

# Modelling the potential geographic distribution of invasive ant species in New Zealand

Darren F. Ward

Received: 10 August 2006 / Accepted: 26 October 2006 / Published online: 29 November 2006  
© Springer Science+Business Media B.V. 2006

**Abstract** Despite their economic and environmental impacts, there have been relatively few attempts to model the distribution of invasive ant species. In this study, the potential distribution of six invasive ant species in New Zealand are modelled using three fundamentally different methods (BIOCLIM, DOMAIN, MAXENT). Species records were obtained from museum collections in New Zealand. There was a significant relationship between the length of time an exotic species had been present in New Zealand and its geographic range. This is the first time such a time lag has been described for exotic ant species, and shows there is a considerable time lag in their spread. For example, it has taken many species several decades (40–60 years) to obtain a distribution of 17–25% of New Zealand regions. For all six species, BIOCLIM performed poorly compared to the other two modelling methods. BIOCLIM had lower AUC scores and higher omission error, suggesting BIOCLIM models under-predicted the potential distribution of each species. Omission error was significantly higher between models fitted with all 19 climate variables compared to those models with fewer

climate variables for BIOCLIM, but not DOMAIN or MAXENT. Widespread species had a greater commission error. A number of regions in New Zealand are predicted to be climatically suitable for the six species modelled, particularly coastal and lowland areas of both the North and South Islands.

**Keywords** Invasive ants · BIOCLIM · DOMAIN · MAXENT · Species distribution modelling

## Introduction

Invasive species are a global problem, affecting agriculture, forestry, fisheries, human health and natural ecosystems (Drake et al. 1989; Mooney and Drake 1986; Sandland et al. 1999; Mack et al. 2000). A fundamental approach to understanding and managing invasive species is to determine their current and potential distribution. There has been a number of recent papers providing an overview to species distribution modelling, or a comparison of modelling methods (Fielding and Bell 1997; Guisan and Zimmermann 2000; Zaniwski et al. 2002; Anderson et al. 2003; Segurado and Araújo 2004; Guisan and Thuiller 2005; Elith et al. 2006; Phillips et al. 2006). Essentially, species distribution modelling aims to predict areas that describe where environmental

---

D. F. Ward (✉)  
School of Biological Sciences, University of  
Auckland, Tamaki Campus, Private Bag 92019,  
Auckland, New Zealand  
e-mail: d.ward@auckland.ac.nz

conditions are suitable for the survival of the species. That is, the potential distribution or fundamental niche (Anderson et al. 2003; Peterson 2003; Guisan and Thuiller 2005).

In general, these modelling methods combine species locality data (geo-referenced coordinates of latitude and longitude from confirmed presence) with environmental variables to create a model of a species requirements for the examined variables (Anderson et al. 2003). The resulting model is then projected onto a GIS map (termed a habitat suitability map), of the study region showing the potential geographic distribution of a species. For invasive species management, habitat suitability maps identify areas where (1) invasive species may actually be present (but are as yet undetected), and (2) where invasive species may disperse to in the future, thus providing assistance for planning and prioritising areas for surveillance. Such information can also assist in determining the extent, cost and likelihood of success of a control programme. Thus, predictive modelling of a species distribution represents an important tool for invasive species management (Anderson et al. 2003).

Invasive ant species are currently receiving considerable attention from around the globe, with increasing evidence of economic and agricultural impacts, health effects on humans, and disruption to natural ecosystems (Williams 1994; Christian 2001; Holway et al. 2002; O'Dowd et al. 2003; Ward and Harris 2005). Although only a handful of invasive ant species are well studied (Holway et al. 2002), there are many other ant species with the opportunity to become invasive. For example, at least 150 species of ants have been accidentally transported by humans to new regions through global trade (McGlynn 1999). However, this number is almost certainly an underestimate (Suarez et al. 2005; Ward et al. 2006).

Climatic variables, especially temperature, rainfall and humidity, play a large role in determining the distribution of ant species. Hölldobler and Wilson (1990) state that every ant species operates within a temperature-humidity envelope. However, this climate envelope is more apparent at large spatial scales, where at local levels it is strongly shaped by species microhabitat

specialisation and strategies to avoid inter-specific competition (Hölldobler and Wilson 1990). On large spatial scales ant abundance is strongly correlated with net primary productivity (a function of solar radiation and rainfall) (Kaspari et al. 2000). Temperature also plays an important role in the abundance of ants by restricting foraging activity and regulating seasonal productivity (Kaspari et al. 2000). Environments with high rainfall reduce the time spent foraging (Vega and Rust 2001). Conversely, in xeric habitats, the lack of water and soil moisture can also limit the distribution of some species (Holway and Suarez 2006).

At the level of the colony, the location and construction of nests play an important role in regulating temperature and humidity (Hölldobler and Wilson 1990). For example, nests can also provide a thermal refuge in hot environments, allowing workers to retreat to a cool nest in the hottest part of the day. Temperature primarily controls the development of the eggs, larvae and pupae (Hartley and Lester 2003). Some ant species are known to move brood vertically within the nest to keep them at the optimum temperature for development (Hölldobler and Wilson 1990). Extremes of temperature are known to severely limit, or stop, the production of workers and reproductive castes, which can ultimately kill the colony (Korzukhin et al. 2001).

Despite the importance of climate variables to the survival and distribution of ants, there have been relatively few attempts to model the distribution of invasive ant species. Hartley and Lester (2003) used climate station records and a degree-day model for each life stage of the Argentine ant, *Linepithema humile*, to examine its potential distribution in New Zealand. Roura-Pascual et al. (2004) also examined the potential distribution of *L. humile* on a global distribution and under climate change scenarios using a genetic algorithm for rule-set prediction (GARP) model.

Pimm and Bartell (1980) provided one of the first distribution models for the red imported fire ant, *Solenopsis invicta*, an invasive species in the south-eastern USA. They used a principal coordinate analysis to model the distribution in Texas on two environmental axes (temperature and rainfall). However, they overestimated the extent

of expansion under dry conditions (Korzukhin et al. 2001). Stoker et al. (1994) used a complex mechanistic model to simulate population and colony growth of *S. invicta* at different temperature and rainfall regimes. More recently Korzukhin et al. (2001) has provided a simulation model for *S. invicta* based on colony growth as a function of soil temperature. The production of female alates (reproductives) of a colony was estimated and this provided an assessment of whether *S. invicta* could survive in different locations. The model of Korzukhin et al. (2001) has also been applied on a global scale by Morrison et al. (2003). Sutherst and Maywald (2005) have also modelled *S. invicta* at a global scale using colony growth and stress parameters in the programme CLIMEX.

These models have allowed the identification of areas that are climatically suitable, as well as providing insight into the factors that may limit the expansion of these two invasive species. In this study, the potential distribution of six invasive ant species in New Zealand is modelled using three fundamentally different methods. New Zealand has a very small native ant fauna of 11 species (Ward 2005). As a consequence, the establishment and subsequent spread of invasive ant species is less likely to be determined by inter-specific competition from native ant species. Thus, climate variables are most likely to be the primary factor in restricting the occurrence of invasive ant species on a large-scale in New Zealand.

## Methods

### Species records

Records of all exotic species established in New Zealand were obtained from an online database (Landcare Research 2006). The database represents records from a 90-year period (from the early 1900s to 2004) of specimens held in museums throughout the country. Specimens are collections made by professional scientists (from Universities and government institutions), amateur entomologists, and members of the public. All records contained information of locality,

year of collection, and the majority (> 90%) of records contained information on habitat, collector, and a map reference. The database consists of over 2,000 species-locality records.

### Spread since arrival

The geographical spread since the arrival of a species was estimated by using the date of the first recorded presence in New Zealand and the number of coded regions ('Crosby codes') each species has been recorded within. Crosby codes are equal-sized regions throughout New Zealand and are primarily used for the retrieval and documentation of entomological specimens in New Zealand collections (Crosby et al. 1998). They are used here as a measure of distribution. There are 29 Crosby regions for New Zealand.

### Modelling potential distribution

#### Data sources

Not all exotic species in New Zealand were modelled. Many species are known from very few records or are very recent establishments. Six invasive species were chosen to be modelled: *Iridomyrmex* sp. (undescribed), *Ochetellus glaber*, *Paratrechina* sp. (undescribed), *Pheidole rugosula*, *Technomyrmex albipes*, and *Tetramorium grassii* (Table 1). These species have been in New Zealand for many decades, are among the most geographically widespread and are increasingly found in native ecosystems.

Environmental data was obtained from WORLDCLIM (version 1.3, <http://www.worldclim.org>) which is explained in detail in Hijmans et al. 2005). WORLDCLIM contains climate data (monthly precipitation and monthly mean, minimum and maximum temperature) at a spatial resolution of 30 arc seconds (~1 × 1 km resolution) obtained by interpolation of climate station records from 1950–2000. From this climate data, 19 climate variables are derived: annual mean temperature [1], mean monthly temperature range [2], isothermality [3], temperature seasonality [4], maximum temperature of warmest month [5], minimum temperature of coldest

**Table 1** The six exotic ants species modelled and the number of training records used in each partition. PCA 1–3 refers to the contribution of the first, second and third PCA axis in explaining the cumulative percentage variation in climate variables. The selected climate variables

refers to variables used in comparison to ‘all variables’, numbered variables are given in the methods. The number of Crosby regions is a measure of the extent of a species distribution in New Zealand (maximum 29)

Species	Training records	PCA 1	PCA 2	PCA 3	Selected variables <sup>a</sup>	Crosby Regions
<i>Iridomyrmex</i> sp.	143	40.5	75.4	87.2	1, 3, 5, 12	8
<i>Ochetellus glaber</i>	125	43.1	76.7	87.1	1, 2, 3, 12	11
<i>Paratrechina</i> sp.	180	42.6	75.8	85.8	1, 2, 3, 7, 12	10
<i>Pheidole rugosula</i>	58	52.0	77.8	86.1	1, 2, 3, 5, 10, 14, 16	9
<i>Technomyrmex albipes</i>	122	40.3	73.7	84.5	1, 2, 3, 4, 12	16
<i>Tetramorium grassii</i>	94	81.9	98.4	99.4	1, 4, 12, 16	7

<sup>a</sup> Annual mean temperature [1], mean monthly temperature range [2], isothermality [3], temperature seasonality [4], maximum temperature of warmest month [5], temperature annual range [7], mean temperature of warmest quarter [10], annual precipitation [12], precipitation of driest month [14], precipitation of wettest quarter [16]

month [6], temperature annual range [7], mean temperature of wettest quarter [8], mean temperature of driest quarter [9], mean temperature of warmest quarter [10], mean temperature of coldest quarter [11], annual precipitation [12], precipitation of wettest month [13], precipitation of driest month [14], precipitation seasonality [15], precipitation of wettest quarter [16], precipitation of driest quarter [17], precipitation of warmest quarter [18], and precipitation of coldest quarter [19]. For further detail see (<http://www.worldclim.org>, or Hijmans et al. 2005). The same climate variables were used in each modelling method. The climate variables represent a combination of annual trends, seasonality and extreme environmental conditions.

### Modelling methods

Three different modelling methods were used. First, BIOCLIM (Nix 1986), uses a climate envelope (a rectilinear volume in environmental space) to summarise the climate at locations where a species has been recorded. BIOCLIM predicts suitable conditions for a species where values of the climate variables fit within the extreme values determined by the set of known locations. Locations where the values lie within the 5–95th percentile of the climate envelope are traditionally classified as ‘core’ regions of suitability. The second method, DOMAIN (Carpenter et al. 1993), uses a distance-based method (the Gower metric) to assess new sites in terms of their

environmental similarity to sites of known presence. DOMAIN produces an index of habitat suitability on a continuous scale (0–100), where higher scores (e.g. > 90) are considered highly suitable. Both BIOCLIM and DOMAIN modelling methods were implemented in DIVA-GIS software (version 5.2, <http://www.diva-gis.org>).

The third method is maximum entropy species distribution modelling (MAXENT, version 2.2), a general-purpose machine learning method (Phillips et al. 2004). Entropy in the context of probability theory and statistics measures the amount of information that is contained in a random variable or unknown quantity. The idea of MAXENT is to estimate the target probability distribution by finding the probability distribution of maximum entropy, that is, the closest to uniform. This is equivalent to finding the maximum likelihood Gibbs distribution. MAXENT software and further information on this method are available from <http://www.cs.princeton.edu/~schapire/maxent> (or see Phillips et al. 2006).

The three modelling methods differ in their theoretical assumptions, modelling procedures, novelty and performance. BIOCLIM is an established method and has been widely used for species distribution modelling (Télez-Valdés and Dávila-Aranda 2003; Meynecke 2004; Beaumont et al. 2005), DOMAIN has not been widely used (Loiselle et al. 2003), and MAXENT has only recently been applied to modelling species distributions (Phillips et al. 2006). Despite using three very different modelling methods, all use the

same basic set of information to model the distribution of a species. That is, a set of samples (species presence) is available from a geographical region, which is linked to a set of features (e.g. climatic variables).

### *Model building and evaluation*

For each species, 10 random partitions were made of species records by bootstrapping. Each partition was created by randomly selecting 75% of the species records as training data. The remaining 25% of species records were set aside for testing the resulting models. This is a split-sample approach (Guisan and Zimmerman 2000). Ten partitions were made to assess the variability of each method and to allow statistical testing of differences in performance (see Phillips et al. 2006). Data was ‘cleaned’ in DIVA-GIS where duplicates records were deleted and only one species occurrence record per grid cell was allowed. Coordinates for species records in the online database are listed as New Zealand map grid references, these were converted to decimal latitude and longitude in DIVA-GIS.

Species records consist of individual point-locality data, that is, presence-only data. However, in order to evaluate models on the basis of error rates, absence data is needed. To overcome the lack of absence data, ‘pseudo-absence’ data is generated which uses random points throughout the study area as assumed absences (Zaniewski et al. 2002). For BIOCLIM and DOMAIN pseudo-absence were generated at random in a 1:1 ratio with the number of presence records. MAXENT uses 10,000 random background points in the study area to define the probability distribution and evaluate model predictions.

In presence/absence models there are two types of prediction errors (Fielding and Bell 1997). False negatives (omission error, under-prediction) result in areas being classified as climatically unsuitable when they are not. Conversely, false positives (commission error, over-prediction) result in areas being classified as climatically suitable when they are not. For invasive species it is more important to minimise false negatives. That is, it is better to predict that

an invasive species will occur in a area (but it never happens) than to predict an invasive will not occur in an area, when it actually could. Models that have a high proportion of presences correctly predicted (i.e. model sensitivity) and a low omission error (false negative rate) should be preferred. Commission error are likely to result from the species not yet having colonised all climatically suitable locations and dispersal limitations (Guisan and Thuiller 2005), and for presence-only modelling apparent commission error will exist, where the species is present in an area but surveys have not been undertaken to confirm this (Anderson et al. 2003).

Omission error was determined through a confusion matrix (Fielding and Bell 1997). A threshold was applied to each modelling method because an upper limit is needed to determine what values represent true presence and true absences (BIOCLIM = 25, DOMAIN = 90, MAXENT = 1, see Phillips et al. 2006). Optimal models were defined as ‘omission error < 0.05’ by the criteria of (Anderson et al. 2003). The area under the Receiver Operating Characteristic curve (AUC) was also used to examine model performance. AUC measures the ability of a model to discriminate between sites where a species is present versus those where it is absent (Fielding and Bell 1997; Elith et al. 2006). It provides a single measure of overall accuracy that is not dependent upon a particular threshold (Fielding and Bell 1997). AUC ranges from 0 to 1, where a score of 1 indicates perfect discrimination, a score of 0.5 implies discrimination that is no better than random. A value of 0.8 for the AUC means that there is a 80% probability that a random selection from the presence records will have a model score greater than a random selection from the absence records.

Another issue in the modelling of species distributions is the number of climatic variables used in modelling (Kriticos and Randell 2001; Beaumont et al. 2005). Using too few, or too many climatic variables may result in incorrect predictions. To examine the influence on the number of climate variables on model performance climate data was generated for each species record in DIVA-GIS. Principle co-ordinate analysis (PCA) within PRIMER v5.0

software (Clarke and Warwick 2005) was used to examine the similarity between the climate variables. Collinearity was examined through a Pearson correlation matrix and subsets of variables with a high average correlation (> 90%) were reduced to a single variable (recommended by Clarke and Warwick 2005). To assist in the interpretation of results the same variables were retained for each species from the PCA. The Wilcoxon signed-rank test was used to examine differences in omission error between models generated from all 19 climate variables (e.g. BIOLCIM-all) and between models with only a 'select' number of climate variables (e.g. BIOC-LIM-select).

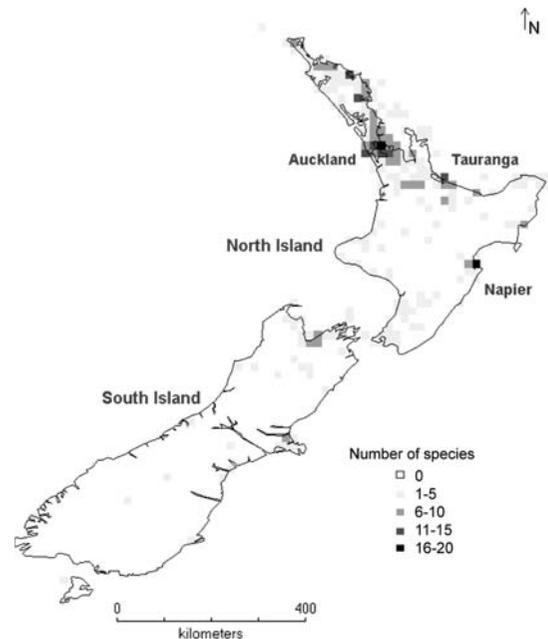
Thus, 360 models were created, using six species, three modelling methods, 10 partitions, and two options (number of climatic variables).

## Results

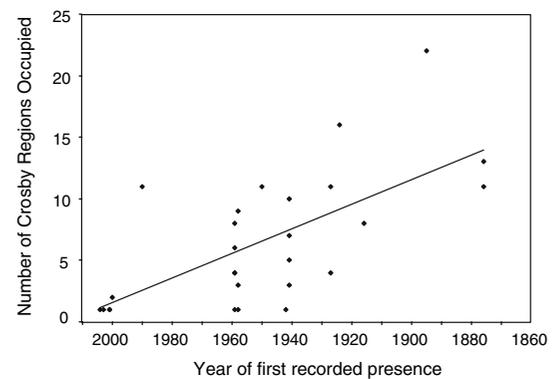
### Exotic ant species in New Zealand

There are twenty-eight exotic ant species present in New Zealand (Ward 2005). Exotic species are predominantly found in the northern regions of the North Island, and to a lesser extent in the coastal lowland regions of the North Island and the northern region of the South Island (Fig. 1). These areas generally represent the warmer areas of New Zealand. The greatest number of exotic species are concentrated in the cities with a large port: Auckland, Tauranga and Napier (Fig. 1).

There has been a relatively constant arrival of exotic species to New Zealand from the late 1800s to the present (Fig. 2). The oldest records of exotic species date from before the 1870s and are thought to have been associated with soil ballast of ships arriving during the early days of European settlement (Brown 1958). The presence of four new species in the last 5 years also serves to highlight the relatively regular establishment of ant species in New Zealand. There was a significant correlation ( $r^2 = 0.436$ ,  $P < 0.01$ ) between the length of time an exotic species has been present in New Zealand and the extent of its current distribution (Fig. 2).



**Fig. 1** The current distribution of exotic species in New Zealand (mapped on  $20 \times 20$  km scale). Points off the main islands represent small offshore islands



**Fig. 2** Relationship between the length of time an exotic species has been present in New Zealand and its current distribution (measured by the number of Crosby regions; each point represents one species,  $y = -0.1x + 201.66$ ,  $r^2 = 0.436$ ,  $P < 0.01$ )

The six species (which are being modelled in the following section) have been in New Zealand an average of 65.5 years (SE  $\pm$  6.2; range 42–84), and occupy an average of 10.2 Crosby regions (SE  $\pm$  1.3; range 7–16), for an average spread of 6.8 years for every Crosby region occupied (SE  $\pm$  0.9; range 5–11).

## Potential distribution

### *Selection of climatic variables*

The first three PCA axes consistently explained ~90% of the variation within climate data for each species (Table 1). The first PCA axis always represented temperature variables and the second precipitation variables. The number of variables selected was relatively consistent for each species, reduced from the original 19, to approximately four variables (range 4–7, Table 1). Across all species, the climate variables selected were: annual mean temperature, mean monthly temperature range, isothermality, temperature seasonality, maximum temperature of warmest month, temperature annual range, mean temperature of warmest quarter, annual precipitation, precipitation of driest month, precipitation of wettest quarter. However, it should be remembered, where high average collinearity existed within a groups of variables, the same variable for each species was consistently selected to represent this group.

For each modelling method the omission error from models of ‘all variables’ was compared to ‘selected variables’ derived from the PCA. For DOMAIN and MAXENT methods there was no significant differences between the average omission error of models with ‘all variables’ compared to ‘selected variables’ (Wilcoxon signed rank test, all species,  $P > 0.125$ ). For BIOCLIM, omission error was significantly higher for models fitted with all 19 climate variables compared to those with fewer climate variables. This was consistent for all species; *Ochetellus glaber* and *Pheidole rugosula* ( $P < 0.05$ ), *Iridomyrmex* sp., *Paratrechina* sp., *Technomyrmex albipes*, *Tetramorium grassii* ( $P < 0.01$ ).

### *Model performance*

From a plot of omission versus commission error, the performance of methods and each model ( $n = 360$ ) can be compared (Fig. 3). BIOCLIM models generally cluster in the upper left, represented by high omission and low commission, suggesting that these models are under-predicting species distributions. DOMAIN and MAXENT

models mix together, and 91% of these models have an omission error less than 0.10.

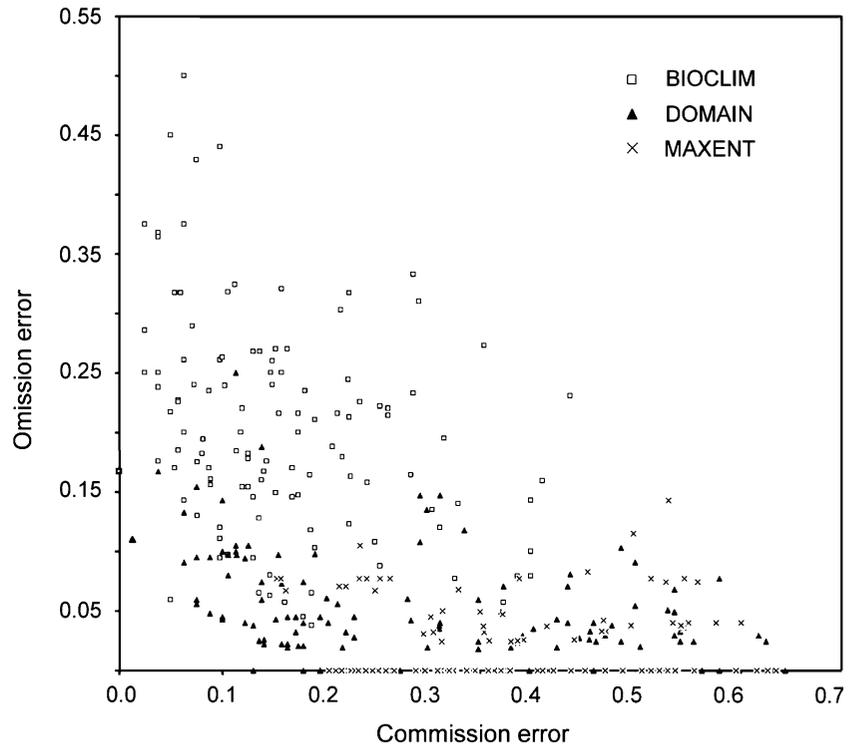
Of models with an omission error of less than 0.05, there were significantly fewer BIOCLIM models ( $n = 2$ ) represented compared to DOMAIN ( $n = 76$ ) and MAXENT ( $n = 96$ ) models ( $\chi^2 = 85.31$ , d.f. = 5,  $P < 0.001$ ). Further examination of these models revealed that species were not spread evenly across the range of observed commission error (Fig. 4). This is a consequence of widespread species having greater commission error (Anderson et al. 2003). The average commission error of these species (for models with omission error of less than 0.05) is highly correlated with the number of Crosby regions ( $r^2 = 0.865$ ), a measure of New Zealand wide distribution, but not with the number of training records ( $r^2 = 0.467$ ).

For each species, MAXENT and DOMAIN consistently performed better than BIOCLIM, with higher average AUC scores (Fig. 5a). AUC scores were not significantly correlated with the number of training records for any modelling method (all methods,  $P > 0.50$ ). Omission errors parallel AUC values, with BIOCLIM having higher average omission error than either DOMAIN or MAXENT (Fig. 5b). Omission error was not significantly correlated with the number of training records for BIOCLIM for MAXENT methods, but it was for DOMAIN ( $r^2 = 0.783$ ,  $P < 0.05$ ).

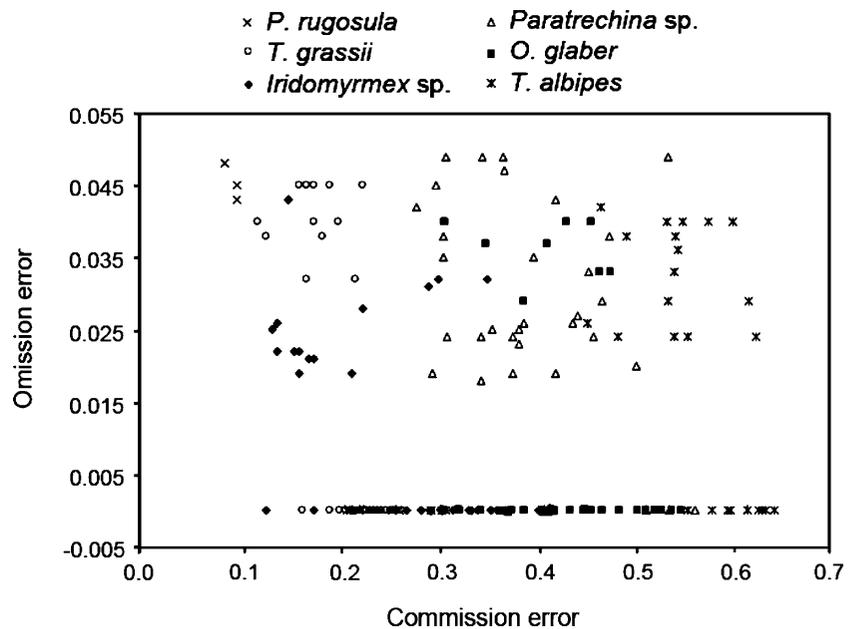
### *Habitat suitability maps*

For each species a single model was selected to create a habitat suitability map of the predicted geographic distribution (Fig. 6). Habitat suitability maps with presence-only data do not predict the probability of presence, but provide relative index of suitability (Anderson et al. 2003). The criteria of selecting ‘optimal models’ from Anderson et al. (2003) was used, that is, the model which was closest to the average commission (of all models with omission error of less than 0.05). For each species, a DOMAIN model was closest to these criteria. However, it should be noted that there were a number of DOMAIN and MAXENT models that could have been used, and the

**Fig. 3** Omission versus commission error for modelling methods. Climate variable options (all vs. select) are not distinguished

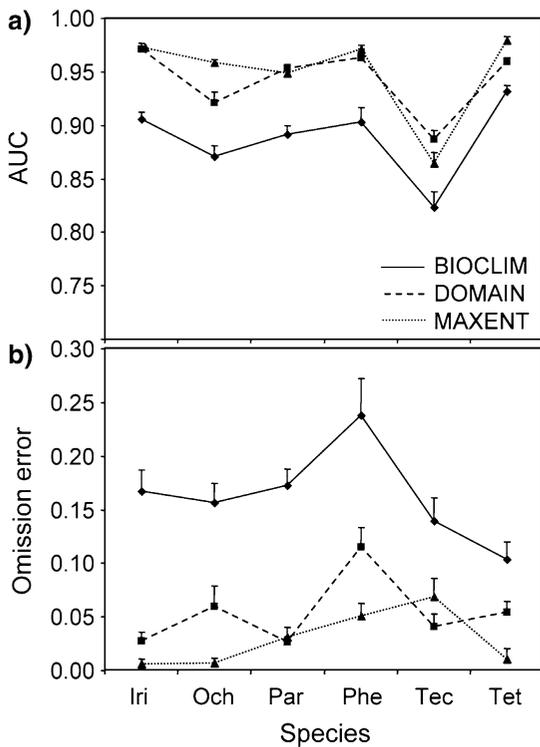


**Fig. 4** The range of commission error for each species from optimal models (omission error less than 0.05)



selection of this single model is not meant to imply that DOMAIN out-performed MAXENT. The number of Crosby regions occupied was counted for the current distribution and was

compared to predicted potential distribution for each species. On average, the current distribution was 47.9% of potential distribution, which was consistent for each species *Iridomyrmex sp.*



**Fig. 5** (a) Average AUC values (± SD) and (b) average omission error (±SD) for each species and modelling method. Codes for species are: Iri = *Iridomyrmex* sp., Och = *Ochetellus glaber*, Par = *Paratrechina* sp., Phe = *Pheidole rugosula*, Tec = *Technomyrmex albipes*, Tet = *Tetramorium grassii*

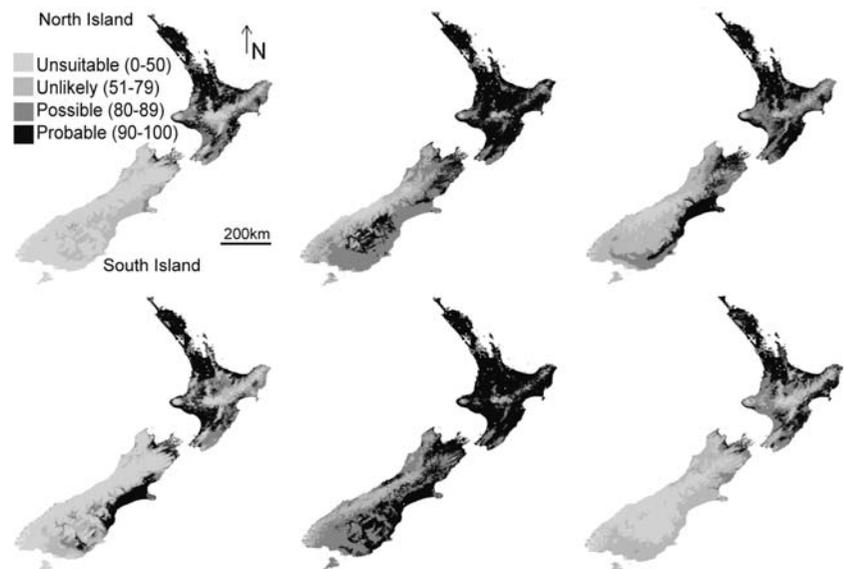
(50.0%), *Ochetellus glaber* (47.8%), *Paratrechina* sp. (45.5%), *Pheidole rugosula* (40.9%), *Technomyrmex albipes* (59.3%), *Tetramorium grassii* (43.8%).

**Discussion**

Presence-only modelling

Museum records have great potential for ecological research, conservation issues, and in the study of invasive species (Loiselle et al. 2003; Suarez and Tsutsui 2004). Several recent studies on the invasive Argentine ant, *Linepithema humile*, have relied on museum collections to track its dispersal and model the potential distribution across local, regional and global scales (Suarez et al. 2001; Roura-Pascual et al. 2004; Ward et al. 2005). Museum records are particularly useful because the records consist of individual point-locality information, which are readily transferable as input data for species distribution modelling methods. However, museum records represent presence-only data. There is almost always no information on where a species is absent. This represents several drawbacks for modelling (Zaniewski et al. 2002); absence data is a

**Fig. 6** The potential geographic distribution of six invasive species in New Zealand. Top, from left to right: *Iridomyrmex* sp., *Ochetellus glaber*, *Paratrechina* sp. Bottom, from left to right: *Pheidole rugosula*, *Technomyrmex albipes*, *Tetramorium grassii*. Each habitat suitability map is based on a DOMAIN model and score. Species presence points are marked by white dots



necessary component of many modelling methods, there may be unknown biases associated with ad hoc or non-systematic data samples, and rare species are often disproportionately present in presence-only records. These generally make presence-only data more difficult to model than systematically gathered presence–absence data (Zaniewski et al. 2002).

However, results from a recent comprehensive comparison of modelling methods found that, although presence-absence data generally outperform presence-only methods, models with presence-only were sufficiently accurate for modelling potential species' distributions and thus for applied use (Elith et al. 2006). Furthermore, there can also be problems with obtaining accurate absence data, especially when the study species is mobile or cryptic (Guisan and Thuiller 2005). These issues are particularly important for ant species, which may frequently move nesting sites and are often very cryptic, with nests underground or under logs. In addition, many ant species in temperate regions show a strong seasonal activity pattern, with less (or no) workers actively foraging in colder periods. Sampling effort also plays an important role in determining whether or not a absence is accurate. In this study, a  $1 \times 1$  km grid was used to classify either presence or absence of a species, however, ant sampling is typically undertaken over a much smaller area (e.g. a  $20 \times 20$  m grid). Thus, insufficient effort or inappropriate sampling can relatively easily result in a false absence.

There are also theoretical reasons for the justification of using presence-only modelling. Presence-only modelling is strongly linked with the fundamental niche of a species (Guisan and Zimmerman 2000; Phillips et al. 2006). Presence-only modelling determines potential habitat suitability, the inclusion of absence data will restrict habitat suitability as the result of historical restrictions, dispersal limitations, extinction and biological interactions (Anderson et al. 2003). Absence data is more useful in determining the realised niche (Guisan and Zimmerman 2000; Anderson et al. 2003; Phillips et al. 2006). Modelling the fundamental niche is more appropriate for invasive species, which may be less restricted by biotic interactions, and because pest manage-

ment authorities are interested in determining the 'maximum' potential distribution of an invasive species. Furthermore, in New Zealand, determining the fundamental niche of invasive ant species may equate closely to the realised niche. This is a consequence of New Zealand having very few native ant species (Ward 2005), and thus invasive ant species are unlikely to be limited by competition with native species.

#### Model performance and climatic variables

In this study, two measurements of model performance were examined. For each species, MAXENT and DOMAIN consistently perform better than BIOCLIM, with higher average AUC scores and lower omission error. Although AUC scores provide a single measure of performance that is independent of a classification threshold, for invasive species omission error also needs to be given significant consideration. For invasive species, high omission errors are considered to be a serious flaw (Guisan and Thuiller 2005), as they result in areas being classified as climatically unsuitable when they are not. Hence, the importance given to omission error in this study.

There are several recent studies which have compared modelling methods involving BIOCLIM, DOMAIN or MAXENT. Loiselle et al. (2003) compared five methods, including BIOCLIM and DOMAIN, to assess the conservation of 11 bird species in Brazil. DOMAIN models were amongst the best performing models, with the highest kappa values, low false-positives and included the greatest number of key areas in reserve designs. In contrast, BIOCLIM performed relatively poorly (Loiselle et al. 2003). In a major comparison of modelling methods, regions and taxa, Elith et al. (2006) reported a general progression of performance (poor to best) from BIOCLIM to DOMAIN to MAXENT (Elith et al. 2006). MAXENT was consistently one of the best performing models.

In this study, BIOCLIM under-predicted the potential distribution of invasive ant species in New Zealand. BIOCLIM had substantially higher omission error than DOMAIN and MAXENT, even when a number of different thresholds were examined within BIOCLIM. A major criticism of

BIOCLIM is how locality records and climatic variables are characterised into an ‘environmental envelope’ in Euclidean space (Carpenter et al. 1993; Kriticos and Randell 2001). As more climatic variables are added, a progressively smaller potential distribution occurs. Consequently, BIOCLIM may tend to under-predict the potential distribution species. In this study, the inclusion of fewer variables led to a larger potential distribution area and thus improved the omission error of models (as more actual presence records were encompassed). Therefore, the number of climate variables included in a model is an important consideration because using too few, or too many may result in incorrect predictions (Beaumont et al. 2005).

Choosing the right climate variables based on the biology of the study species also plays an important role in robust modelling (Beaumont et al. 2005; Guisan and Thuiller 2005). Although the number of environmental variables currently available as digital environmental layers is relatively few, they provide many of the variables that are strongly thought to commonly influence species macro-distributions (Anderson et al. 2003). Several climate variables are known to play a key role in the biology of ant species (Kaspari et al. 2000; Vega and Rust 2001; Holway and Suarez 2006). At large spatial scales, the tolerances of ant species are generally correlated with climate and major habitat (Hölldobler and Wilson 1990). Microhabitat specialisation and inter-specific competition play an increasing role at finer scales.

#### Distribution of exotic ants in New Zealand

Detailed studies of newly arrived exotic species and their subsequent spread are greatly lacking in the invasion biology literature (Puth and Post 2005). The continued study of these early introductions may prove useful insights into the process of invasion for ant species. This is the first time such a time lag has been described for exotic ant species, and suggests there is a considerable time lag in their spread. For example, it has taken many species several decades (40–60 years) to obtain a distribution of 5–8 Crosby regions (17–25% of all regions). Even after 100 years, the number of Crosby regions occupied is ~50% of maximum.

There are several caveats to this approach; primarily, the reliability of the year of introduction, the completeness of distribution records, and whether all species would reach the maximum number of Crosby regions. It is also likely that some species will spread faster than others, as a consequence of human-mediated spread. The Argentine ant (point x, y; 1990, 11; Fig. 2) is one such species (Ward et al. 2005). A recent analysis of weeds (exotic plants) in New Zealand has also shown a linear increase in distribution with the number of years since the species naturalised (Williams and Cameron 2006). Such data suggests that it takes most naturalised plants more than a century after naturalisation to appear in all environmentally suitable regions (Williams and Cameron 2006).

The six ant species modelled are predicted to be ubiquitous in the northern regions of the North Island, although areas with large stands of Kauri forest appear not to be suitable. Coastal lowland regions of the North Island are highly suitable for all species, although *Technomyrmex albipes* and *Ochetellus glaber* have the potential to extend inland and inhabit considerable areas of the middle and lower North Island. The suitability of the South Island is low for *Iridomyrmex* sp. and *Tetramorium grassii*, and these species should remain restricted to the very upper regions of the South Island. However, the other four species have the potential to inhabit sizeable areas of the South Island, particularly the eastern lowland (drier) areas of Canterbury. These distributions correspond very well to an intolerance of cooler mountainous regions. The distribution of these species in the South Island is more problematic because there are fewer occurrence records and there is greater uncertainty whether these records are permanent self-sustaining populations.

Comparing the number of Crosby regions of current and predicted distribution suggests that each species is currently distributed in only half the number of regions it could potentially inhabit. Thus, it will take many more decades before these species have reached an equilibrium in their regional-scale distribution (assuming these species are not already present in these areas but are undetected).

## Conclusions

Exotic ant species are potentially a significant threat to the New Zealand biota, which has evolved in the absence of a large and dominant native ant fauna. Thus, the continued study of exotic ants in New Zealand is warranted, particularly the potential distribution of species and their ecological impacts. At present there is no information regarding the ecological impact of exotic ant species in New Zealand and scant information on other aspects of their biology. Several avenues exist to improve the accuracy and value of habitat suitability maps for exotic ant species in New Zealand. The inclusion of soil moisture and temperature information are likely to be particularly useful as this has been an important variable in *Solenopsis invicta* modelling (Korzukhin et al. 2001), but such information is still being developed for New Zealand. Overlays of other environmental data could also prove useful to improve predictions, particularly vegetation coverage. However, there is also the need for widespread surveys for the presence of exotic ant species in several areas of the country. Such information will help evaluate modelling performance and also reduce apparent commission error, areas where there are no records but the species is actually present. Colony level information on the development of different life stages and nesting behaviour would also greatly assist in determining the environmental tolerances of exotic ant species.

**Acknowledgements** Many thanks to Richard Harris for access to database records, and to Steve Phillips and Robert Hijmans for their assistance with software. Thanks to Jacqueline Beggs, Mick Clout and Margaret Stanley for comments on earlier versions of this manuscript. This work was supported by the Entomological Society of New Zealand, Landcare Research (FRST C09X0507), and a FRST doctoral scholarship.

## References

- Anderson RP, Lew D, Peterson AT (2003) Evaluating predictive models of species' distributions: criteria for selecting optimal models. *Ecol Modell* 162:211–232
- Beaumont LJ, Hughes L, Poulsen M (2005) Predicting species distribution: use of climatic parameters in BIOCLIM and its impact on predictions of species' current and future distributions. *Ecol Modell* 186:250–269
- Brown WL (1958) A review of the ants of New Zealand. *Acta Hymenopterologica* 1:1–50
- Carpenter G, Gillison AN, Winte J (1993) DOMAIN: a flexible modeling procedure for mapping potential distributions of plants, animals. *Biodivers Conserv* 2:667–680
- Christian CE (2001) Consequences of a biological invasion reveal the importance of mutualism for plant communities. *Nature* 413:635–638
- Clarke KR, Warwick RM (2005) Change in marine communities: an approach to statistical analysis and interpretation, 2nd edn. Plymouth Marine Laboratory
- Crosby TK, Dugdale JS, Watt JC (1998) Area codes for recording specimen localities in the New Zealand subregion. *N Z J Zool* 25:175–183
- Drake JA, Mooney HA, di Castri F, Groves RH, Kruger FJ, Rejmanek M, Williamson M (1989) Biological invasions: a global perspective. John Wiley and Sons, New York
- Elith J, Graham CH, Anderson RP, Dudík M, Ferrier S, Guisan A, Hijmans RJ, Huettmann F, Leathwick JR, Lehmann A, Li J, Lohmann LG, Loiselle BA, Manion G, Moritz C, Nakamura M, Nakazawa Y, Overton J.McC, Peterson AT, Phillips SJ, Richardson K, Scachetti-Pereira R, Schapire RE, Soberón J, Williams S, Wisz MS, Zimmermann NE (2006) Novel methods improve prediction of species' distributions from occurrence data. *Ecography* 29:129–151
- Fielding AH, Bell JF (1997) A review of methods for the assessment of prediction errors in conservation presence/absence models. *Environ Conserv* 24:38–49
- Guisan A, Thuiller W (2005) Predicting species distribution: offering more than simple habitat models. *Ecol Lett* 8:993–1009
- Guisan A, Zimmermann NE (2000) Predictive habitat distribution models in ecology. *Ecol Modell* 135:147–186
- Hartley S, Lester PJ (2003) Temperature-dependent development of the Argentine ant, *Linepithema humile* (Mayr) (Hymenoptera: Formicidae): a degree-day model with implications for range limits in New Zealand. *New Zeal Entomol* 26:91–100
- Hijmans RJ, Cameron SE, Parra JL, Jones PG, Jarvis A (2005) Very high resolution interpolated climate surfaces for global land areas. *Int J Climatol* 25:1965–1978
- Hölldobler B, Wilson EO (1990) The ants. Harvard University Press, Cambridge
- Holway DA, Suarez AV (2006) Homogenization of ant communities in Mediterranean California: the effects of urbanisation and invasion. *Biol Conserv* 127:319–326
- Holway DA, Lach L, Suarez AV, Tsutsui ND, Case TJ (2002) The causes and consequences of ant invasions. *Annu Rev Ecol Syst* 33:181–233
- Landcare Research (2006) New Zealand ant distribution data. Accessed April 2006 <http://www.landcarere-search.co.nz/research/biosecurity/stowaways/Ants/distribution/index.asp>
- Loiselle BA, Howell CA, Graham CH, Goerck JM, Brooks T, Smith KG, Williams PH (2003) Avoiding

- pitfalls of using species distribution models in conservation planning. *Conserv Biol* 17:1591–1600
- Kaspari M, Alonso A, O'Donnell S (2000) Three energy variables predict ant abundance at a geographical scale. *Proc R Soc Lond B* 267:485–489
- Korzukhin MD, Porter SD, Thompson LC, Wiley S (2001) Modelling temperature-dependent range limits for the fire ant *Solenopsis invicta* (Hymenoptera: Formicidae) in the United States. *Environ Entomol* 30:645–655
- Kriticos DJ, Randell RP (2001) A comparison of systems to analyse potential weed distribution. In: Groves RH, Panetta FD, Virtue JG (eds) *Weed risk assessment*. CSIRO Publishing, Collingwood
- Mack RN, Simberloff D, Lonsdale WM, Evans H, Clout MN, Bazzaz FA (2000) Biotic invasions: casues, epidemiology, global consequences, and control. *Ecol Appl* 10:689–710
- Meynecke JO (2004) Effects of global climate change on geographic distributions of vertebrates in North Queensland. *Ecol Modell* 174:347–357
- McGlynn TP (1999) The worldwide transfer of ants: geographical distribution and ecological invasions. *J Biogeogr* 26:535–548
- Mooney HA, Drake JA (eds) (1986) *Ecology of biological invasions of North America and Hawaii*. Springer-Verlag, New York
- Morrison LW, Porter SD, Daniels E, Korzukhin MD (2003) Potential global range expansion of the invasive fire ant, *Solenopsis invicta*. *Biol Invasions* 6:183–191
- Nix H (1986) A biogeographic analysis of Australian Elapid snakes. In: Longmore R (ed) *Snakes: atlas of Elapid snakes of Australia*. Bureau of Flora and Fauna, Canberra
- O'Dowd DJ, Green PT, Lake PS (2003) Invasional 'meltdown' on an oceanic island. *Ecol Lett* 6:812–817
- Peterson AT (2003) Predicting the geography of species' invasions via ecological niche modeling. *Q Rev Biol* 78:419–433
- Phillips SJ, Dudik M, Schapire RE (2004) A maximum entropy approach to species distribution modelling. In: *Proceedings of the 21st International Conference on Machine Learning*. ACM Press, New York
- Phillips SJ, Anderson RP, Schapire RE (2006) Maximum entropy modeling of species geographic distributions. *Ecol Modell* 190:231–259
- Pimm SL, Bartell DP (1980) Statistical model for predicting range expansion of the Red Imported Fire Ant, *Solenopsis invicta*, in Texas. *Environ Entomol* 9:653–658
- Puth LM, Post DM (2005) Studying invasion: have we missed the boat?. *Ecol Lett* 8:715–721
- Roura-Pascual N, Suarez AV, Gómez C, Pons P, Touyama Y, Wild AL, Peterson AT (2004) Geographical potential of Argentine ants (*Linepithema humile* Mayr) in the face of global climate change. *Proc R Soc Lond B* 271:2527–2535
- Sandland OT, Schei PJ, Viken A (eds) (1999) *Invasive species and biodiversity management*. Kluwer Academic Publishers, Dordrecht
- Segurado P, Araújo MB (2004) An evaluation of methods for modelling species distributions. *J Biogeogr* 31:1555–1568
- Stoker RL, Ferris DK, Grant WE, Folse LJ (1994) Simulating colonization by exotic species: a model of the red imported fire ant (*Solenopsis invicta*) in North America. *Ecol Modell* 73:281–292
- Suarez AV, Tsutsui ND (2004) The value of museum collections for research and society. *Bioscience* 54:66–74
- Suarez AV, Holway DA, Case TJ (2001) Patterns of spread in biological invasions dominated by long-distance jump dispersal: Insights from Argentine ants. *Proc Natl Acad Sci* 98:1095–1100
- Suarez AV, Holway DA, Ward PS (2005) The role of opportunity in the unintentional introduction of non-native ants. *Proc Natl Acad Sci* 102:17032–17035
- Sutherst RW, Maywald G (2005) A climate model of the Red Imported Fire Ant, *Solenopsis invicta* Buren (Hymenoptera: Formicidae): implications for invasion of new regions, particularly Oceania. *Environ Entomol* 34:317–335
- Télez-Valdés O, Dávila-Aranda P (2003) Protected areas and climate change: a case study of the Cacti in the Tehuacán-Cuicatlán Biosphere Reserve, México. *Conserv Biol* 17:846–853
- Vega SJ, Rust MK (2001) The Argentine ant—a significant invasive species in agricultural, urban and natural environments. *Sociobiology* 37:3–25
- Ward DF (2005) Changes to the classification of ants (Hymenoptera: Formicidae). *The Weta* 30:16–18
- Ward DF, Harris R (2005) Invasibility of native habitats by Argentine Ants, *Linepithema humile*, in New Zealand. *N Z J Ecol* 29:215–219
- Ward DF, Beggs JR, Clout MN, Harris RJ, O'Connor S (2006) The diversity and origin of exotic ants arriving to New Zealand via human-mediated dispersal. *Diversity and Distributions*. 12:601–609
- Ward DF, Harris RJ, Stanley MC (2005) Human-mediated range expansion of Argentine ants *Linepithema humile* (Hymenoptera: Formicidae) in New Zealand. *Sociobiology* 45:401–407
- Williams DF (1994) *Exotic ants: biology, impact, and control of introduced species*. Westview Press, Boulder Colorado
- Williams PA, Cameron EK (2006) Creating gardens: the diversity and progression of European plant introductions. In: Allen RB and Lee WG (eds) *Biological invasions in New Zealand*. Springer-Verlag, London
- Zaniewski AE, Lehmann A, Overton JMcC (2002) Predicting species spatial distributions using presence-only data: a case study of native New Zealand ferns. *Ecol Modell* 157:261–280