Identifying hotspots of alien plant naturalisation in Australia: approaches and predictions.

Aaron J Dodd[1,2,*], Michael A McCarthy[3], Nigel Ainsworth[4] and Mark A Burgman[1]

[1] Centre of Excellence for Biosecurity Risk Analysis, School of BioSciences, The University of Melbourne, Parkville, 3010, Australia

[2] Victorian Department of Economic Development, Jobs, Transport and Resources, 475 Mickleham Rd, Attwood, 3049, Australia;

* Corresponding author: Tel.: +61 3 9217 4333; Fax.: +61 3 9217 4322; E-mail: aaron.dodd@ecodev.vic.gov.au

[3] Centre of Excellence for Environmental Decisions, School of BioSciences, The University of Melbourne, Parkville, 3010, Australia;

[4] Victorian Department of Economic Development, Jobs, Transport and Resources, 1 Spring Street, Melbourne, Victoria 3000 , Australia;

Article type: Original Paper

Word count: 8,233 *Number of tables and figures: 2 + 4* *Number of references:* 108

(A) **ACKNOWLEDGEMENTS**

The authors would like to acknowledge Jane Catford, Jane Elith, Frith Jarrad, Petr Pysek, John Wilson and an additional anonymous reviewer for their insightful comments on earlier versions of this manuscript. Alison Vaughan, Niels Klazenga and Anna Monro helped us understand the nuances of both AVH and APC. Jane Elith provided advice regarding point process models. This research was supported by a Victorian Life Sciences Computation Initiative (VLSCI) grant [VR0284] on its Peak Computing Facility at the University of Melbourne, an initiative of the Victorian Government, Australia. This research was also supported by an Australian Research Council (ARC) Future Fellowship to M.M. and the ARC Centre of Excellence for Environmental Decisions.

(A) **ABSTRACT**

The early detection of newly naturalised alien species is vital to ensuring the greatest chance of their successful eradication. Understanding where species naturalise most frequently is the first stage in allocating surveillance effort. Using Australia's Virtual Herbarium, we compiled the collection records for all plant species in Australia. We controlled for potential spatial biases in collection effort to identify areas that have an elevated rate of first records of alien species' occurrence in Australia. Collection effort was highly variable across Australia, but the most intense collection effort occurred either close to herbaria (located in cities) or in remote natural environments. Significant clusters of first records of occurrence were identified around each state's capital city, coinciding with higher collection effort. Using Poisson point process modelling, we were able to determine the relative influence of environmental and anthropogenic factors on the spatial variation in the risk of species naturalisation. Effort-corrected naturalisation risk appeared to be strongly related to land use, road and human population densities, as well as environmental factors such as average temperature and rainfall. Our paper illustrates how the risk of naturalisation at a location can be estimated quantitatively. Improved understanding of factors that contribute to naturalisation risk enhances allocation of surveillance effort, thereby detecting novel species sooner, and increasing the likelihood of their eventual eradication.

Keywords alien flora, herbarium, invasive plants, pathway analysis, sampling bias, surveillance.

(A) **INTRODUCTION**

The human mediated exchange of species is an important economic, social and environmental issue (Pfeiffer 2008; Pimentel et al. 2005; Vilà et al. 2011). Despite considerable investment in pre-border biosecurity measures such as import regulations tied to weed risk assessment (Andersen et al. 2004b; Gordon et al. 2008; Pheloung et al. 1999), alien plant species richness continues to increase in all regions across the globe (Dodd et al. 2015b; Lambdon et al. 2008; Maroyi 2012; Rojas-Sandoval and Acevedo-Rodríguez 2015; Wu et al. 2004). Whilst many of the recently detected species likely arrived prior to the strengthening of biosecurity management approaches internationally (Caley et al. 2008; Phillips et al. 2010), the ongoing entry and establishment of new species requires management agencies to continue to improve their biosecurity practices (Cacho and Hester 2011; Hester et al. 2013; Moore et al. 2010).

Preventing species introduction in the first instance is often the most cost effective way of managing the impact of alien species (Finnoff et al. 2007; Kim et al. 2006). However, preventing species exchange altogether is not practical in modern economies (see Hulme 2009; Virtue et al. 2004) and is, in fact, prohibited without technical justification under the World Trade Organisation Agreement on the Application of Sanitary and Phytosanitary Measures (the SPS Agreement). The majority of invasive species research has, therefore, focussed on questions relating to either species' invasiveness (Elton 1958; Hayes and Barry 2008; Kolar and Lodge 2001; Rejmánek and Richardson 1996; Sakai et al. 2001; Williamson and Fitter 1996) or community invasibility (Catford et al. 2011; Davis et al. 2000; Fridley et al. 2007; Levine and D'Antonio 1999; Lonsdale 1999; Tilman 1997), because these questions are vital in underpinning decision making regarding the regulation of species both pre- and post-border (Andersen et al. 2004a; Hulme 2012; Simberloff 2003; Weber et al. 2009).

Once a species has established a population, eradication (sensu Newsom 1978) might avoid future impacts most cost-effectively despite its high expense (Harris et al. 2001; Wittenberg and Cock 2001). Several factors influence the relative cost effectiveness of eradication (Dodd et al. 2015a; Hester et al. 2013; Panetta 2009), but the area occupied by the species when first detected is particularly important (Pluess et al. 2012; Rejmanek and Pitcairn 2002) as treatment cost and the probability of success both depend on the size of the infested area. This implies that early detection of a species increases the likelihood of successful eradication (Harris et al. 2001; Myers et al. 2000; Simberloff 2003).

However, few published studies focus on detecting alien plant species in the early stages of naturalisation (Huang et al. 2012; Sullivan et al. 2004). The vast majority of research on surveillance design has focussed on economic optimisation (Bogich et al. 2008; Epanchin-Niell et al. 2012; Hauser and McCarthy 2009) or maximising detection (Baxter and Possingham 2011; Kaplan et al. 2014; Moore et al. 2014), and usually in the context of a single species programme (see Hauser and McCarthy 2009; Kaplan et al. 2014). The relative scarcity of published surveillance methods pertaining to multi-species naturalisation likely reflects a lack of spatial data (Genovesi 2005; Huang et al. 2012). Traditionally, most of the information (where it exists) is held by management agencies, although this is changing (see Tingley et al. 2014).

Herbaria play a vital role in the detection and consequent identification of newly naturalised plant species (Pyšek et al. 2013; Pyšek et al. 2004; Stajsic and Vaughan 2007). In Victoria, Australia, 76% of newly naturalised plant species detected between 2004 and 2007 were first detected by herbarium staff (Stajsic and Vaughan 2007), with all other species reported as transitioning from casual to naturalised (sensu Blackburn et al. 2011) during that period also lodged with herbaria. The recent digitisation of herbarium records, and their collation into online databases such as Australia’s Virtual Herbarium (CHAH 2014a), makes these records relatively accessible (Lavoie 2013). By reconstructing the histories of alien species from historical records it may be possible to infer the circumstances facilitating their naturalisation (see Huang et al. 2012; Sullivan et al. 2004) and to tailor surveillance programmes aimed at finding them.

Here we aim to identify hotspots of species naturalisation and to explore the influence of environmental and anthropogenic factors in creating them. Using Australia's Virtual Herbarium (CHAH 2014a), we compile records of species collection effort for the entire flora of Australia. We then control for potential spatial biases in collection effort (sensu Lavoie et al. 2012; Moerman and Estabrook 2006; Rich 2006; Yang et al. 2014) to identify areas that have an elevated rate of first records of species occurrence. Finally, we analyse these data to determine the influence of environmental and anthropogenic factors on the risk of species naturalisation. By analysing this large dataset, we aim to identify robust patterns that can inform future surveillance effort targeting newly-naturalised alien species.

(A) **METHODS**

We used the Australian Plant Census to identify the naturalised alien plant species in Australia's Virtual Herbarium database. We then condensed the dataset, from species records to collection events, to remove duplicate information and control for spatial biases. Finally, we used Poisson point process modelling to determine the influence of environmental and anthropogenic factors on the spatial variation in the risk of species naturalisation.

(B) **Data collection**

The complete database of catalogued flora records held by Australian herbaria was exported from Australia's Virtual Herbarium (AVH) (CHAH 2014a). Records in this master dataset were filtered to include only correctly georeferenced records of vascular plants (tracheophytes) collected in Australia and identified to at least the rank of species. After filtering, 3,074,544 species records were exported for analysis. Records were then imported into the R software environment for statistical computing and graphics (R Core Team 2013) where the naturalisation status of each species [AVH processed name] was sourced from the Australian Plant Census (APC) (CHAH 2014b).

The AVH 'processed name' indicates the currently accepted APC species name for a specimen derived from the original 'supplied name' provided by the holding herbaria. This matching process ensures that each record in AVH is systematically checked for taxonomic issues such as invalidity, synonymy and re-classification therefore ensuring the record reflects the species' current, nationally agreed, taxonomy. All families except the Orchidaceae (which has one naturalised species) have been treated by APC, with extensive updates in 2014 (CHAH 2014b). APC follows a similar framework to Blackburn et al. (2011), with alien species regarded as 'naturalised' once they form self-sustaining populations. 'Casual' species (those with surviving individuals but not self-sustaining populations) and 'cultivated' species were not considered in this analysis as they have not been uniformly treated by herbaria.

Using the APC naturalisation status, the master dataset was then split into three subsets for analysis. The first subset, hereafter referred to as the 'native' species, contained 2,737,139 records of 20,376 species not considered naturalised by APC and therefore native to Australia. The second subset, hereafter referred to as the 'naturalised' species, contained 237,541 records of 2,699 species considered naturalised by APC and therefore 'alien' to Australia. The third subset, hereafter referred to as the 'first record of occurrence' dataset, is a subset of the naturalised dataset containing only the earliest herbarium record for each of the 2,699 naturalised species in Australia. Importantly, this dataset does not imply the exact date or location of introduction nor naturalisation (which are unknown), merely the first specimen held by herbaria. The first records dataset is analogous to those used in Pyšek et al. (2003) and Sullivan et al. (2004) and identical to Dodd et al. (2015b).

Where multiple species records were collected during a visit to a location, these were aggregated into a single 'collection event'. By removing these duplicate records we ensured that the events are independent, satisfying the assumptions of parametric analyses. This reduced the 3,074,544 species records to 1,273,445 unique collection events. For the purposes of our analysis, we used the number of discrete collection events to represent collection effort; similar to Hyndman et al. (2015) and Yang et al. (2014). One (or more) first records of occurrence were detected during 2,555 collection events and these events are hereafter referred to as 'detection events'. Because herbarium collection is known to be spatially biased and largely unstructured (Moerman and Estabrook 2006; Rich 2006), individual detection events are likely to be separated in space and/or time from the true first instance of naturalisation. However, when aggregated together and adjusted for these biases, we expect that the frequency of detection events at a spatial location is likely to be a robust indicator of the true naturalisation risk.

(B) **Spatio-temporal distribution of events**

Locations of the 2,699 first records of occurrence were initially examined visually, by plotting the data at the Australian scale. Points were then projected from geographic coordinates to the Australian Albers (equal area conic) coordinate system (EPSG: 3577) and formatted as point pattern objects to facilitate their analysis using the spatstat R package (Baddeley and Turner 2005; R Core Team 2013). In order to analyse the data as a single cohort, we first needed to rule out significant temporal variation in the point pattern. Using Ripley's Cross K-Function (Baddeley et al. 2000) we found that the first records of separate temporal cohorts were significantly (Monte Carlo p-values <0.02) clustered together, implying a dependence on spatial covariates. As these spatial patterns were clustered (rather than independent or dispersed) over time, we considered it appropriate to analyse the data as a single cohort.

The first step towards identifying areas with an elevated risk of naturalisation is determining the spatial distribution of both detection and collection events. These were estimated by calculating the kernel smoothed intensity (points per unit area) of the point pattern (Diggle 1985). The smoothing kernel was edge corrected, with the bandwidth automatically selected by cross-validation according to Berman and Diggle (1989). We qualitatively chose 10 km x 10 km pixels as the scale for the analysis of intensity as this was large enough to minimise issues caused by points with low spatial accuracy, whilst small enough to still allow for meaningful inference.

(B) **Spatial statistics**

Once the spatial intensity of detection and collection events had been identified, we trialled two approaches to identify areas with higher than expected frequencies of detection events given collection effort. The first was a spatial scan test for clustering in a spatial point pattern (Kulldorff 1997). A binomial scan test identifies locations that have a significantly higher than expected proportion of points of one type (i.e. detection events) within a set radius given the spatial location of all points (collection events) in the pattern (Baddeley and Turner 2005; Kulldorff 1997). The statistical significance of a cluster is indicated by the likelihood ratio test statistic of its spatial location as calculated by simulation. For our analysis the radius was set to 10 km with the Monte Carlo p-value calculated from 10,000 simulations.

The second approach was to calculate the spatially-varying probability of detection events (Bivand et al. 2008; Kelsall and Diggle 1995) using spatstat's relative risk function (Baddeley and Turner 2005). In our context, the probability was calculated by comparing the intensity of detection events with the intensity of collection events at the pixel scale. As with the intensity estimates above, the smoothing kernel was edge corrected, with the bandwidth automatically selected by cross-validation according to Berman and Diggle (1989). The pixel size was increased to 20 km x 20 km for this analysis to avoid over-fitting in areas with unusually low collection effort. This also aligned the scales of the two spatial statistics (both to 20 km widths) making their results directly comparable.

(B) **Model construction and evaluation**

To determine the influence of environmental and anthropogenic factors on the spatial variation in the risk of species naturalisation, we constructed a series of 5 km x 5 km raster grids for factors previously suggested to influence naturalisation success. Only Sullivan et al. (2004) and Moodley et al. (2014) have looked directly at plant naturalisation, so we used a selection of factors previously found to be influential in studies of gross [country-scale] naturalisation rates (Dalmazzone 2000; Hulme 2009; Pyšek et al. 2010; Vila and Pujadas 2001), invasibility (Catford et al. 2011; Huang et al. 2012; Lin et al. 2007) and herbarium biases (Lavoie et al. 2012; Moerman and Estabrook 2006; Rich 2006; Yang et al. 2014) to ensure an adequate range of variables (Table 1).

We used inhomogenous Poisson Point Process Modelling (Renner et al. 2015) to test the relationship between the explanatory variables and the intensity of detection events. Poisson point process modelling (Poisson PPM) differs from conventional Poisson regression in that it models the response variable's intensity rather than its count (Baddeley and Turner 2006; Renner et al. 2015). Poisson PPM is particularly well suited to this application because it can include an exposure variable (such as the intensity of collection events) as an offset, allowing us to model the rate of detection events per unit of collection effort (Baddeley and Turner 2006). Hyndman et al. (2015) used a similar approach to control for collection effort.

Prior to model fitting, the explanatory variables were examined using a scatterplot matrix to identify pairwise correlations. Average annual days of frost and average daily solar exposure were excluded from the dataset due to their strong (r>0.7) Pearson correlation with average yearly rainfall and average daily mean temperature respectively. Human population density was log-transformed to enable us to sensibly fit a linear response to it. Due to our use of the log link function, the intensity of collection events was calculated at 30 km x 30 km pixels (an additional increase in scale) to prevent over-fitting the more complex model where the intensity of collection events was close to zero.

Models were fitted with maximum likelihood using the Berman-Turner approximation in the spatstat package (Baddeley and Turner 2005). The full model (containing all explanatory variables and the offset) was trained using the data rich south-eastern states of Australia (Victoria, New South Wales, Tasmania and the Australian Capital Territory). Performance was evaluated using the Akaike Information Criterion (Burnham and Anderson 2002) with the best performing model selected using backwards stepwise selection. Checks of model assumptions, including spatial residuals (Baddeley et al. 2005), were also performed. Final model evaluation was undertaken by predicting the expected probability (and conditional intensity) for two adjacent states (Queensland and South Australia) and comparing the results with the observed pattern of first records of occurrence.

Unless otherwise specified, all data processing and analyses were undertaken in the R software environment for statistical computing and graphics (R Core Team 2013) with the following packages installed: reshape2 (Wickham 2007), plyr (Wickham 2011), data.table (Dowle et al. 2014) and doParallel (Revolution Analytics and Weston 2014) for data management; sp (Pebesma and Bivand 2005), rgdal (Bivand et al. 2013), rgeos (Bivand and Rundel 2013), maptools (Bivand and Lewin-Koh 2013), raster (Hijmans and van Etten 2013) and geostatsp (Brown 2015) for spatial operations; ggplot2 (Wickham 2009) and scales (Wickham 2012) for plotting. The R script used for the analysis is included in Appendix S1.

(A) RESULTS

(B) Spatio-temporal distribution of events

At the continental scale, the majority of the 2,699 first records of occurrence were located in the eastern half of the continent, close to the coast, with the greatest density of points in areas adjacent to capital cities (Figure 1). This spatial pattern remained relatively constant across 30-year cohorts, with significant clustering between temporal cohorts (data not shown). Queensland had the highest number of first records (673), followed by New South Wales (538), South Australia (502) and Victoria (444). In contrast, Western Australia had the highest number of collection events (345,176), followed by Queensland (253,310), New South Wales (216,261) and South Australia (177,600; data not shown).

In order to visualise the relatively fine spatial patterns (relative to Australia's large size), we use the state of Victoria here as a consistent basis for presenting our approach. As was observed at the continental scale (Figure 1), the spatial location of detection events was concentrated around the capital city (Melbourne) with fewer, scattered distributions elsewhere (Figure 2a). This pattern resulted in a kernel smoothed intensity (number of first records per square metre) ranging from 0 to 2.97×10^{-7} events m^{-2} (Figure 2b) with the most intense areas corresponding to those in Figure 2a. Collection effort was highly variable across the state (Figure 2c) with the intensity of collection events ranging from 0 to 9.18×10^{-6} events m^{-2} with more than twenty areas of elevated intensity (Figure 2d). These broad patterns were consistent across each of the states (data not shown).

(B) Hotspot (cluster) detection

Significant clusters of detection events were identified by the spatial scan test in areas adjacent to major human population centres across each of the states (Figure 2e). In most instances, clusters correlated with the highest intensity of detection events identified in Figure 2b and the results of the two approaches were spatially similar. The maximum value of the likelihood ratio test statistic for areas in the eastern states was between 200 and 300 (Figure 2e) with Monte Carlo p-values < 0.0001 indicating significantly higher proportions of detection events in these locations than expected given the spatial distribution of collection events. Although the number of detection events in each state reached into the hundreds (Figure 2a), the number of clusters identified was usually less than five (Figure 2e).

In contrast, the spatially-varying probability analysis (Figure 2f) indicated a more widespread distribution of elevated naturalisation risk. The probability of detection events ranged from 0 to 6% (1 in 17 events) compared to an overall base rate of 0.2% (1 in 500 events). Species were more likely to be detected along major roadways, particularly where roads intersected at rural townships. The risk of species detection was also higher around major urban centres (Figure 2f), coinciding with the clusters identified at those locations by the scan test (Figure 2e). This trend remained even after their higher collection effort (Figures 2c-d) had been accounted for.

(B) **Evaluation of the factors influencing hotspot location**

The best performing model included all ten of the candidate variables and the offset for collection effort, although the influence of proximity to ports and watercourses was not significant (Table 2). Human population and road densities, land use intensity, average rainfall and proximity to railways all increased the probability of a detection event. Conversely, proximity to herbaria, average temperature and increasing Normalised Difference Vegetation Index (NDVI) negatively influenced the probability. The estimated model coefficient for each parameter and its Z score is included in Table 2.

Using the model trained on events in the four south-eastern states, we predicted the probability of detection events across the adjacent states of South Australia (Figure 3b) and Queensland (Figure 4b). In both states, the expected probability ranged from 0% to 3%, with the highest probability found in areas along the coastline and adjacent to capital cities similar to the broad patterns identified in Figure 1. We were unable to formally test these predictions using conventional methods (chi-square or Kolmogorov–Smirnov tests) due to the unavoidable presence of zeroes in both the observed and predicted patterns.

Notwithstanding, the observed first records in both states (Figures 3a & 4a) nearly all occurred in areas predicted to have an elevated probability of detection events (Figures 3b & 4b). The model appeared to over-estimate the probability of detection events in central areas of South Australia and under-estimate them in far-north Queensland near Cairns. However, these discrepancies disappeared when collection effort was used to predict the intensity of detection events (effectively re-biasing the predictions), indicating that collection effort has likely influenced the observed patterns of naturalisation (Figures 3c & 4c).

(A) **DISCUSSION**

By modelling plant naturalisation as a Poisson point process our analysis consistently showed that alien plants are most frequently first recorded around major human population centres, even when corrected for collection (sampling) effort. We demonstrate how the rate of first detection is significantly influenced by anthropogenic factors such as human population and road density, land use and distance to railways as well as environmental factors such as average rainfall and temperature. Finally, our paper illustrates how the risk of naturalisation can be estimated spatially and discusses how these estimates could be used to inform surveillance programmes targeting newly naturalised alien species.

(B) **Separating spatial biases from naturalisation hotspots**

First records of occurrence were most frequently located in areas with the highest human influence. This is not at all surprising, given that alien species are by definition introduced (either accidentally or deliberately) by humans (Blackburn et al. 2011) and this basic relationship has previously been demonstrated by Sullivan et al. (2004) for New Zealand. In Australia, this has resulted in a pattern of first occurrence that largely follows the eastern coastal fringe with isolated events in the north and west consistent with the distribution of human settlement (Figure 1). These patterns have remained relatively constant over time, despite changing purposes of species introduction (Dodd et al. 2015b; Essl et al. 2011). This may indicate that any spatial changes are occurring at a finer scale than our analysis, or [because most introduction pathways are inherently associated with humans] that the spatial patterns of first occurrence are more heavily influenced by the patterns of human settlement than the specific pathway by which the species arrives.

However, collection effort is also known to be spatially biased towards areas close to where people, particularly botanists, live (Aikio et al. 2010; Moerman and Estabrook 2006; Rich 2006; Yang et al. 2014). Because the likelihood of detection increases with additional effort (Hauser and McCarthy 2009; Moore et al. 2014), it is important to determine whether these patterns of first occurrence reflect differences in propagule pressure, ecosystem invasibility, or collection effort (Aikio et al. 2010; Hulme 2012; Lavoie et al. 2012). As expected, our analysis clearly indicated that the intensity of collection effort was highly variable across our study area (Figures 2c-d), apparently similar to China (Yang et al. 2014) and the British Isles (Rich 2006), with elevated collection in natural areas and areas with high human population density respectively. One exception is the area of elevated collection around the rural town of Dimboola where the amateur botanist Felix M. Reader collected over 10,000 specimens in the 1890s. Whilst Reader didn't detect first occurrences at a substantially higher rate than average, the intensity of his collection effort resulted in the naturalisation hotspot found at Dimboola (Figures 2b & 2e) and is a clear demonstration of the 'botanist effect' (sensu Moerman and Estabrook 2006). Fortunately, collection effort is highly variable across the study area allowing us to effectively disentangle it from propagule pressure and human-mediated modification of the environment.

Of the two approaches we used to control for elevated collection effort, the spatially-varying probability estimates (Bivand et al. 2008; Kelsall and Diggle 1995) had the most potential for correctly guiding surveillance effort. Because the scan test (Kulldorff 1997) is designed only to detect clusters, it tended to reduce the observed occurrences into too few areas (clusters) with elevated naturalisation risk (Figure 2e). In comparison to the spatially varying probability estimates (Figure 2f), this would result in type II errors when allocating surveillance effort as many of the observed occurrences would not have been detected. Nonetheless, both approaches indicated that naturalisation risk was elevated in areas adjacent to major human population centres, even once controlled for collection effort (Figures 2e & 2f).

(B) **Factors influencing naturalisation risk**

Inhomogeneous Poisson point process modelling builds on the spatially varying probability estimates by considering the intensity of detection events to be a function of not just the intensity of collection events but also of spatial covariates (Baddeley and Turner 2006). One of the limitations of our analysis, as well as similar previous analyses (Yang et al. 2014), is that the collection events and some of the spatial covariate data necessarily relate to different time ranges (collection events have mostly occurred over the last 120 years whilst the anthropogenic data are from the last ten years). However, as discussed above, the observed patterns of naturalisation haven't changed substantially over time nor have the majority of our predictive variables such as ports, major roads and railways (which are drawn at the 1:1M scale). We therefore consider the influence of this limitation to be minimal.

Of the factors included in our regression model, anthropogenic variables had the greatest effect on the intensity of detection events (Table 2). As expected, human population density had the largest influence the probability of occurrence, reinforcing that species exchange is a human mediated phenomenon and that human activities facilitate alien species establishment (Blackburn et al. 2011; Pyšek et al. 2010; Sullivan et al. 2004). Commonly used indicators of human activity, such as road density and distance to railways, were also influential, reflecting their known importance as dispersal pathways, as well as their correlation with habitat disturbance (Catford et al. 2011; Lin et al. 2007; Vila and Pujadas 2001). It is therefore reasonable to conclude that proximity to these features also increases naturalisation risk.

Land use class (ABARES 2010) also significantly influenced the probability of occurrence (Table 2). As the intensity of the land use increased from 'conservation and natural environments', through increasing agricultural intensity (natural, dryland and irrigated) to 'intensive uses', the probability also increased. Increasing land use intensity is generally considered to be correlated with both increased disturbance and propagule pressure (Catford et al. 2011; Moodley et al. 2014; Vila and Pujadas 2001), so this result is intuitive. Indeed, the influence of these factors on alien species naturalisation [establishment] and invasion [spread] is generally agreed (Catford et al. 2011; Dalmazzone 2000; Hulme 2009; Moodley et al. 2014; Pyšek et al. 2010; Sullivan et al. 2004; Vila and Pujadas 2001).

The role of the environmental variables including rainfall and temperature was less clear. Our finding that rates of first occurrence were significantly higher in cooler and wetter climates (Table 2) supports the previous observation that temperate mainland areas are more invaded than tropical ones (Lonsdale 1999; Pyšek and Richardson 2006; van Kleunen et al. 2015). However, these climates also correspond with the most populous areas of human settlement in Australia (ABS 2012). As such, temperate regions also correspond with the most intense agriculture and the most dense road, rail and port networks. Analysing the data by biogeographic, rather than geo-political, boundaries may help understand the influence of climate on plant naturalisation. By identifying the first occurrences in each biogeographic region it may be possible to compare the rates of first occurrence and species-area relationships across environmental gradients at a finer scale than previously possible (see Huang et al. 2012; Pyšek et al. 2010; van Kleunen et al. 2015).

One advantage of point process models is that we can estimate the risk of naturalisation at a relatively fine scale. We were able to test the predictive value of our model by estimating the probability of detecting a first occurrence in the states adjacent to the study area. The model performed well in comparison to the observed events; indicating elevated risk in nearly all areas where first records of occurrence were observed and conversely, indicating reduced risk in areas where events were not detected (Figures 3 & 4). A small number of areas were either over- or under- predicted, although this is to be expected given the highly random nature of long distance dispersal events leading to naturalisation (Aikio et al. 2010). These deviations also largely disappeared once historical collection effort was included as a predictor, indicating that the observed spatial pattern of first occurrence in Australia is clearly influenced by collection effort.

(B) **Implications for designing surveillance programmes**

We believe our analysis is an important step towards the 'integrated risk maps' envisaged by Hulme (2009). However, the scale of our results is still higher than the underlying biological processes (Catford et al. 2009) and, as discussed above, may also be too coarse to detect the fine-scale temporal changes possibly occurring due to changing pathways of introduction (Essl et al. 2011). This is likely to be at least partly due to the unknown spatial and temporal separation between the first records of occurrence and the true first instance of naturalisation (Hyndman et al. 2015). Notwithstanding these limitations, the spatial patterns identified by our analysis appear robust enough to inform the spatial allocation of resources to general awareness activities and passive surveillance programs (Cacho and Hester 2011).

The logical extension of our work then is to look at finer scale patterns within identified areas of elevated risk, possibly in a multi-scale framework similar to Kaplan et al. (2014). Using Victoria as an example, we could select one of the identified hotspots (e.g. the north eastern suburbs of Melbourne or the agricultural area between Longerenong and Warracknabeal; Figure 2e-f) and re-scale the analysis down to just that area; potentially identifying the hotspots within the hotspot. It would also be interesting to separate species by their suspected purpose or pathway of introduction (Dodd et al. 2015b; Essl et al. 2011; Virtue et al. 2004) and contrast their spatial patterns.

The early detection of newly naturalised species is critical to ensure the best chance of successful eradication (Harris et al. 2001; Myers et al. 2000; Simberloff 2003). However, as we have demonstrated, the base rate of species detection through herbarium lodgement is extremely low (1 in >500 events) meaning finding new species can be like finding a needle in a haystack (Schmidt et al. 2010). Should agencies ultimately allocate surveillance resources to areas where the probability of detection is highest, with consideration to optimising costs and expected benefits (sensu Hauser and McCarthy 2009), our analysis indicates that there is potential to improve upon this default rate by tenfold.

(A) REFERENCES

ABARES (2010) Guidelines for land use mapping in Australia: principles, procedures and definitions. 7 edn. Australian Bureau of Agricultural and Resource Economics and Sciences, Canberra

ABARES (2014) Catchment scale land use of Australia. Australian Bureau of Agricultural and Resource Economics and Sciences. http://www.agriculture.gov.au/ABARES/aclump/pages/land-use/data-download.aspx. Accessed 09/02/2015

ABS (2012) Census of Population and Housing: Mesh Block Counts, 2011. Australian Government Bureau of Statistics. http://www.abs.gov.au/AUSSTATS/abs@.nsf/DetailsPage/2074.02011. Accessed 16/2/2015

Aikio S, Duncan RP, Hulme PE (2010) Herbarium records identify the role of long-distance spread in the spatial distribution of alien plants in New Zealand. Journal of Biogeography 37:1740-1751

Andersen MC, Adams H, Hope B, Powell M (2004a) Risk analysis for invasive species: General framework and research needs. Risk Analysis 24:893-900

Andersen MC, Adams H, Hope B, Powell M (2004b) Risk assessment for invasive species. Risk Analysis 24:787-793

Baddeley A, Turner R (2005) spatstat: An R Package for Analyzing Spatial Point Patterns. Journal of Statistical Software 12:1-42

Baddeley A, Turner R (2006) Modelling spatial point patterns in R. In: Baddeley A, Gregori P, Mateu Mahiques J, Stoica R, Stoyan D (eds) Case studies in spatial point process modeling. Lecture notes in statistics. Springer-Verlag, New York, pp 23-74

Baddeley A, Turner R, Møller J, Hazelton M (2005) Residual analysis for spatial point processes (with discussion). Journal of the Royal Statistical Society: Series B (Statistical Methodology) 67:617-666

Baddeley AJ, Møller J, Waagepetersen R (2000) Non-and semi-parametric estimation of interaction in inhomogeneous point patterns. Statistica Neerlandica 54:329-350

Baxter PW, Possingham HP (2011) Optimizing search strategies for invasive pests: Learn before you leap. Journal of Applied Ecology 48:86-95

Berman M, Diggle P (1989) Estimating weighted integrals of the second-order intensity of a spatial point process. Journal of the Royal Statistical Society: Series B (Statistical Methodology) 51:81–92

Bivand RS, Keitt T, Rowlingson B (2013) rgdal: Bindings for the Geospatial Data Abstraction Library. R package version 0.8-9. The Comprehensive R Archive Network, http://CRAN.R-project.org/package=rgdal

Bivand RS, Lewin-Koh N (2013) maptools: Tools for reading and handling spatial objects. R package version 0.8-23. The Comprehensive R Archive Network, http://CRAN.R-project.org/package=maptools

Bivand RS, Pebesma EJ, Gómez-Rubio V, Pebesma EJ (2008) Applied spatial data analysis with R. Use R! Springer, New York

Bivand RS, Rundel C (2013) rgeos: Interface to Geometry Engine - Open Source (GEOS). R package version 0.2-19. The Comprehensive R Archive Network, http://CRAN.R-project.org/package=rgeos

Blackburn TM et al. (2011) A proposed unified framework for biological invasions. Trends in Ecology and Evolution 26:333-339

Bogich TL, Liebhold AM, Shea K (2008) To sample or eradicate? A cost minimization model for monitoring and managing an invasive species. Journal of Applied Ecology 45:1134-1142

BoM (2015) Climate and oceans data and analysis. Australian Government Bureau of Meteorology. http://www.bom.gov.au/climate/data-services/. Accessed 16/02/2015

Brown PE (2015) Model-Based Geostatistics the Easy Way. Journal of Statistical Software 63:1-24

Burnham KP, Anderson DR (2002) Model selection and multimodel inference: a practical information-theoretic approach. Springer Science & Business Media, New York

Cacho OJ, Hester SM (2011) Deriving efficient frontiers for effort allocation in the management of invasive species. Australian Journal of Agricultural and Resource Economics 55:72-89

Caley P, Groves RH, Barker R (2008) Estimating the invasion success of introduced plants. Diversity and Distributions 14:196-203

Catford JA, Jansson R, Nilsson C (2009) Reducing redundancy in invasion ecology by integrating hypotheses into a single theoretical framework. Diversity and Distributions 15:22-40

Catford JA, Vesk PA, White MD, Wintle BA (2011) Hotspots of plant invasion predicted by propagule pressure and ecosystem characteristics. Diversity and Distributions 17:1099-1110

CHAH (2014a) Australia's Virtual Herbarium. Council of Heads of Australasian Herbaria. http://avh.ala.org.au/. Accessed 19/08/2014

CHAH (2014b) Australian Plant Census. Council of Heads of Australasian Herbaria. http://www.chah.gov.au/apc/index.html. Accessed 04/09/2014

Dalmazzone S (2000) Economic factors affecting vulnerability to biological invasions. In: Perrings C, Willamson M, Dalmazzone S (eds) The economics of biological invasions. Edward Elgar Publishing Limited, Cheltenham, pp 17-30

Davis MA, Grime JP, Thompson K (2000) Fluctuating resources in plant communities: a general theory of invasibility. Journal of Ecology 88:528-534

Diggle PJ (1985) A kernel method for smoothing point process data. Journal of the Royal Statistical Society: Series C (Applied Statistics) 34:138–147

Dodd AJ, Ainsworth N, Burgman MA, McCarthy MA (2015a) Plant extirpation at the site scale: implications for eradication programmes. Diversity and Distributions 21:151-162 doi:10.1111/ddi.12262

Dodd AJ, Burgman MA, McCarthy MA, Ainsworth N (2015b) The changing patterns of plant naturalisation in Australia. Diversity and Distributions 21:1038-1050 doi:10.1111/ddi.12351

Dowle M, Short T, Lianoglou S, Srinivasan A (2014) data.table: Extension of data.frame. R package version 1.9.2. The Comprehensive R Archive Network, http://CRAN.R-project.org/package=data.table

Elton CS (1958) The ecology of invasions by animals and plants. Methuen and Co Ltd, London

Epanchin-Niell RS, Haight RG, Berec L, Kean JM, Liebhold AM (2012) Optimal surveillance and eradication of invasive species in heterogeneous landscapes. Ecology Letters 15:803-812

Essl F et al. (2011) Socioeconomic legacy yields an invasion debt. Proceedings of the National Academy of Sciences 108:203-207

Finnoff D, Shogren JF, Leung B, Lodge D (2007) Take a risk: Preferring prevention over control of biological invaders. Ecological Economics 62:216-222

Fridley JD et al. (2007) The invasion paradox: Reconciling pattern and process in species invasions. Ecology 88:3-17

Genovesi P (2005) Eradications of invasive alien species in Europe: A review. Biological Invasions 7:127-133

Geoscience Australia (2004) Global Map Australia 1M 2001, 3rd edn. Commonwealth of Australia, Canberra

Gordon DR, Onderdonk DA, Fox AM, Stocker RK (2008) Consistent accuracy of the Australian weed risk assessment system across varied geographies. Diversity and Distributions 14:234-242

Harris S, Brown J, Timmins S (2001) Weed surveillance - how often to search? Science for Conservation. Department of Conservation, Wellington

Hauser CE, McCarthy MA (2009) Streamlining 'search and destroy': Cost-effective surveillance for invasive species management. Ecology Letters 12:683-692

Hayes KR, Barry SC (2008) Are there any consistent predictors of invasion success? Biological Invasions 10:483-506 doi:10.1007/s10530-007-9146-5

Hester SM, Cacho OJ, Panetta FD, Hauser CE (2013) Economic aspects of post-border weed risk management. Diversity and Distributions 19:580-589

Hijmans RJ, van Etten J (2013) raster: Geographic data analysis and modeling. R package version 2.1-37. The Comprehensive R Archive Network, http://CRAN.R-project.org/package=raster

Huang D, Zhang R, Kim KC, Suarez AV (2012) Spatial Pattern and Determinants of the First Detection Locations of Invasive Alien Species in Mainland China. PLoS One 7:e31734 doi:10.1371/journal.pone.0031734

Hulme PE (2009) Trade, transport and trouble: managing invasive species pathways in an era of globalization. Journal of Applied Ecology 46:10-18 doi:10.1111/j.1365-2664.2008.01600.x

Hulme PE (2012) Weed risk assessment: A way forward or a waste of time? Journal of Applied Ecology 49:10-19 doi:10.1111/j.1365-2664.2011.02069.x

Hyndman R, Mesgaran M, Cousens R (2015) Statistical issues with using herbarium data for the estimation of invasion lag-phases. Biological Invasions:1-11

Kaplan H, van Niekerk A, Le Roux JJ, Richardson DM, Wilson JR (2014) Incorporating risk mapping at multiple spatial scales into eradication management plans. Biological invasions 16:691-703

Kelsall JE, Diggle PJ (1995) Non-parametric estimation of spatial variation in relative risk. Statistics in Medicine 14:2335-2342

Kim CS, Lubowski RN, Lewandrowski J, Eiswerth ME (2006) Prevention or control: Optimal government policies for invasive species management. Agricultural and Resource Economics Review 35:29-40

Kolar CS, Lodge DM (2001) Progress in invasion biology: predicting invaders. Trends in Ecology and Evolution 16:199-204

Kulldorff M (1997) A spatial scan statistic. Communications in Statistics — Theory and Methods 26:1481–1496

Lambdon PW et al. (2008) Alien flora of Europe: species diversity, temporal trends, geographical patterns and research needs. Preslia 80:101-149

Lavoie C (2013) Biological collections in an ever changing world: Herbaria as tools for biogeographical and environmental studies. Perspectives in Plant Ecology, Evolution and Systematics 15:68-76

Lavoie C, Saint-Louis A, Guay G, Groeneveld E, Villeneuve P (2012) Naturalization of exotic plant species in north-eastern North America: trends and detection capacity. Diversity and Distributions 18:180–190

Levine JM, D'Antonio CM (1999) Elton revisited: a review of evidence linking diversity and invasibility. Oikos 87:15-26

Lin W, Cheng X, Xu R (2007) Fast economic development accelerates biological invasions in China. PLoS One 11:e1208

Lonsdale WM (1999) Global patterns of plant invasions and the concept of invasibility. Ecology 80:1522-1536

Maroyi A (2012) The casual, naturalised and invasive alien flora of Zimbabwe based on herbarium and literature records. Koedoe 54:30-36 doi:10.4102/koedoe.v54i1.1054

Moerman DE, Estabrook GF (2006) The botanist effect: counties with maximal species richness tend to be home to universities and botanists. Journal of Biogeography 33:1969-1974

Moodley D, Geerts S, Rebelo T, Richardson DM, Wilson JR (2014) Site-specific conditions influence plant naturalization: The case of alien Proteaceae in South Africa. Acta Oecologica 59:62-71

Moore AL, McCarthy MA, Parris KM, Moore JL (2014) The optimal number of surveys when detectability varies. PLoS One 9:e115345 doi:10.1371/journal.pone.0115345

Moore JL, Rout TM, Hauser CE, Moro D, Jones M, Wilcox C, Possingham HP (2010) Protecting islands from pest invasion: optimal allocation of biosecurity resources between quarantine and surveillance. Biological Conservation 143:1068-1078

Myers JH, Simberloff D, Kuris AM, Carey JR (2000) Eradication revisited: Dealing with exotic species. Trends in Ecology and Evolution 15:316-320

Natural Earth (2014) 1:10m Cultural Vectors - Ports. Natural Earth Data. http://www.naturalearthdata.com/downloads/10m-cultural-vectors/.

Newsom LD (1978) Eradication of plant pests - Con. Bulletin of the Entomological Society of America 24:35-40

Panetta FD (2009) Weed eradication - an economic perspective. Invasive Plant Science and Management 2:360-368

Pebesma EJ, Bivand RS (2005) Classes and methods for spatial data in R. R News 5:9-13

Pfeiffer JM (2008) Biological invasions and biocultural diversity: linking ecological and cultural systems. Environmental Conservation 35:281-293

Pheloung P, Williams PA, Halloy SR (1999) A weed risk assessment model for use as a biosecurity tool evaluating plant introductions. Journal of Environmental Management 57:239-251

Phillips ML, Murray BR, Leishman MR, Ingram R (2010) The naturalization to invasion transition: Are there introduction-history correlates of invasiveness in exotic plants of Australia? Austral Ecology 35:695-703

Pimentel D, Zuniga R, Morrison D (2005) Update on the environmental and economic costs associated with alien-invasive species in the United States. Ecological Economics 52:273-288

Pluess T, Cannon R, Jarosik V, Pergl J, Pysek P, Bacher S (2012) When are eradication campaigns successful? A test of common assumptions. Biological Invasions 14:1365-1378

Pyšek P et al. (2013) Hitting the right target: taxonomic challenges for, and of, plant invasions. Aob Plants 5:plt042 doi:10.1093/aobpla/plt042

Pyšek P et al. (2010) Disentangling the role of environmental and human pressures on biological invasions across Europe. Proceedings of the National Academy of Sciences 107:12157-12162

Pyšek P, Richardson DM (2006) The biogeography of naturalization in alien plants. Journal of Biogeography 33:2040-2050

Pyšek P, Richardson DM, Rejmanek M, Webster GL, Williamson M, Kirschner J (2004) Alien plants in checklists and floras: towards better communication between taxonomists and ecologists. Taxon 53:131-143 doi:10.2307/4135498

Pyšek P, Sadlo J, Mandak B, Jarosik V (2003) Czech alien flora and the historical pattern of its formation: what came first to Central Europe? Oecologia 135:122-130 doi:10.1007/s00442-002-1170-7

R Core Team (2013) R: A language and environment for statistical computing vol 3.0. R Foundation for Statistical Computing, Vienna

Rejmanek M, Pitcairn MJ (2002) When is eradication of exotic pest plants a realistic goal? In: Veitch CR, Clout MN (eds) Turning the tide: the eradication of invasive species. IUCN SSC Invasive Species Specialist Group, Gland, pp 249-253

Rejmánek M, Richardson DM (1996) What attributes make some plant species more invasive? Ecology 77:1655-1661

Renner IW et al. (2015) Point process models for presence-only analysis. Methods in Ecology and Evolution 6:366-379 doi:10.1111/2041-210x.12352

Revolution Analytics, Weston S (2014) doParallel: Foreach parallel adaptor for the parallel package. R package version 1.0.8. The Comprehensive R Archive Network, http://CRAN.R-project.org/package=doParallel

Rich TCG (2006) Floristic changes in vascular plants in the British Isles: Geographical and temporal variation in botanical activity 1836-1988. Botanical Journal of the Linnean Society 152:303-330

Rojas-Sandoval J, Acevedo-Rodríguez P (2015) Naturalization and invasion of alien plants in Puerto Rico and the Virgin Islands. Biological Invasions 17:149-163

Sakai AK et al. (2001) The population biology of invasive species. Annual Review of Ecology and Systematics 32:305-332

Schmidt D, Spring D, Mac Nally R, Thomson JR, Brook BW, Cacho O, McKenzie M (2010) Finding needles (or ants) in haystacks: Predicting locations of invasive organisms to inform eradication and containment. Ecological Applications 20:1217-1227

Simberloff D (2003) How much information on population biology is needed to manage introduced species? Conservation Biology 17:83-92

Stajsic V, Vaughan AC (2007) The role of the National Herbarium of Victoria (MEL) in the documentation of new weeds. In: Young K, Raymond K (eds) Weed Society of

Victoria Third Biennial Conference 'Earth Wind Fire Water and Weeds', Bendigo, 2007. The Weed Society of Victoria, Frankston, pp 70-74

Sullivan JJ, Williams PA, Cameron EK, Timmins SM (2004) People and time explain the distribution of naturalized plants in new zealand. Weed Technology 18:1330-1333

Tilman D (1997) Community invasibility, recruitment limitation, and grassland biodiversity. Ecology 78:81-92

Tingley R, Weeks AR, Smart AS, van Rooyen AR, Woolnough AP, McCarthy MA (2014) European newts establish in Australia, marking the arrival of a new amphibian order. Biological Invasions 17:31-37 doi:10.1007/s10530-014-0716-z

van Kleunen M et al. (2015) Global exchange and accumulation of non-native plants. Nature

Vilà M et al. (2011) Ecological impacts of invasive alien plants: a meta-analysis of their effects on species, communities and ecosystems. Ecology Letters 14:702-708 doi:10.1111/j.1461-0248.2011.01628.x

Vila M, Pujadas J (2001) Land-use and socio-economic correlates of plant invasions in European and North African countries. Biological Conservation 100:397-401

Virtue JG, Bennett SJ, Randall RP (2004) Plant introductions in Australia: how can we resolve 'weedy'conflicts of interest. In: Sindel BM, Johnson SB (eds) Proceedings of the 14th Australian Weeds Conference, Wagga Wagga, 2004. Weed Society of New South Wales, Sydney, pp 42-48

Weber J, Panetta FD, Virtue J, Pheloung P (2009) An analysis of assessment outcomes from eight years' operation of the Australian border weed risk assessment system. Journal of Environmental Management 90:798-807

Wickham H (2007) Reshaping data with the reshape package. Journal of Statistical Software 21:1-20

Wickham H (2009) ggplot2: elegant graphics for data analysis. Springer, New York

Wickham H (2011) The split-apply-combine strategy for data analysis. Journal of Statistical Software 40:1-29

Wickham H (2012) scales: Scale functions for graphics. R package version 0.2.3. The Comprehensive R Archive Network, http://CRAN.R-project.org/package=scales

Williamson MH, Fitter A (1996) The characters of successful invaders. Biological Conservation 78:163-170

Wittenberg R, Cock M (2001) Invasive alien species: a toolkit of best prevention and management parctices. GISP/CAB International, Wallingford

Wu SH, Hsieh CF, Chaw SM, Rejmanek M (2004) Plant invasions in Taiwan: Insights from the flora of casual and naturalized alien species. Diversity and Distributions 10:349-362 doi:10.1111/j.1366-9516.2004.00121.x

Yang W, Ma K, Kreft H (2014) Environmental and socio-economic factors shaping the geography of floristic collections in China. Global Ecology and Biogeography 23:1284-1292

(A) **SUPPORING INFORMATION**

Additional Supporting Information may be found in the online version of this article:

Appendix S1 R script for the analysis.

Table 1 Summary information for the factors selected for use in the analysis

Factor	Categories / Units	Source format	Conversion	Final format	Data Source
Environmental					
Average yearly rainfall	Measurement (mm)	0.025 degree raster	Projection	5000 m raster	(BoM 2015)
Average daily mean temperature	Measurement (°C)	0.025 degree raster	Projection	5000 m raster	(BoM 2015)
Average annual days of frost	Count (days)	0.005 degree raster	Projection	5000 m raster	(BoM 2015)
Average daily solar exposure	Measurement (MJ m^{-2})	0.005 degree raster	Projection	5000 m raster	(BoM 2015)
Normalised Difference Vegetation Index	Index	0.250 degree raster	Projection	5000 m raster	(BoM 2015)
Distance to nearest watercourse	Measurement (m)	1:1M Line & Polygon	Minimum Distance	5000 m raster	(Geoscience Australia 2004)
Anthropogenic					
Distance to nearest railway	Measurement (m)	1:1M Line	Minimum Distance	5000 m raster	(Geoscience Australia 2004)
Distance to nearest port	Measurement (m)	Point	Minimum Distance	5000 m raster	(Natural Earth 2014)
Distance to nearest herbaria	Measurement (m)	Point	Minimum Distance	5000 m raster	(CHAH 2014a)
Road density	Measurement (km $pixel^{-1}$)	1:1M Line	Total Distance	5000 m raster	(Geoscience Australia 2004)
Human population density	Measurement (n ha^{-1})	Polygon	Density	5000 m raster	(ABS 2012)
Land Use	ALUM Land Use Class[a]	50 m raster	Mode	5000 m raster	(ABARES 2014)

[a]The ALUM classification was sourced from ABARES (2010).

Fig. 1 Map of Australia showing the first records of occurrence for the 2,699 naturalised alien plant species known in Australia. Colour indicates the year in which the record was collected between 1800 and 2010. Map is geographic

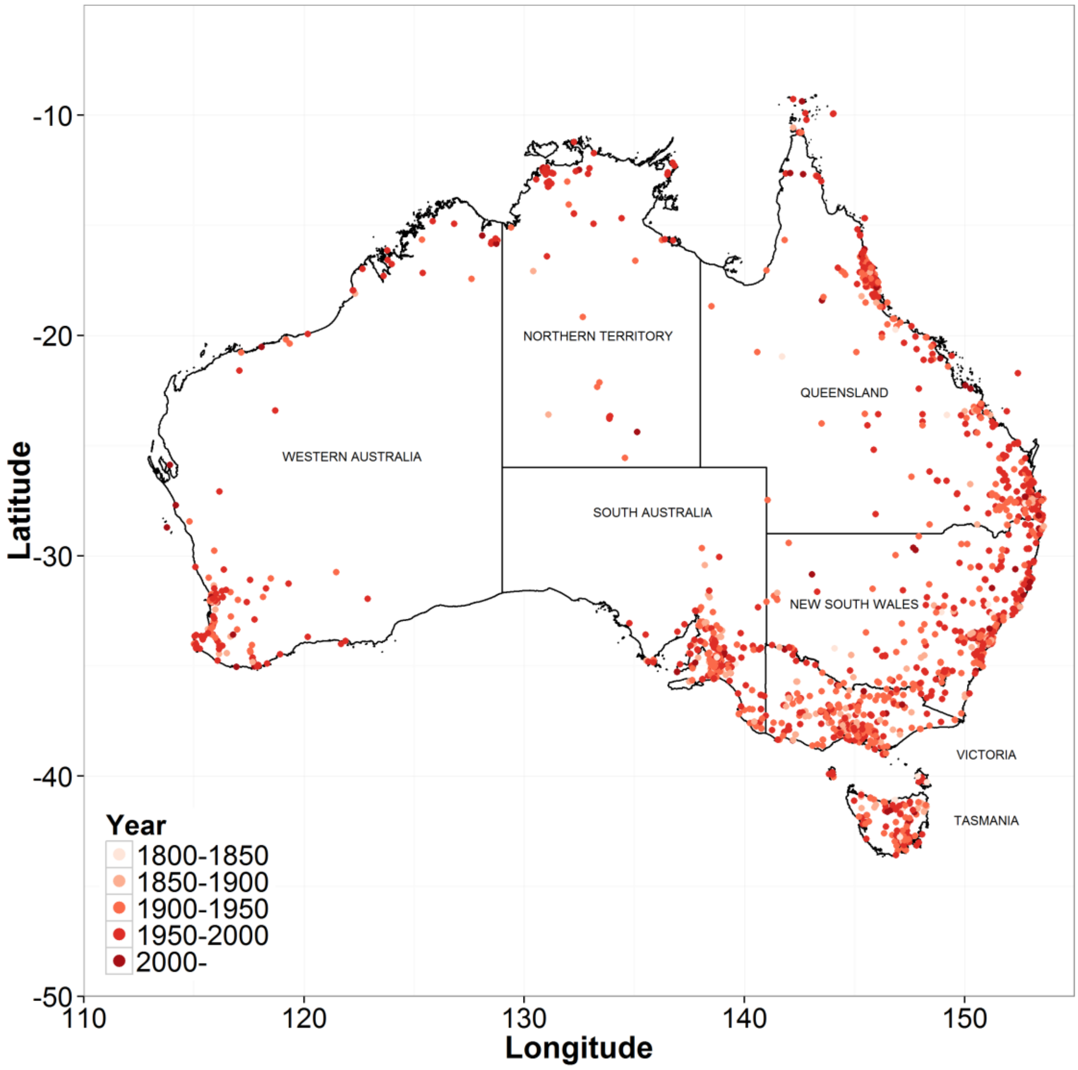

Fig. 2 Detection events (a); kernel smoothed intensity (events per square metre) of detection events (b); collection events (c); kernel smoothed intensity (events per square metre) of collection events (d); likelihood ratio test statistic for locations identified to have higher than expected proportion of detection events than expected given the collection effort (e); and the probability of a detection event occurring (f) in the state of Victoria, Australia. All records collected between 1800 and 2010 are shown in comparison to major roadways. Melbourne (state capital city) is indicated by a cross. Map projection is Australian Albers.

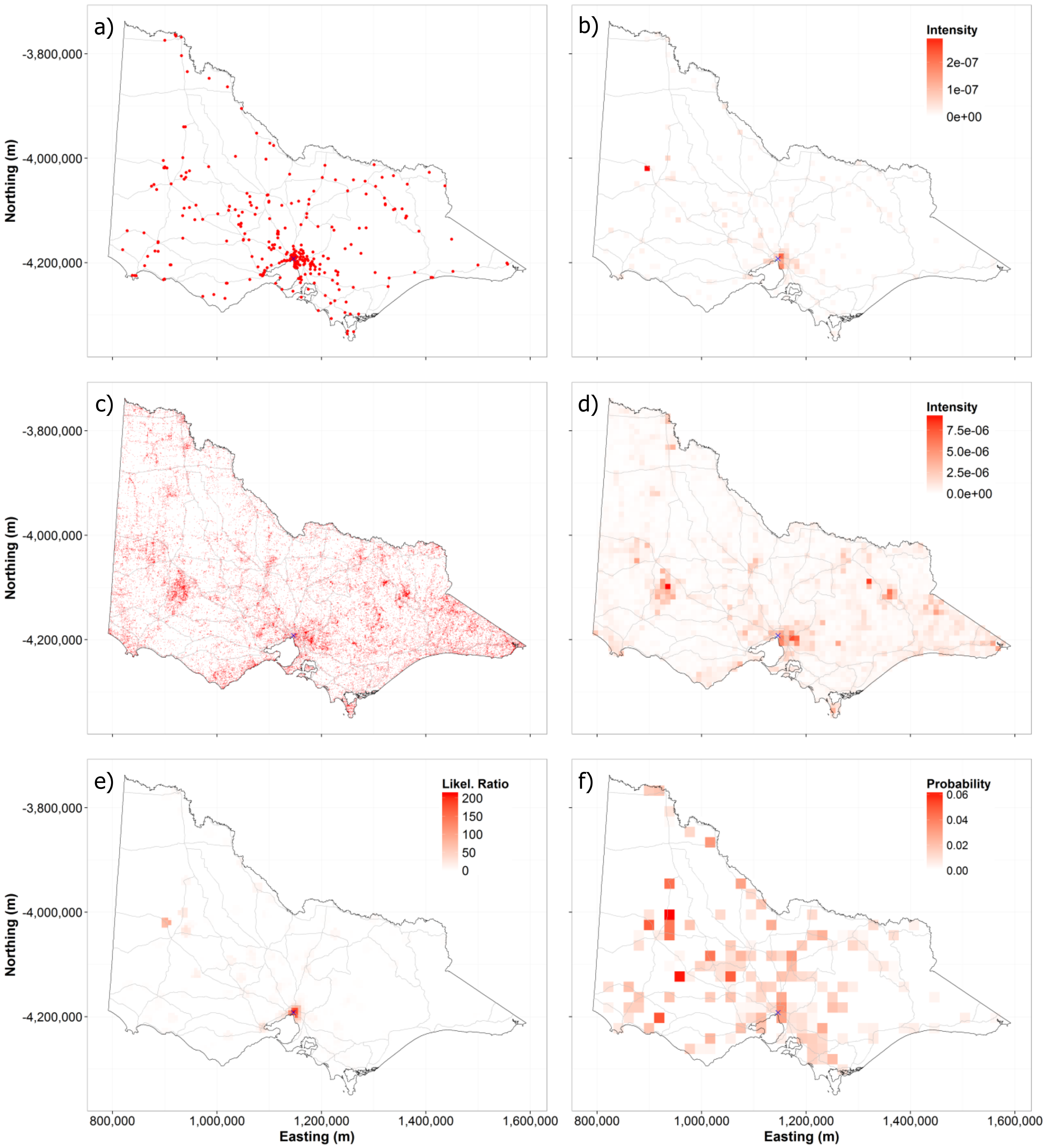

Table 2 Relative importance of model variables in the best performing model. The categorisation of Z-test scores was "*" p <0.05, "**" p<0.01, and "***" p <0.001

Factor	**Estimate**	**S.E.**	**Z-test**	**Z-score**
Intercept	-3.294553e+00	2.816843e-01	***	-11.6959064
Environmental				
Average yearly rainfall (mm)	4.615e-04	1.152e-04	***	4.003
Average daily mean temperature (°C)	-1.099e-01	1.418e-02	***	-7.750
Average annual days of frost (n)	N/A			
Average daily solar exposure (MJ m^{-2})	N/A			
Normalised Difference Vegetation Index	-2.797e+00	3.411e-01	***	-8.198
Distance to nearest watercourse (m)	-2.176e-05	1.413e-05		-1.539
Anthropogenic				
Distance to nearest railway (m)	-4.455e-06	1.374e-06	**	-3.240
Distance to nearest port (m)	-7.767e-07	4.068e-07		-1.909
Distance to nearest herbaria (m)	3.888e-06	4.528e-07	***	8.587
Road density (km $pixel^{-1}$)	2.948e-02	4.332e-03	***	6.805
Human population density (people ha^{-1})	2.445e-01	1.679e-02	***	14.563
Land Use – 2 (Natural area agriculture)	-3.877e-01	1.685e-01	*	-2.301
Land Use – 3 (Dryland agriculture)	2.810e-01	1.023e-01	**	2.745
Land Use – 4 (Irrigated agriculture)	4.699e-01	2.316e-01	*	2.029
Land Use – 5 (Intensive uses)	8.636e-02	1.229e-01		0.702
Land Use – 6 (Water)	-7.173e-01	4.586e-01		-1.564

Fig. 3 Observed location of the first records of occurrence for alien plant species found in the State of South Australia, Australia (a); compared with the predicted probability of a detection event given constant collection effort (b); and the predicted intensity of detection events conditional on historic collection effort (c). Adelaide (state capital city) is indicated by a cross. Pixel size is 5km x 5km. Map projection is Australian Albers

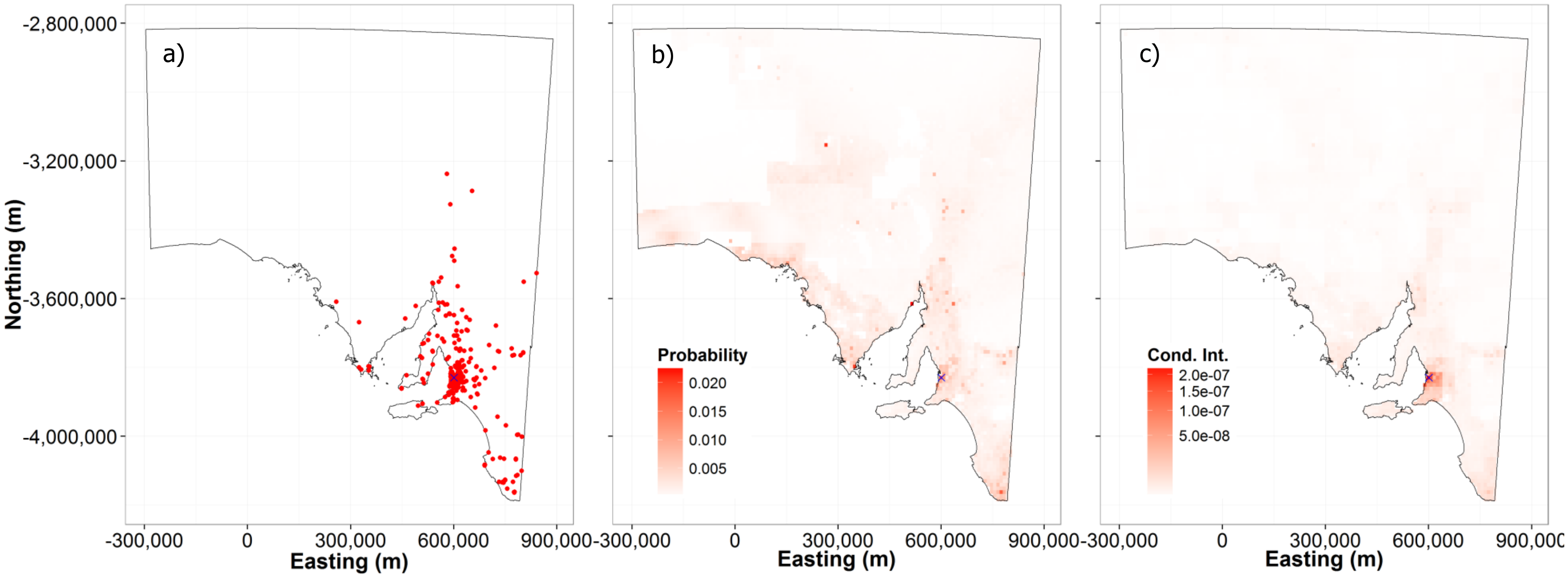

Fig. 4 Observed location of the first records of occurrence for alien plant species found in the State of Queensland, Australia (a); compared with the predicted probability of a detection event given constant collection effort (b); and the predicted intensity of detection events conditional on historic collection effort (c). Brisbane (state capital city) is indicated by a cross. Pixel size is 5km x 5km. Map projection is Australian Albers