgrass             |       |       |       |       |       |       |       |       |       |       |       | +   |       |       |
| rockcov            |       |       |       |       |       |       |       |       |       |       |       |       |       |       |
| totalba            |       |       |       |       |       |       |       |       |       |       |       |       |       |       |
| wet                |       |       |       |       |       |       |       |       |       |       |       |       |       |       |

Notes: For BRT results, only variables with 5 or greater relative influence are listed. Variables used in remotely sensed models are in the upper section of the table and variables used in field-informed models are in the lower section. The symbol “+” denotes a positive response to a variable, and the symbol “−” denotes a negative response. Columns “B” and “M” refer to the islands of Bathurst and Melville, respectively.
which showed little ability to differentiate between site priorities, the GLM inputs allowed for some level of discrimination between high- and low-quality sites in a hypothetical reserve design across 1,000 runs (Figure 4b).

4 | DISCUSSION

Our results highlight the influence of data source and modelling method on identifying spatial conservation priorities. The species

**FIGURE 3** (a) Comparative plot of counterpart variables canopy cover (field-informed) and persistent green cover (remotely sensed). Values shown are measurements of the percentage of total cover. (b) Comparative plot of counterpart variables canopy height (field-informed) and vegetation height (remotely sensed). Values shown are measurements of height to the nearest metre.

**FIGURE 4** (a) Plotted relationship between the frequency of site selection in a Marxan site irreplaceability ranking using the “best” field-derived model inputs (x-axis) and remotely sensed model inputs (y-axis). Each circle represents a survey site, while numerical values represent the number of times a given site was selected for inclusion in a hypothetical reserve design across 1,000 runs. A site that occurs more frequently across repeat runs is deemed to be more irreplaceable. (b) Plotted relationship between the frequency of site selection in a Marxan site irreplaceability ranking using the “best” BRT inputs (x-axis) and GLM inputs (y-axis). Each circle represents a survey site, while numerical values represent the number of times a given site was selected for inclusion in a hypothetical reserve design across 1,000 runs. A site that occurs more frequently across repeat runs is deemed to be more irreplaceable.
records we used in this study are relatively high quality, being based on targeted surveys that use methods known to be successful in trapping these species. These were a good basis for testing the effect of predictor variables and modelling methods. Spatial planning scenarios—in the form of Marxan irreplaceability rankings—were shown to be sensitive to variation in model inputs from both data type (field-informed or remotely sensed) and modelling method (GLMs or BRTs). For SDM outputs, models built using field-informed measurements had superior predictive capacity over those utilizing remotely sensed data; they also established that these species records can be modelled reasonably given relevant predictor variables. Models generated using remotely sensed data lead to outputs with low predictive capacity. Despite extensive modelling efforts and careful consideration of methods to address and minimize the challenges of our data, our study suggests that existing species-occurrence and remotely sensed data used in this study are insufficient to provide reliable SDMs for the target species on the Tiwi Islands. Our following discussion explores the likely reasons for these results. While this might have been a particularly challenging location and set of species to model, our study shows the strong need for models to be carefully evaluated before they are used to support conservation decisions. It should also be noted that as conceptual modelling guided the process of variable inclusion/exclusion to some degree, our results may not be directly applicable to purely exploratory analyses.

Both field-informed and remotely sensed models selected variables relating to vegetation composition and water access to be important drivers of small mammal distributions. These results are broadly supportive of the existing literature regarding the ecology of the target species (Aplin et al., 2008; Braithwaite & Griffiths, 1996; Firth, Woinarski, & Noske, 2006; Friend, 1987; Friend & Taylor, 1985; Gordon, 1974; Kerle, 1985,2008; Kerle & How, 2008; Woinarski et al., 2003,1999). Despite these broad agreements between the field-informed and remotely sensed models in the types of variables that were most important, models still varied in their estimates of probability of occupancy at survey sites. Remotely sensed models suffered from poor predictive ability across all species except M. burtoni, with a mean deviance explained (across species) of 8.9% compared to a mean of 22.5% for field-informed models. This difference in predictive capacity was particularly apparent for P. delicatulus and R. tunneyi, possibly due to field-informed predictors representing the particular local habitat conditions at the time of survey, which may be most relevant to very small rodents with restricted home ranges. While it is possible that remotely sensed fine-grain variables might have been able to provide similar measures to the field-informed data because the technology to remotely capture local habitat conditions now exists (Woodget, Carbonneau, Visser, & Maddock, 2014), such data were unfortunately unavailable for this study. Hence, we interpret these results as a comparison of existing field-informed data with existing remotely sensed data. It should not be used as a broader commentary on whether remotely sensed data are useful in SDMs.

The high predictive capacity of the models for M. burtoni was consistent with existing knowledge for this species; M. burtoni is more closely associated with restricted habitats (primarily rain forests and mangroves, as evident in Figure 2a) in this setting compared with the other species considered, which occur broadly and with more nuanced variation across the extensive eucalypt woodlands and open forests. However, even for this species the remotely sensed variables did not predict as well as field-informed ones, and there are clear differences between the per-site predicted probabilities from both models (see Figure 2b), which then affect the estimates of site irreplaceability. This implies that these models based on the available remotely sensed variables are not reliable enough to provide a strong basis for conservation planning. It supports the call for adequate assessments of the explanatory or predictive ability of models (MacNally, Duncan, Thomson, & Yen, 2017) to enable sound conservation and management planning.

Relatively few studies are able to compare the predictive performance of field-informed and remotely sensed SDMs. Similarly to the results of our study, Johnston, Elmore, Mokany, Lisk, and Fitzpatrick (2017) found that remotely sensed variables were broadly inferior to field-informed variables as predictors in low-order stream biodiversity models. While our results support the notion that, generally, the use of proximal predictors for the drivers of a species’ ecology is preferable to distal predictors (Austin, 2002; Leathwick & Whitehead, 2001), the extent to which the poor performance of these remotely sensed models are directly due to an inability of the existing remotely sensed predictors to capture the proximal processes observed using field-informed measurements is difficult to determine. The two major types of environmental data in this study (field-informed and remotely sensed) were collected across—and therefore describe—different temporal windows. Field-informed data were collected in the period between 2000 and 2003, while remotely sensed data span numerous, often more lengthy time frames (i.e., vegetation height was measured over the period of 2000 to 2010). The low correlations exhibited between field-informed measurements of canopy height and canopy cover when compared to their remotely sensed counterparts, vegetation height and persistent green cover (Figure 3) suggest that, in this instance, we should be cautious of the ability of remotely sensed predictors to accurately reflect what is measured on the ground, or at least the notion that they are measuring the same thing. Inherently, different components of the vegetation (i.e., from below and from above the tree canopy) are being measured, and uncertainty in predictor variables can be introduced by both instrument error and spatial scaling (Huston, 2002).

Spatial planning priorities—represented in this study by Marxan site irreplaceability rankings—were shown to be sensitive to both data type and modelling method. Comparisons between rankings generated using different data types revealed a low correlation, likewise for rankings generated using different modelling methods. The shape of this relationship between data types suggests that while field-informed model inputs were able, to some degree, to distinguish between valuable sites and those which offered little value to the species considered, remotely sensed model inputs were less able to discern such differences. Similarly, GLM inputs appear more able to discern these differences than BRT inputs. It is unsurprising that GLMs typically outperformed BRTs in model evaluation statistics (and subsequent utility in spatial prioritisation scenarios), given the small number of presence
records available for many of the target species. Simple models are considered preferable when data are limited (Merow et al., 2014).

In theory, areas identified as being of high conservation value using both data types and modelling methods could be considered more robust investments, or potentially as "sites of interest" in future targeted surveys. However, while congruence between predictions from differing input types gives some confidence, large uncertainties surrounding the ability of predictions to correctly identify priority areas remain. Further, since field-informed variables are not available island-wide, the field-informed models are of limited use in conservation planning. The Tiwi landscape may not be well suited to species distribution modelling, as it is subject to relatively large stochastic events (e.g., cyclones that cause a change in vegetation structure) and environmental variation across the islands is generally subtle, largely composed of minor variation of *E. miniata* and *E. tetrodonta* open forest. Occurrence may be better explained by continually-shifting factors such as resource availability (e.g., occurrence and/or abundance of particularly-favoured grass or fruit-bearing shrub species, hollow size and abundance, etc.), species co-occurrence, and predation risks, including feral cats (Davies et al., 2017). Future studies could model cat distribution independently and include the results as a covariate in SDM attempts, similar to the approach of Davies et al. (2017).

When uncertainty surrounding SDM predictions are high, attempts at spatial prioritisation that lack a detailed and realistic assessment of limitations may be misleading. In the face of such uncertainties, assuming the worst-case scenario—that some SDM predictions may hold little value for informing management decisions—may be the best strategy. This touches upon an already difficult decision-making process in conservation planning; while not enough data may lead to poor conservation decisions (Possingham, Grantham, & Rondinini, 2007), some studies have demonstrated diminishing returns on investment in data collection and suggested that early conservation actions may be most beneficial (Grantham et al., 2008; Grantham, Wilson, Molianen, Rebelo, & Possingham, 2009). Indeed, it is not uncommon that conservation planning is done without SDMs (Tulloch et al., 2016). Ultimately, the decision to use SDMs as conservation planning tools should be made on a case-by-case basis based on evidence around model performance. In some cases, funding may be better directed to detailed autecological studies of specific species to determine the driving causes of their decline and to develop and implement management responses to such threats. While the financial costs associated with obtaining such information can be great, so too are the potentially undesirable—and typically irreversible—consequences which may result from suboptimal decision-making, resulting in less desirable conservation outcomes.

Nevertheless, the modelling process itself can provide important information, including structuring information through conceptual modelling to identify potential key drivers, and revealing data and knowledge gaps. Our exploration into some of the underlying assumptions of model construction, performance and efficacy demonstrates the need for considered modelling—and model evaluation—to support robust conservation decisions. The low cross-validated AUC values and deviance explained by remotely sensed models highlighted the limitations of existing GIS and species data to provide reliable and useful predictions for conservation planning on the Tiwi Islands. In a broader context, our study supported notions that models should be adequately assessed for useful explanatory or predictive ability (MacNally et al., 2017), and demonstrated how field measurements can be a useful tool in this process.

**ACKNOWLEDGEMENTS**

First and foremost we wish to thank the Tiwi Land Council and the Tiwi traditional owners for their ongoing support of scientific research on their land. We are also particularly grateful to Kate Hadden for her organisational efforts, enabling us to obtain data. We acknowledge the contributions of the Northern Territory Government for allowing the use of their biodiversity survey data, and its many original collectors for their efforts. WLM was supported by sponsorship from the Australian Geographic Society; BM by fellowships from the Australian Research Council (ARC; DE130100434 and FT170100004), and JE by ARC project DP160101003 and University of Melbourne’s Indigenous Research Initiative.

**DATA ACCESSIBILITY**

The data that support the findings of this study are available from the Tiwi Land Council (TLC), but restrictions apply to the availability of these data, which were used under licence for the current study, and are not publicly available. For licensing details please contact the Tiwi Land Council.

[Correction statement added on 17 May 2019 after first online publication: The data accessibility section has been updated in this version]

**ORCID**

*William La Marca* https://orcid.org/0000-0001-6513-1587

**REFERENCES**

Akaike, H. (1973). Information theory as an extension of the maximum likelihood principle. In B. N. Petrov, & F. Csaki (Eds.), Second international symposium on information theory (pp. 267–281). Budapest, Hungary: Akademiai Kiado.

Aplin, K., Dickman, C., Helgen, K., Winter, J., Ellis, M., & Burbidge, A. (2008). Melomys burtoni. The IUCN Red List of Threatened Species 2008: e.T13116A3410301. 2015. https://doi.org/10.2305/IUCN.UK.2008.RLTS.T13116A3410301.en

Austin, M. P. (2002). Spatial prediction of species distribution: An interface between ecological theory and statistical modelling. *Ecological Modelling*, 157, 101–118. https://doi.org/10.1016/S0304-3800(02)00205-3

Ball, I. R., Possingham, H. P., & Watts, M. (2009). Marxan and relatives: Software for spatial conservation prioritisation. In A. Molianen, K. A. Wilson, & H. P. Possingham (Eds.), *Chapter 14: Spatial conservation prioritisation: Quantitative methods and computational tools* (pp. 185–195). Oxford, UK: Oxford University Press.
Fournier, A., Barbet-Massin, M., Rome, Q., & Courchamp, F. (2017). Predicting species distribution combining multi-scale drivers. *Global Ecology and Conservation, 12*, 215–226. https://doi.org/10.1016/j.gecco.2017.11.002

Franklin, J. (2010). *Mapping species distributions: Spatial inference and prediction*. Cambridge, UK: Cambridge University Press.

Friend, G. R. (1987). Population ecology of mesembriomys Gouldii (Rodentia: Muridae) in the wet-dry tropics of the Northern Territory. *Australian Wildlife Research, 14*, 293–303. https://doi.org/10.1071/WR9870293

Friend, G. R., & Taylor, J. A. (1985). Habitat preferences of small mammals in tropical open-forest of the Northern Territory. *Australian Journal of Ecology, 10(2)*, 173–185. https://doi.org/10.1111/j.1442-9993.1985.tb00879x

Gordon, G. (1974). Movements and activity of the short-nosed bandicoot Isoodon macrourus Gould (Marsupialia). *Mammalia, 38*(3), 405–432.

Grantham, H. S., Molianen, A., Wilson, K. A., Pressey, R. L., Rebelo, T. G., & Possingham, H. P. (2008). Diminishing return on investment for biodiversity data in conservation planning. *Conservation Letters, 1*, 190–198. https://doi.org/10.1111/j.1755-263X.2008.00029.x

Grantham, H. S., Wilson, K. A., Molianen, A., Rebelo, T., & Possingham, H. P. (2009). Delaying conservation actions for improved knowledge: How long should we wait? *Ecology Letters, 12*, 293–301.

Griffiths, A. D., Koenig, J., Carrol, F., & Price, O. (2002). Activity area and day-time tree use of the black-footed tree-rat Mesembriomys gouldii. *Australian Mammalogy, 23*(2), 181–183.

Guillera-Arroita, G., Lahoz-Monfort, J. J., Elith, J., Gordon, A., Kujala, H., Lentini, P. E., .... Wintle, B. A. (2015). Is my species distribution model fit for purpose? Matching data and models to applications. *Global Ecology and Biogeography, 24*, 276–292. https://doi.org/10.1111/geb.12268

Guisan, A., Tingley, R., Baumgartner, J. B., Naujokaitis-Lewis, I., Sutcliffe, P. R., Tulloch, A. I. T., .... Buckley, Y. M. (2013). Predicting species distributions for conservation decisions. *Ecology Letters, 16*, 1424–1435. https://doi.org/10.1111/ele.12189

Guisan, A., & Zimmerman, N. E. (2000). Predictive habitat distribution models in ecology. *Ecological Modelling, 135*, 147–186. https://doi.org/10.1016/S0304-3800(00)00354-9

Hanley, J. A., & McNeil, B. J. (1982). The meaning and use of the area under a Receiver Operating Characteristic (ROC) curve. *Radiology, 143*, 29–36. https://doi.org/10.1148/radiology.143.1.7063747

Hastie, T., & Tibshirani, R. (1990). In D. R. Cox, D. V. Hinkley, D. Rubin, & B. W. Silverman (Eds.), *Generalized additive models*. Monographs on statistics and applied probability. London, UK: Chapman & Hall.

Hastie, T., Tibshirani, R., & Friedman, J. H. (2009). *The elements of statistical learning: Data mining, inference, and prediction*, 2nd ed. New York, NY: Springer-Verlag.

Heikkinen, R. K., Luoto, M., Araujo, M. B., Virkkala, R., Thuiller, W., & Sykes, M. T. (2006). Methods and uncertainties in bioclimatic envelope modelling under climate change. *Progress in Physical Geography, 30*, 751–777. https://doi.org/10.1177/0309133306071957

Hollingsworth, I. (2003). *Land capability study of the Tiwi Islands*. Darwin, NT: EWL Sciences Pty Ltd for the Northern Territory Department of Infrastructure Planning and Environment.

Hooker, S., & Innes, J. (1995). Ranging behaviour of forest-dwelling ship rats, Rattus rattus, and effects of poisoning with brodifacoum. *New Zealand Journal of Zoology, 22*, 291–304.

Huston, M. A. (2002). Critical issues for improving predictions. In: J. M. Scott, P. J. Heglund, M. L. Morrison, M. G. Raphael, W. A. Wall, & F. B. Samson (Eds.), *Predicting species occurrences: Issues of accuracy and scale* (pp. 7–24). Covelu, CA: Island Press.

Johnston, M. R., Elmore, A. J., Mokany, K., Lisk, M., & Fitzpatrick, M. C. (2017). Field-measured variables outperform derived alternatives in Maryland stream biodiversity models. *Diversity and Distributions, 23*(1054), 1066. https://doi.org/10.1111/ddi.12598

Kerle, J. A. (1985). Habitat preference and diet of the Northern Brushtail Possum Trichosurus anhemensis in the Alligator Rivers Region, NT. *Proceedings of the Ecological Society of Australia*, 13, 161–176.

Kerle, J. A. (1998). The population dynamics of a tropical possum, *Trichosurus vulpecula anhemensis*. Collett. *Wildlife Research, 25*, 171–182. https://doi.org/10.1071/WR96113

Kerle, J. A. (2008). Grassland Melomys, *Melomys burtoni*. In S. Van Dyck, & R. Strahan (Eds.), *The mammals of Australia*, 3rd ed. (pp. 667–669). Sydney, NSW: Reed New Holland.

Kerle, J. A., & How, R. A. (2008). Common Brushtail Possum *Trichosurus vulpecula*. In S. Van Dyck, & R. Strahan (Eds.), *The mammals of Australia*, 3rd ed. (pp. 274–276). Chatswood, NSW: Reed New Holland.
