

# Models of spatiotemporal variation in rabbit abundance reveal management hotspots for an invasive species

Stuart C Brown<sup>1\*</sup>, Konstans Wells<sup>2</sup>, Emilie Roy-Dufresne<sup>1</sup>, Susan Campbell<sup>3</sup>, Brian Cooke<sup>4</sup>, Tarnya Cox<sup>5</sup>, Damien A. Fordham<sup>1,6</sup>

<sup>1</sup> The Environment Institute and School of Biological Sciences, University of Adelaide, Adelaide, South Australia 5005, Australia

<sup>2</sup> Department of Biosciences, Swansea University, Swansea SA2 8PP, UK

<sup>3</sup> Biosecurity and Regulation, Primary Industries and Regional Development, Albany, Western Australia 6330, Australia

<sup>4</sup> Institute for Applied Ecology, University of Canberra, Canberra, ACT 2601, Australia

<sup>5</sup> Vertebrate Pest Research Unit, NSW Department of Primary Industries, Orange, New South Wales 2800, Australia

<sup>6</sup>Center for Macroecology, Evolution, and Climate, Natural History Museum of Denmark, University of Copenhagen, Copenhagen, Denmark

\* Corresponding author. Email: [s.brown@adelaide.edu.au](mailto:s.brown@adelaide.edu.au)

**Acknowledgements:** This research was funded by an Australian Research Council (ARC) Linkage Project (LP12020024) and Future Fellowship (FT140101192). We would like to thank Greg Mutze for discussions and advice regarding preliminary analysis and outputs.

**Author contributions:** D.A.F led the project. S.C.B and K.W did the analysis and discussed the results with all authors. S.C.B drafted the manuscript and all authors provided comments.

**Title:** Models of spatiotemporal variation in rabbit abundance reveal management hotspots for an invasive species

**Running title:** Spatiotemporal hotspots of rabbit abundance

**Abstract:**

Aim: The European rabbit (*Oryctolagus cuniculus*) is a notorious economic and environmental pest species in its invasive range. To better understand the population and range dynamics of this species, long-term abundance data has been collected across a broad range of climatic and environmental condition in Australia. We analysed this time series data to determine whether inter-annual variation in climatic conditions can be used to map historic, contemporary, and potential future fluctuations in rabbit abundance from regional to continental scales.

Location: Australia

Method: We compiled spotlight monitoring data at 116 unique sites, collected over 41 years from locations in different climate and vegetation zones. We constructed a hierarchical Bayesian regression model of relative abundance that corrected for observation error and seasonal biases. The corrected abundances were regressed against environmental and disease variables in order to project high spatiotemporal resolution, continent-wide rabbit abundances.

Results: We show that rabbit abundance in Australia is highly variable in space and time, being driven primarily by inter-annual variation in temperature and precipitation in concert with the prevalence of a non-pathogenic virus. Moreover, we show that inter-annual variation in local spatial abundances can be mapped effectively at a continental scale using highly resolved spatiotemporal predictors, allowing “hotspots” of persistently high rabbit abundance to be identified. Importantly, cross-validated model performance was fair to excellent within and across distinct climate zones.

Primary conclusion: Long-term monitoring data for invasive species can be used to map fine-scale spatiotemporal fluctuations in abundance patterns when accurately accounting for inherent sampling biases. Our analysis provides ecologists and pest managers with a clearer understanding of the determinants of rabbit abundance in Australia, offering an important new tool for predicting spatial abundance patterns of invasive species at the near-term temporal scales that are directly relevant to resource management.

**Keywords:** historic climate change, invasion hotspot, invasive species management, long-term monitoring, N-mixture model, *Oryctolagus*, random forests

## **Introduction:**

Long term datasets are vitally important for the successful management of native and invasive species (Yoccoz, Nichols, & Boulinier, 2001). In a macro-ecological context, such datasets should ideally report the abundance of the species through time at multiple sites, capturing a broad range of niche requirements (Magurran et al., 2010). These types of spatiotemporal explicit datasets, whilst being rare (Whittaker et al., 2005), can help to identify climatic, environmental and biotic processes that determine species range limits and drive the structure and dynamics of geographic ranges (Caughley, Grice, Barker, & Brown, 1988; Gaston, 2003). Furthermore they can be used to predict spatiotemporal variation in population fluctuations, which is crucial for effective conservation (Channell & Lomolino, 2000) and pest and invasive species management (Strayer, Eviner, Jeschke, & Pace, 2006).

Differences in duration, methodology, frequency of sampling, spatial grain and spatial extent within and between studies, makes systematic comparisons of inferences of long-term population trends, from multiple studies problematic (Magurran et al., 2010; Pagel et al., 2014). Notwithstanding these difficulties, the integration of long-term spatially explicit datasets from disparate sources, with varying spatial and temporal scales, is becoming more common in macroecological studies as a means of forecasting patterns of spatial abundance in time and space (Mellin, Russell, Connell, Brook, & Fordham, 2012), and to better understand ecological impacts of invasive species (Strayer et al., 2006). The development of statistical approaches that allow sampling and temporal biases (that result in false variation; Martin et al., 2005), among other sources of uncertainty in abundance time series data (e.g., imperfect detection; Joseph, Elkin, Martin, & Possingham, 2009), to be identified and explicitly modelled, is facilitating better estimates of spatiotemporal variation in local-to-regional species abundance patterns (Joseph et al., 2009; Magurran, 2007; Schurr et al., 2012).

Although abundance data are still typically analysed using the average or maximum count per unit transect (Buckland, Rexstad, Marques, & Oedekoven, 2015), N-mixture models (models with a mixed binomial-Poisson likelihood), which estimate relative abundance based on replicated counts (Royle, 2004), are being used with increasing frequency to explicitly account for detection probability in spatially explicit estimates of abundance (Aubry et al., 2012; Guillera-Arroita, 2017), including invasive species (Wells, O'Hara, et al., 2016). Such models essentially produce a bias-corrected abundance estimate, that can then be used to examine trends in population fluctuations and abundances. To do this, corrected abundance data are combined with environmental and climatic information,

and where available, data on biotic interactions (i.e. covariate effects; Royle, 2004), to determine potential drivers of abundances. These statistical relationships can then be used to make projections of abundance, providing opportunities to assess population trends and geographic range limits across space and time (Joseph et al., 2009; Schurr et al., 2012).

The abundances and subsequent ecological impacts of invasive species are typically highly variable through time in their invasive range (Sofaer, Jarnevich, & Pearse, 2018), making them highly suited for N-mixture modelling type approaches (Guillera-Arroita, 2017; Joseph et al., 2009). For example, the occupancy and abundance rates of invasive species can change over short time scales as a result of shifting evolutionary pressures (Cox, 2013; Phillips, Brown, Webb, & Shine, 2006), modifications in community composition (Mutze, Cooke, & Jennings, 2016), and changes in species interactions (Lurgi, Ritchie, & Fordham, 2018). This temporal variability means that invasive species are typically not in equilibrium with their environment (Elith, Kearney, & Phillips, 2010). Therefore, the collection of long-term count data at spatiotemporal resolutions that allow for acute and chronic effects of abiotic and biotic conditions on abundance to be teased apart using robust statistical approaches is vital for generating estimates of population variation for invasive species. These estimates will lead to better ecological and economic outcomes if the mechanisms that affect population fluctuations of invasive species can be identified and managed accordingly (Strayer et al., 2006).

Abundance, not probability of occurrence, is the currency of invasive species management. This is because ecological mechanisms of range limits and shifts can be better understood using abundance data (Fordham et al., 2012) and spatiotemporal estimates of abundance and subsequent variability are frequently used to maximise the efficacy of on-ground management efforts to control invasive species (Simberloff, 2003). For example, efforts for managing rabbits in their invasive range are often reduced when population size is perceived to be relatively low, allowing resources to be prioritised elsewhere (Wells, Cassey, et al., 2016). Moreover, time series abundance estimates are important for the development of spatially explicit and dynamic population models for invasive species management (Mellin et al., 2016), enabling management intervention strategies to be directly tested (Fordham, Akcakaya, Araujo, Keith, & Brook, 2013). This is because they enable key demographic rates to be estimated, including finite rates of population growth (Hone, 1999), and for models to be calibrated (Cabral & Schurr, 2010). Therefore, modelling spatiotemporal abundance patterns of invasive species provides resource managers with some distinct advantages over modelling occurrence.

The European rabbit (*Oryctolagus cuniculus*) is an invasive pest species in several regions around the world (Cooke, 2012); but is a threatened species in its native range on the Iberian peninsula, where it is a multifunctional keystone species that engineers habitat and provides prey to consumers (Delibes-Mateos, Delibes, Ferreras, & Villafuerte, 2008; Fordham, Akçakaya, et al., 2013). In Australia, *O. cuniculus* occurs across > 70 % of the mainland (Stodart & Parer, 1988) and is well adapted to the broad range of climatic conditions of the central and southern parts of the continent (Cooke, Brennan, & Elsworth, 2018), making it one of Australia's most abundant invasive vertebrate species (Cooke, 2012). It negatively effects Australia's unique biodiversity directly through outcompeting native species for food and shelter (Cooke, 2012), altering vegetation composition (Mutze et al., 2016), and, indirectly, by maintaining invasive predator populations (Lurgi et al., 2018). Furthermore, rabbits adversely affect the agro-economy with an attributed economic loss of approximately A\$200m annually (Cooke, Chudleigh, Simpson, & Saunders, 2013).

Recently, more than 50 years of historical and contemporary rabbit abundance survey data, from across the invasive range of rabbits in Australia, was combined with high resolution weather, climate and environmental information to create the Australian National Rabbit Database (Roy-Dufresne, E., personal communication). This dataset provides new and important opportunities to better manage rabbits in Australia through a stronger understanding of their population fluctuations and range dynamics. To date, there has been no Australia-wide assessment of spatiotemporal prediction of rabbit abundances and fluctuations. Such an assessment is needed to identify where, and under which environmental conditions and with what temporal lags rabbits attain high abundance, due to population recruitment and resistance to environmental perturbations.

Here we combine spotlight count data from 116 sites collected over 41 years (1972 – 2012 inclusive) for mainland Australia for *O. cuniculus*. This period occurs approximately 20 years after the introduction of myxomatosis to mainland Australia, while Rabbit Haemorrhagic Disease Virus (RHDV) was introduced in late 1995 (Pech & Hood, 1998), midway through the focal period. We apply a multi-level hierarchical N-mixture model to this data to “correct” for important biases (namely detection error and bias due to seasonal reproduction and abundance fluctuations) in annual estimates of abundance; and then use machine learning algorithms to determine the relative influence of climate, environment and disease on spatiotemporal variation in rabbit abundance in Australia. Our two-step approach provides a computationally efficient method for generating range-wide, high-resolution spatially (~5 km) and temporally (annual) explicit projections of abundance for a well-

established invasive species, which can then be used to reveal hotspots of abundance and to elucidate the drivers of the structure and dynamics of their geographic ranges.

## **Methods:**

The key steps used to analyse time series count data for rabbits and predict spatiotemporal ‘hotspots’ of abundance are shown schematically and described in Fig. S1.

### *Spotlight data:*

The spotlight data collated in the Australian National Rabbit Database (Roy-Dufresne, E., personal communication) covers the broad range of climatic and environmental conditions that rabbits inhabit in Australia (Fig. 1). The data was collated from several different sources, mostly regional researchers affiliated with state-level governmental institutions. We only consider data records with complete information regarding: (1) geolocation; (2) survey dates; (3) transect length (limited to plausible values  $\geq 1$  and  $\leq 100\text{km}$ ); and (4) rabbit encounters (0 counts were included). This resulted in 116 unique “sites” (after aggregating all count data into  $5 \text{ km}^2$  grid cell units) for which we had enough information to model rabbit abundances in space and time.

### *Climate and environmental data*

The aggregated site data was matched to a data frame containing information on total daily precipitation (mm), and minimum and maximum temperature ( $^{\circ}\text{C}$ ) data from the TERN e-MAST facility (<http://portal.tern.org.au>). These climatic variables have been shown to influence the burrow emergence behaviour of rabbits, and therefore affect detection probability of rabbits during spotlight surveys (Ballinger & Morgan, 2002). Furthermore, grid-cell information on the location and length of the transect, annual climatic conditions (e.g. previous summer total precipitation, previous winter average minimum temperature), environmental conditions (e.g. proportion of agricultural land within the  $5\text{km}^2$  grid cell, proportion of clay soils), and disease proxies (e.g. prevalence of non-pathogenic RCV-A1 virus; Liu et al., 2014) was also joined to the aggregated data.

Full details of the climatic and environmental variables used in our analysis are shown in Table S1.

### *Climate zone*

To account for regional differences in seasonal rabbit population growth and recruitment (Hone, 1999; Mutze et al., 2002; Wells, O'Hara, et al., 2016), the spotlight records were grouped into climate zones, where long term climatic conditions were similar. The climate zones were identified by first doing a principal component analysis on 12 climatic variables

(e.g. 30 year mean maximum temperature, precipitation seasonality), and then an iterative test to identify the number of “climate zones” required to adequately describe the diversity of climatic conditions. Using a k-means clustering algorithm (R Core Team, 2017), the iterative test involved minimising the within-cluster sum of squared errors (SSE) for a range of potential cluster solutions (2 – 12 clusters) and identifying the point of inflection where the SSE levelled off. Since rabbit reproduction and regional population fluctuations are often highly seasonal (Mutze et al., 2002; Wells, O’Hara, et al., 2016), the spotlight abundance data was grouped by climatic zones where long-term average climate and resource availability is likely to be similar (Fig. 1). In doing so, we assume that all rabbit populations from a climate zone follow similar seasonal population fluctuations through synchronous timing of reproduction. These variables (mean  $\pm$  S.D) for each climate zone are shown in Table S2. The climate zones are available in Appendix I.

#### *Correcting abundance count data:*

We established a generalised multi-level hierarchical mixture model in a Bayesian framework to derive an index of standardised relative yearly rabbit abundance free of well-established and important sampling biases that is comparable across surveys. Specifically, we used an  $N$ -mixture model (Royle, 2004) to model spotlight counts  $c(t,k)$  for each of the repeated surveys  $k$  during a given year in for cell  $r$  at time step  $t$  as:

$$c(t,k) \sim \text{Bin}[N_r(t), p(t,k)] \text{ and } N_r(t) \sim \text{Pois}[\Lambda(r,t)] \#(1)$$

where, we assumed each count to be a random draw based on the underlying true but unknown abundance  $N_r(t)$ , governed by the Poisson density  $\Lambda(r,t)$ .

We modelled the detection probability  $p(t,k)$  with a *logit*-link function as:

$$\text{logit}[p(t,k)] = \mu(r) + \eta(c) + \nu_1(ds) + \nu_2(t_{max}) + \nu_3(t_{min}) + \nu_4(pr_{day}) \#(2)$$

where,  $\mu(s)$  accounts for variation in detection probability between cells ( $r$ ),  $\eta(c)$  accounts for variation between climate zones ( $c$ ),  $\nu(ds)$ ,  $\nu(t_{max})$ ,  $\nu(t_{min})$ , and  $\nu(pr_{day})$  allow for variation between different data sources (corresponding to different survey teams), and daily weather variables. Note that  $\Lambda(r,t)$  represents the relative local abundance at the time of observation and is not corrected for transect length nor seasonal bias in counts.

The spotlight surveys were conducted in different seasons and along transects of different lengths, therefore, the estimate of detection probability needs to account for spatiotemporal variability in recruitment (Wells, O’Hara, et al., 2016) and survey efforts (Ballinger & Morgan, 2002; Barrio, Acevedo, & Tortosa, 2009). Doing so will reduce potentially important sources of bias in estimates (Link, Schofield, Barker, & Sauer, 2018).

We used random regression models, following Wilson, Kruuk, and Coltman (2005), to account for potentially important variation due to seasonal bias and different transect lengths. For this, we modelled  $\Lambda(r,t)$  on a log-scale (natural logarithm):

$$\ln[\Lambda(r,t)] = \omega(r,y,m) + \ln[transL(i)] \#(3)$$

where,  $\omega(r, y, m)$  is the intercept of the random regression model, which can vary across cells ( $r$ ), years ( $y$ ), and months ( $m$ ); and  $transL(i)$  is the corresponding length of the transect in km for the respective sample.

We modelled  $\omega(r, y, m)$  further based on a two-level mean/hyperprior models with Gaussian error of the general form  $N[H, \sigma^2]$  with mean  $H$  and variance  $\sigma^2$  in order to derive an estimate of the most likely yearly abundance independent of the seasonal bias in counts:

$$\omega(r, y, m) \sim N[H_{RY.\omega}(r, y), \sigma_{R.\omega}(r)^2] \#(4)$$

and

$$H_{RY.\omega}(r, y) \sim N[H_{R.\omega}(r), \sigma_{R.\omega}(r)^2] \#(5)$$

where,  $H_{RY.\omega}(r, y)$  denotes the population-level annual ‘average’ density distribution for each site and year, while the respective variance  $\sigma_{R.\omega}(r)^2$  is a random estimate of the inter-seasonal variance in population fluctuations for each cell ( $r$ ) according to the variation in monthly counts from different sites.  $H_{R.\omega}(r)$  denotes the site-level ‘average’ relative abundance and  $\sigma_{R.\omega}(r)^2$  interannual variance in population fluctuations.

Our model was fitted in a Bayesian framework using Markov Chain Monte Carlo (MCMC) sampling in the software JAGS 4.3.0, operated via the R software (v 3.5.1; R Core Team, 2017) with the ‘rjags’ package (Plummer, 2016). Priors were specified as  $\sigma \sim \text{Gamma}(0.01, 0.01)$  and  $H_{R.\omega}(r) \sim N(0, 10)$ .

We applied posterior predictive model diagnostics to assess whether the model assumptions were good approximations of the data generating process. Bayesian  $P$ -values  $\sim 0.5$  indicate a good model fit whereas values close to 0 or 1 indicate an increasing discrepancy between model predictions and observation data (Gelman, Meng, & Stern, 1996). We calculated 5000 values as posterior distributions for all parameters of interest after mixing and convergence of two parallel MCMC chains was confirmed visually and with the Gelman-Rubin diagnostic (most values  $< 1.2$ ) after a burn-in of 50,000 MCMC samples.

#### *Modelling rabbit abundances in space and time:*

Modelled corrected abundances ( $H_{RY.\omega}(r, y)$ ; Fig. S2) – extracted from the Bayesian model as the mode of the posterior density distributions – were aligned with the key predictor variables in Table S1. We opted to use a readily available and simple machine learning

approach, random forests (RF; Breiman, 2001), to generate predictions of rabbit density. This decision was made because compared to most Bayesian methods, RFs enable a broad range of model and variable selection and cross-validation procedures that are not commonly used in an MCMC context due to practical computing time issues (Kéry & Schaub, 2011).

Essentially, we use the hierarchical model for correcting abundances for various sources of bias, while we use RFs for exploring and accounting for any non-linear patterns between corrected abundances and climate and environmental correlates in spatial predictions.

For simplicity, we categorised the annual model corrected abundance estimates into ‘low density’ (corresponding to  $\leq 45$  rabbits/km $^2$ , assuming an effective transect width of 0.1 km) and ‘high density’ (corresponding to  $> 45$  rabbits/km $^2$ ). These values were chosen as there was a natural break in the calculated density values, and it has been demonstrated that at densities  $< 50$  rabbits/km $^2$ , damage to perennial native vegetation is minimised across both temperate and arid environments (Mutze et al., 2016). This binary variable allowed us to use a two-step regression model (Mellin et al., 2012) for predicting rabbit abundance through time and space.

A binary classification RF model was first created using the aforementioned density classes. Due to the unbalanced nature of the classes *a posteriori* tuning of the probability threshold for classification (default = 0.50) was performed to maximise model accuracy. Using this approach, once the classification RF model had been trained (using the default threshold), a series of thresholds between 0 and 1 were tested to minimise the distance to a perfect model (i.e. where sensitivity and specificity = 1). A separate regression RF model was then built which incorporated the RF predicted density classes as an additional correlate for predicting rabbit abundance. The two-step procedure necessitated the use of RF predicted rabbit density for cell  $r$  at time  $t_{-1}$ .

To train and assess the performance of our RF models, we used 1,000 repeats of a Monte-Carlo cross-validation procedure, whereby the data was randomly split (without replacement) into 80/20 training/test splits. Models consisting of 2001 trees were built on the 80% splits and tested on the 20% holdout test samples. Correlations between the environmental and climatic correlates (Table S1) were tested, and with a few exceptions, were not strongly positive or negatively correlated (Fig. S3). As random forests are able to deal with correlated variables (Breiman, 2001), we did not perform any feature selection or reduction on the correlates included in the model. Spatial autocorrelation in the residuals of predicted RF abundance was conducted using Moran’s I statistic.

Spatially and temporally explicit estimates of rabbit density and abundance (i.e. maps), were then predicted for Australia between 1972 and 2012. Hotspots of high rabbit abundance were identified for each year that abundance predictions were made by approximating a cumulative relative frequency distribution (CRFD) curve as a function of relative rabbit abundance plotted against the relative frequency distribution of abundance for a given year (Bartolino, Maiorano, & Colloca, 2011a). The count data for the CRFD curve was pooled across all sites within each year. Unlike traditional hotspot identification methods which require a user defined threshold (Bartolino et al., 2011a; Bartolino, Maiorano, & Colloca, 2011b; Cayuela, Gálvez-Bravo, Carrascal, & de Albuquerque, 2011), the CRFD method identifies the hotspot threshold as a function of the CRFD curve. The method involves calculating the slope of the tangents to the CRFD curve, with the hotspot threshold being identified as the value where the slope of the tangent is 45° (Bartolino et al., 2011a). Using the thresholds for each year, the predicted abundance maps were then converted to a series of binary hotspot maps. The proportion of time spent in “hotspot” conditions across the study period was then determined for each cell by counting the number of times a cell was classified as a hotspot, divided by the number of years.

#### *Climate envelope and confidence mask*

To constrain predictions of corrected density and abundance to climate conditions used to train the model and, in doing so, avoid model extrapolation, we identified all cells for Australia that fell within the extremes (i.e. highest and lowest) of the climatic variables that were used to define the climate zones. In doing so we constrain our predictions of density and abundance to cells that fall within the bounds of our long-term climatic variables and limit extrapolation to “unseen” climatic conditions (Liu et al., 2014). All maps in Appendix I are clipped to this climate envelope.

To identify regions of “low confidence”, we did a PCA using the top 10 environmental and climatic variables that were considered important for both the classification and regression RF models, and using the first 3-components as pseudo-3d-coordinates, we identified individual grid cells in 3-dimensional space that were outside a 3-dimensional convex-hull of the sampling sites within each of the climate zones. That is, 3-dimensional polygons were constructed, with vertices defined by the pseudo-3d-coordinates of all cells for which we had samples within each climate zone. Any additional cells (i.e. grid cells where no monitoring had occurred) that exist inside the 3d boundary of this polygon can be considered inside the “environmental space” of the study sites within each climate zone.

Cells outside this 3d polygon, have climatic or environmental conditions not seen at the sites used within the climate zone. This process allowed us to identify grid-cells where our predictions of corrected density and abundance were outside the environmental and climatic space of all monitoring sites in the study, essentially creating a mask of low confidence where predictions were likely to have been extrapolated in environmental and/or climatic space, rather than interpolated. The confidence mask is supplied in Appendix I.

## **Results:**

### *Drivers of detection*

Coefficient estimates for the drivers of detection varied in magnitude and sign across climate zones (Table 1). Within 5 of the 9 climate zones (1, 2, 4, 6, and 7) an increase in maximum daily temperature decreased the probability of detection (mean = -0.23, S.D. = 0.11), whilst for the remaining 4 zones (3, 5, 8, and 9) it resulted in an increase in detection probability (mean = 0.28, S.D. = 0.18). Increases in minimum daily temperature (except for climate zone 4) and daily total precipitation, resulted in slight increases in detection probability (minimum daily temperature: mean = 0.04, S.D. = 0.03; precipitation: mean = 0.05, S.D. = 0.02).

Gelman-Rubin diagnostics for hyperpriors indicated good convergence of the detectability model: (i) maximum daily temperature = 1.08 (upper CI = 1.32); (ii) minimum daily temperature = 1.17 (upper CI = 1.59); (iii) daily precipitation = 1.09 (upper CI = 1.31).

### *Determinants of spatiotemporal patterns of density and abundance*

The five most important predictor variables for the classification RF model of rabbit density (high/low) were: temperature seasonality, proportion of irrigated agriculture within the grid-cell, two year lagged average temperature, distance to permanent water, and the prevalence of the benign RCV-A1 calicivirus (Fig. S4). Evaluation of these variables with Accumulated Local Effects (ALE; Apley, 2016) plots (Fig. S5) suggests that at sites with lower temperature seasonality and lower mean annual temperatures in the years ( $n = 2$ ) prior to the survey, there was an increased probability of high rabbit densities. Furthermore, at the few sites that were sampled on irrigated land, sites with as little as ~7.5% irrigated agriculture within a 5km grid-cell were 60% more likely to have high rabbit densities than a conditional averaged site (Fig. S5). The probability of a site having high density decreased with distance from permanent water until  $\geq 20\text{km}$ . Although some sites that had high densities were situated very far away from permanent water sources ( $> 150\text{km}$ ; Fig. S5). As benign RCV-A1 prevalence increased above ~20% the probability of a site having high abundance was ~25% greater than the average site (Fig. S5). Taken together the results from the

classification RF model suggest that at sites with stable, cool temperatures, with abundant pastures (owing to irrigation, and proximity to permanent water sources), and prevalent RCV-A1, rabbit densities will be higher. Cross-validated performance of the classification model was excellent across climate zones and in most cases fair to excellent within climate zones (Table 2). The mode of the predicted rabbit densities and the mean prediction probability confidence (1972 – 2012) is shown in Fig. 2. The complete time series of density maps is shown in Appendix I.

The five most important variables influencing model predictions of rabbit abundance (excluding rabbit density class from the previous year; Fig. S4) were: precipitation the year prior, average winter minimum temperature the year prior, two year lagged total precipitation, precipitation seasonality, and the prevalence of benign RCV-A1. These variables exhibited similar importance within the regression RF model (Fig. S4). ALE plots of the regression model (Fig. 3) suggest that at sites where one-year lagged precipitation was  $\gtrsim 400\text{mm}$  the year prior, and one-year lagged minimum winter temperatures  $\gtrsim 4^\circ$ , there was an increase in rabbit abundance. As for the classification model, examination of two year lagged average temperature suggests that at sites with lower mean annual temperatures ( $\lesssim 16^\circ \text{C}$ ) rabbit abundances are higher. Like with the classification RF, this suggests that in cool, wet areas rabbit abundances will be higher. However, for the regression RF the effects of precipitation are likely acting as a proxy for pasture availability. In contrast to the classification model, the plot of the association between abundance and the prevalence of benign RCV-A1 suggests that as virus prevalence increases, abundances will decrease (Fig. 3). Regression model performance was excellent across climate zones and in most cases fair to excellent within climate zones (Table 1). With some exceptions, the 95% prediction intervals of the regression RF model were able to capture the variability in the corrected abundance counts (Fig. 4). Moran's I suggested no residual spatial autocorrelation ( $I = 0.001$ ,  $z = 0.085$ ,  $p = 0.93$ ). The mean of the predicted rabbit abundances and the mean standard deviation of the abundances through time (1973 – 2012) is shown in Fig. 2. All abundance maps are available in Appendix I.

#### *Variable interactions*

Examination of second order interactions for the regression model (Fig. 3b) suggest that for sites with  $> 20\%$  prevalence of benign RCV-A1 and with mean annual temperatures below  $\sim 16^\circ \text{C}$ , the combined effects are positive on abundance. Conversely, second order interactions between precipitation and RCV-A1 (Fig. 3c) suggest that at sites with  $\lesssim 20\%$

prevalence, and annual precipitation is  $> 600\text{mm/year}$ , there is a net positive effect on rabbit abundance suggesting increased precipitation, and therefore an increase in potential pasture availability, could lead to higher abundances at sites where there is low immunity to lethal RHDV. Furthermore, after the main effects of the variables (namely precipitation and RCV-A1 prevalence) have been accounted for, there is a negligible effect of RCV-A1 prevalence on rabbit abundance in areas of high rainfall ( $> 600\text{mm/year}$ ) as illustrated by the relative “flatness” and values of the interaction surface (Fig. 3c).

#### *Hotspots of abundance*

Predicted maps of rabbit density and abundance show high levels of interannual variation along the east coast of Australia, with density classes and abundances fluctuating between high and low levels inland from the coast for a distance of  $\sim 300\text{km}$  (Fig. 2). Hindcasts of abundance for the southern coastline of mainland Australia showed lower inter-annual variability compared to the eastern coastline. Inland areas tended to be relatively stable, however, these areas were generally inside the low-confidence mask.

Analysis of temporal hotspot conditions suggest small inter-annual variability in the threshold used to identify hotspots of rabbit abundance across time (mean = 12, S.D. = 2.7; Fig. 3). A Mann-Kendall trend test suggested a very small but non-significant increase in the hotspot threshold across all years ( $\tau = 0.03$ ,  $p = 0.89$ ) suggesting rabbit numbers are decreasing as the threshold used to identify a hotspot has increased. Broadly, hotspots of rabbit abundance followed similar patterns to maps of predicted abundance (Fig. 5). All hotspot maps are provided in Appendix I.

#### **Discussion:**

We show that combining multi-level hierarchical Bayesian models with simple machine learning methods is an effective, and computationally tractable, method for modelling the structure and dynamics of the geographical distribution of a wide-ranging invasive vertebrate species. By allowing us to project, effectively, spatiotemporal variation in the abundance of rabbit populations across their invasive range in Australia, our approach provides an important new tool for identifying hotspots of persistently high rabbit abundance at timescales (annual to decadal) that are directly relevant to management (Cooke, 2012; Hobday, Hartog, Timmiss, & Fielding, 2010). The identification of potential past and future hotspots will allow pest managers to better plan for and conduct control actions for rabbits in areas and at times of greatest need (Simberloff, 2003).

A robust understanding of spatiotemporal variability in abundance patterns across an invasive species range is invaluable for studying the processes of invasion and for pest management (Strayer et al., 2006). However, this requires spatial abundance estimates that account for inherent biases in long-term count data (Joseph et al., 2009; Link et al., 2018). Our use of a hierarchical Bayesian N-mixture model, coupled with a hyperprior-based model, allowed us to explicitly account for important variation in detection probability in long-term estimates of abundance from sites covering a large spatiotemporal extent. Although similar approaches have been applied to rabbit spotlight count data before (Aubry et al., 2012; Wells, O'Hara, et al., 2016), they have not then been extended to account for seasonal bias in counts from multiple surveys at multiple sites.

Our Bayesian hierarchical N-mixture model, employed at multiple sites across the invasive range of rabbits in Australia, generated a standardised relative yearly rabbit abundance across space and time, corrected for ‘false variation’ as a result of imperfect detection and seasonal fluctuations in abundance (due to seasonal birth pulses) in some areas. Using these corrected abundances in a machine learning framework (random forest models; RF) provided a highly computationally efficient method for identifying the spatiotemporal drivers of heterogeneity in rabbit abundance across their invasive range. Furthermore, the use of RF permitted the analysis of non-linear relationships between, and within, predictor variables, that may have otherwise been missed (Apley, 2016; Breiman, 2001).

The comprehensive nature of the statistical modelling undertaken here, covering 41 years and approximately 3.8 million square kilometres, which is unique for a terrestrial invasive vertebrate, allowed us to directly model potential drivers of rabbit abundance in space and time. Our results support previous research, showing that rabbit populations are influenced by pasture availability and the water content of pastures (Cooke, 2012), temperature (attaining higher densities in cooler temperatures; Gilbert et al., 1987) and to an extent, the effects of RCV-A1 (Cooke, Duncan, et al., 2018). The classification RF showed that RCV-A1 prevalence was positively correlated with high rabbit densities (high/low), because RCV-A1 provides partial protection from lethal RHDV (Strive et al., 2013). Since the prevalence of RCV-A1 is strongly correlated with rainfall (Liu et al., 2014), rabbits in cooler, wetter areas are likely to be found in higher abundances due to food availability, and these same populations are likely to be more resistant to infection with lethal RHDV. Although, RCV-A1 also influenced estimates of abundance from the regression RF, prevalence of RCV-A1 was negatively associated with rabbit abundance – a result likely due to complex interactions between rabbit populations, climatic conditions, disease vectors, and

epizootics (Cooke, 1983; Liu et al., 2014). This result requires further investigation, using new approaches that can directly tease apart these interactions, which are often mutually reinforcing (Wells et al., 2018). The presence of European rabbit fleas was not important in determining either density or abundance, but has been shown previously to decrease rabbit numbers as a vector of the myxoma virus (Cooke, 1983). These results are likely a result of a lack of spatial and temporally explicit data on the release, establishment and expansion of European fleas over the study period.

Spatial and temporally explicit estimations of abundance are an important component in ecological modelling, because a change in population abundance over time and space is brought about by changes in movement and/or birth and death rates. Identifying and quantifying the temporal variability in populations of invasive species can help to better target management actions (Simberloff, 2003) – particularly when responses to removal of invasive species happen on varying time scales (Cooke, 2012; Lurgi et al., 2018; Strayer et al., 2006), and different control measures will have differing population level effects (Wells, Cassey, et al., 2016). Just as importantly, quantifying variability in population abundances across the range of a species can unlock the ecological mechanisms responsible for spatiotemporal variation in abundance across a species geographic distribution (Caughley et al., 1988). This information is essential for optimising management actions aimed at controlling invasive species (Mellin et al., 2016).

The concept of a hotspot is focused on the stability of a species population, or a measure of biodiversity, across space and over time (Mittermeier, Turner, Larsen, Brooks, & Gascon, 2011). Beyond rabbits, our hotspots tool could be used to strengthen theory underpinning important concepts in invasion biology, including whether species' populations with high but variable abundance are more sensitive to control interventions than populations with high and less variable abundances (Wells, Cassey, et al., 2016). The hypothesis being, that populations that are consistent hotspots of abundance, are likely to be more resilient to environmental fluctuations and, therefore, better able to recover from management interventions if not completely eradicated (Mutze, Kovaliski, Butler, Capucci, & McPhee, 2010). Likewise, the tool could be used to test whether hotspots of abundance tend to occur more frequently at the centre of the range of an invasive species, as would be expected, based on theory (i.e., species abundance distribution; Lawton, 1993)

Identifying priority areas for control activities is a key component in managing invasive species, and hindcasts of hotspots such as those created here identify spatial clusters where a higher than expected number of rabbits can be found historically. Efforts towards

eradication and control of rabbits in areas that are consistently identified as hotspots (i.e. high proportion of time as a hotspot) should be given top priority and rabbits of all age classes should be targeted (Wells, Cassey, et al., 2016). Near term climate forecasts from regional climate models are now readily available at the temporal and spatial resolutions needed to effectively manage biodiversity in response to global change (Tabor & Williams, 2010). Integrating these forecasts into our hotspot analysis would allow areas for future rabbit management to be prioritised, based on knowledge of what the abundance pattern is projected to be given forecast climate conditions, with the caveat that inter-warren population competition and structure can occur at finer spatial resolutions than the 5km<sup>2</sup> spatial scale used in this study (Lurgi, Wells, Kennedy, Campbell, & Fordham, 2016). A comparable approach is commonly used in fisheries management to set catch rates in space and time, with good success (Hobday et al., 2010). Furthermore, these same spatiotemporal estimates of density and abundance can be integrated into spatially explicit models for informing pest management (Mellin et al., 2016), examining multispecies interactions (Lurgi et al., 2018), or for inferring patterns of spatial variation in demographic rates (Hone, 1999).

The spotlight count data in the Australian National Rabbit Database includes some studies set up to record trends in rabbit population abundances after active (e.g. warren ripping) or passive (e.g., release of RHDV) population control. The timing and extent of these management actions are not extensively documented, potentially biasing abundance records used in this analysis. While we tried to account for this variation by using data source as a detectability covariate, the absence of detailed information on type and timing of management intervention prevented a more thorough analysis of this potential source of uncertainty. This bias could possibly have resulted in lower frequencies of high abundance records for some areas, and therefore it is possible that some sites were not correctly identified as hotspots in years directly following management intervention. However, our CRFD hotspot thresholds are well below those shown to influence vegetation community composition (Mutze et al., 2016), and as such any identified hotspots are still valuable for management as these areas would still be identified as priority control sites at abundances lower than those known to alter, for example, vegetation communities.

Our modelling framework for rabbits is applicable to any species for which long-term spatially explicit monitoring data and high temporal resolution climate data are available. It is flexible with regards to spatial grain and extent, and benefits from the utilisation of computationally efficient machine learning methods for generating maps of spatial abundance, and for estimating variable interactions that link patterns of abundances to

potential drivers and processes. The long-term management of invasive species relies on informed future projections of abundance and a good understanding of the processes that underpin the structure and dynamics of a species past and present-day distribution in its invasive range. Tools such as those created in this study promise to help guide adaptive management approaches by permitting inferences to be made regarding likely future population trends at scales relevant to management.

## References:

- Apley, D. W. (2016). Visualizing the effects of predictor variables in black box supervised learning models. *arXiv preprint arXiv:1612.08468*.
- Aubry, P., Pontier, D., Aubineau, J., Berger, F., Léonard, Y., Mauvy, B., & Marchandea, S. (2012). Monitoring population size of mammals using a spotlight-count-based abundance index: How to relate the number of counts to the precision? *Ecological indicators*, 18, 599-607. doi:10.1016/j.ecolind.2012.01.019
- Ballinger, A., & Morgan, D. G. (2002). Validating two methods for monitoring population size of the European rabbit (*Oryctolagus cuniculus*). *Wildlife Research*, 29(5), 431-437. doi:10.1071/wr01055
- Barrio, I. C., Acevedo, P., & Tortosa, F. S. (2009). Assessment of methods for estimating wild rabbit population abundance in agricultural landscapes. *European Journal of Wildlife Research*, 56(3), 335-340. doi:10.1007/s10344-009-0327-7
- Bartolino, V., Maiorano, L., & Colloca, F. (2011a). A frequency distribution approach to hotspot identification. *Population Ecology*, 53(2), 351-359. doi:10.1007/s10144-010-0229-2
- Bartolino, V., Maiorano, L., & Colloca, F. (2011b). Frequency distribution curves and the identification of hotspots: response to comments. *Population Ecology*, 53(4), 603-604. doi:10.1007/s10144-011-0273-6
- Breiman, L. (2001). Random forests. *Machine learning*, 45(1), 5-32. doi:10.1023/a:1010933404324
- Buckland, S. T., Rexstad, E. A., Marques, T. A., & Oedekoven, C. S. (2015). *Distance sampling: methods and applications*. Heidelberg: Springer.
- Cabral, J. S., & Schurr, F. M. (2010). Estimating demographic models for the range dynamics of plant species. *Global Ecology and Biogeography*, 19(1), 85-97. doi:10.1111/j.1466-8238.2009.00492.x
- Caughley, G., Grice, D., Barker, R., & Brown, B. (1988). The edge of the range. *Journal of Animal Ecology*, 771-785.
- Cayuela, L., Gálvez-Bravo, L., Carrascal, L. M., & de Albuquerque, F. S. (2011). Comments on Bartolino et al. (2011): limits of cumulative relative frequency distribution curves for hotspot identification. *Population Ecology*, 53(4), 597. doi:10.1007/s10144-011-0272-7
- Channell, R., & Lomolino, M. V. (2000). Dynamic biogeography and conservation of endangered species. *Nature*, 403(6765), 84-86. doi:10.1038/47487
- Cooke, B. (1983). Changes in the Age-Structure and Size of Populations of Wild Rabbits in South Australia, Following the Introduction of European Rabbit Fleas, *Spilopsyllus cuniculi* (Dale), as Vectors of Myxomatosis. *Wildlife Research*, 10(1), 105-120. doi:10.1071/WR9830105
- Cooke, B. (2012). Rabbits: manageable environmental pests or participants in new Australian ecosystems? *Wildlife Research*, 39(4), 279-289. doi:10.1071/WR11166
- Cooke, B., Brennan, M., & Elsworth, P. (2018). Ability of wild rabbit, *Oryctolagus cuniculus*, to lactate successfully in hot environments explains continued spread in Australia. *Wildlife Research*, 45(3), 267-273. doi:10.1071/wr17177
- Cooke, B., Chudleigh, P., Simpson, S., & Saunders, G. (2013). The Economic Benefits of the Biological Control of Rabbits in Australia, 1950–2011. *Australian Economic History Review*, 53(1), 91-107. doi:10.1111/aehr.12000
- Cooke, B., Duncan, R. P., McDonald, I., Liu, J., Capucci, L., Mutze, G. J., & Strive, T. (2018). Prior exposure to non-pathogenic calicivirus RCV-A1 reduces both infection rate and mortality from rabbit haemorrhagic disease in a population of wild rabbits in

- Australia. *Transboundary and Emerging Diseases*, 65(2), e470-e477. doi:10.1111/tbed.12786
- Cox, G. W. (2013). *Alien Species and Evolution: The Evolutionary Ecology of Exotic Plants, Animals, Microbes, and Interacting Native Species*. Washington: Island Press.
- Delibes-Mateos, M., Delibes, M., Ferreras, P., & Villafuerte, R. (2008). Key role of European rabbits in the conservation of the Western Mediterranean basin hotspot. *Conservation Biology*, 22(5), 1106-1117. doi:10.1111/j.1523-1739.2008.00993.x
- Elith, J., Kearney, M., & Phillips, S. (2010). The art of modelling range-shifting species. *Methods in Ecology and Evolution*, 1(4), 330-342. doi:10.1111/j.2041-210X.2010.00036.x
- Fordham, D., Akçakaya, H., Araújo, M., Brook, B., Brodie, J., Post, E., & Doak, D. (2012). Modelling range shifts for invasive vertebrates in response to climate change. In Brodie J, Post E, & D. D (Eds.), *Wildlife Conservation in a Changing Climate*. Chicago: University of Chicago Press.
- Fordham, D., Akcakaya, H. R., Araujo, M. B., Keith, D. A., & Brook, B. W. (2013). Tools for integrating range change, extinction risk and climate change information into conservation management. *Ecography*, 36(9), 956-964. doi:10.1111/j.1600-0587.2013.00147.x
- Fordham, D., Akçakaya, H. R., Brook, B. W., Rodríguez, A., Alves, P. C., Civantos, E., . . . Araújo, M. B. (2013). Adapted conservation measures are required to save the Iberian lynx in a changing climate. *Nature Climate Change*, 3, 899. doi:10.1038/nclimate1954
- Gaston, K. J. (2003). *The structure and dynamics of geographic ranges*. New York: Oxford University Press.
- Gelman, A., Meng, X. L., & Stern, H. (1996). Posterior predictive assessment of model fitness via realized discrepancies. *Statistica sinica*, 6(4), 733-760.
- Gilbert, N., Myers, K., Cooke, B. D., Dunsmore, J. D., Fullagar, P. J., Gibb, J. A., . . . Wood, D. H. (1987). Comparative Dynamics of Australasian Rabbit-Populations. *Wildlife Research*, 14(4), 491-503.
- Guillera-Arroita, G. (2017). Modelling of species distributions, range dynamics and communities under imperfect detection: advances, challenges and opportunities. *Ecography*, 40(2), 281-295. doi:10.1111/ecog.02445
- Hobday, A. J., Hartog, J. R., Timmiss, T., & Fielding, J. (2010). Dynamic spatial zoning to manage southern bluefin tuna (*Thunnus maccoyii*) capture in a multi-species longline fishery. *Fisheries Oceanography*, 19(3), 243-253. doi:10.1111/j.1365-2419.2010.00540.x
- Hone, J. (1999). On rate of increase (r): patterns of variation in Australian mammals and the implications for wildlife management. *Journal of Applied Ecology*, 36(5), 709-718. doi:10.1046/j.1365-2664.1999.00439.x
- Joseph, L. N., Elkin, C., Martin, T. G., & Possingham, H. P. (2009). Modeling abundance using N-mixture models: the importance of considering ecological mechanisms. *Ecological Applications*, 19(3), 631-642. doi:10.1890/07-2107.1
- Kéry, M., & Schaub, M. (2011). *Bayesian Population Analysis using WinBUGS: A Hierarchical Perspective*. Oxford: Academic Press.
- Lawton, J. H. (1993). Range, population abundance and conservation. *Trends in Ecology & Evolution*, 8(11), 409-413. doi:10.1016/0169-5347(93)90043-O
- Link, W. A., Schofield, M. R., Barker, R. J., & Sauer, J. R. (2018). On the robustness of N-mixture models. *Ecology*, 99(7), 1547-1551. doi:10.1002/ecy.2362
- Liu, J., Fordham, D. A., Cooke, B. D., Cox, T., Mutze, G., & Strive, T. (2014). Distribution and prevalence of the Australian non-pathogenic rabbit calicivirus is correlated with

- rainfall and temperature. *PLOS ONE*, 9(12), e113976. doi:10.1371/journal.pone.0113976
- Lurgi, M., Ritchie, E. G., & Fordham, D. A. (2018). Eradicating abundant invasive prey could cause unexpected and varied biodiversity outcomes: The importance of multispecies interactions. *Journal of Applied Ecology*, 55(5), 2396-2407. doi:10.1111/1365-2664.13188
- Lurgi, M., Wells, K., Kennedy, M., Campbell, S., & Fordham, D. A. (2016). A Landscape Approach to Invasive Species Management. *PLOS ONE*, 11(7), e0160417. doi:10.1371/journal.pone.0160417
- Magurran, A. E. (2007). Species abundance distributions over time. *Ecology letters*, 10(5), 347-354. doi:10.1111/j.1461-0248.2007.01024.x
- Magurran, A. E., Baillie, S. R., Buckland, S. T., Dick, J. M., Elston, D. A., Scott, E. M., . . . Watt, A. D. (2010). Long-term datasets in biodiversity research and monitoring: assessing change in ecological communities through time. *Trends in Ecology & Evolution*, 25(10), 574-582. doi:10.1016/j.tree.2010.06.016
- Martin, T. G., Wintle, B. A., Rhodes, J. R., Kuhnert, P. M., Field, S. A., Low-Choy, S. J., . . . Possingham, H. P. (2005). Zero tolerance ecology: improving ecological inference by modelling the source of zero observations. *Ecology letters*, 8(11), 1235-1246. doi:10.1111/j.1461-0248.2005.00826.x
- Mellin, C., Lurgi, M., Matthews, S., MacNeil, M. A., Caley, M. J., Bax, N., . . . Fordham, D. A. (2016). Forecasting marine invasions under climate change: Biotic interactions and demographic processes matter. *Biological Conservation*, 204, 459-467. doi:10.1016/j.biocon.2016.11.008
- Mellin, C., Russell, B. D., Connell, S. D., Brook, B. W., & Fordham, D. A. (2012). Geographic range determinants of two commercially important marine molluscs. *Diversity and Distributions*, 18(2), 133-146. doi:10.1111/j.1472-4642.2011.00822.x
- Mittermeier, R. A., Turner, W. R., Larsen, F. W., Brooks, T. M., & Gascon, C. (2011). Global Biodiversity Conservation: The Critical Role of Hotspots. In F. E. Zachos & J. C. Habel (Eds.), *Biodiversity Hotspots: Distribution and Protection of Conservation Priority Areas* (pp. 3-22). Berlin: Springer
- Mutze, G., Bird, P., Kovaliski, J., Peacock, D., Jennings, S., & Cooke, B. (2002). Emerging epidemiological patterns in rabbit haemorrhagic disease, its interaction with myxomatosis, and their effects on rabbit populations in South Australia. *Wildlife Research*, 29(6), 577-590. doi:10.1071/wr00100
- Mutze, G., Cooke, B., & Jennings, S. (2016). Estimating density-dependent impacts of European rabbits on Australian tree and shrub populations. *Australian Journal of Botany*, 64(2), 142-152. doi:10.1071/bt15208
- Mutze, G., Kovaliski, J., Butler, K., Capucci, L., & McPhee, S. (2010). The effect of rabbit population control programmes on the impact of rabbit haemorrhagic disease in south-eastern Australia. *Journal of Applied Ecology*, 47(5), 1137-1146. doi:10.1111/j.1365-2664.2010.01844.x
- Pagel, J., Anderson, B. J., O'Hara, R. B., Cramer, W., Fox, R., Jeltsch, F., . . . McMahon, S. (2014). Quantifying range-wide variation in population trends from local abundance surveys and widespread opportunistic occurrence records. *Methods in Ecology and Evolution*, 5(8), 751-760. doi:10.1111/2041-210x.12221
- Pech, R. P., & Hood, G. M. (1998). Foxes, rabbits, alternative prey and rabbit calicivirus disease: consequences of a new biological control agent for an outbreaking species in Australia. *Journal of Applied Ecology*, 35(3), 434-453. doi:10.1046/j.1365-2664.1998.00318.x

- Phillips, B. L., Brown, G. P., Webb, J. K., & Shine, R. (2006). Invasion and the evolution of speed in toads. *Nature*, 439, 803. doi:10.1038/439803a
- Plummer, M. (2016). rjags: Bayesian Graphical Models using MCMC. Retrieved from <https://CRAN.R-project.org/package=rjags>
- R Core Team. (2017). R: A language and environment for statistical computing (Version 3.5.1). Vienna, Austria: R Foundation for Statistical Computing. Retrieved from <https://www.R-project.org/>
- Royle, J. A. (2004). N-mixture models for estimating population size from spatially replicated counts. *Biometrics*, 60(1), 108-115. doi:10.1111/j.0006-341X.2004.00142.x
- Schurr, F. M., Pagel, J., Cabral, J. S., Groeneveld, J., Bykova, O., O'Hara, R. B., . . . Zimmermann, N. E. (2012). How to understand species' niches and range dynamics: a demographic research agenda for biogeography. *Journal of Biogeography*, 39(12), 2146-2162. doi:10.1111/j.1365-2699.2012.02737.x
- Simberloff, D. (2003). How Much Information on Population Biology Is Needed to Manage Introduced Species? *Conservation Biology*, 17(1), 83-92. doi:10.1046/j.1523-1739.2003.02028.x
- Sofaer, H. R., Jarnevich, C. S., & Pearse, I. S. (2018). The relationship between invader abundance and impact. *Ecosphere*, 9(9), e02415. doi:10.1002/ecs2.2415
- Stodart, E., & Parer, I. (1988). *Colonisation of Australia by the rabbit Oryctolagus cuniculus (L.)*: Canberra, CSIRO.
- Strayer, D. L., Eviner, V. T., Jeschke, J. M., & Pace, M. L. (2006). Understanding the long-term effects of species invasions. *Trends in Ecology & Evolution*, 21(11), 645-651. doi:10.1016/j.tree.2006.07.007
- Strive, T., Elsworth, P., Liu, J., Wright, J. D., Kovaliski, J., & Capucci, L. (2013). The non-pathogenic Australian rabbit calicivirus RCV-A1 provides temporal and partial cross protection to lethal Rabbit Haemorrhagic Disease Virus infection which is not dependent on antibody titres. *Veterinary research*, 44(1), 51-51. doi:10.1186/1297-9716-44-51
- Tabor, K., & Williams, J. W. (2010). Globally downscaled climate projections for assessing the conservation impacts of climate change. *Ecological Applications*, 20(2), 554-565. doi:10.1890/09-0173.1
- Wells, K., Cassey, P., Sinclair, R. G., Mutze, G. J., Peacock, D. E., Lacy, R. C., . . . Fordham, D. A. (2016). Targeting season and age for optimizing control of invasive rabbits. *Journal of Wildlife Management*, 80(6), 990-999. doi:10.1002/jwmg.21093
- Wells, K., Fordham, D. A., Brook, B. W., Cassey, P., Cox, T., O'Hara, R. B., & Schwensow, N. I. (2018). Disentangling synergistic disease dynamics: Implications for the viral biocontrol of rabbits. *Journal of Animal Ecology*, 87(5), 1418-1428. doi:10.1111/1365-2656.12871
- Wells, K., O'Hara, R. B., Cooke, B. D., Mutze, G. J., Prowse, T. A., & Fordham, D. A. (2016). Environmental effects and individual body condition drive seasonal fecundity of rabbits: identifying acute and lagged processes. *Oecologia*, 181(3), 853-864. doi:10.1007/s00442-016-3617-2
- Whittaker, R. J., Araújo, M. B., Jepson, P., Ladle, R. J., Watson, J. E. M., & Willis, K. J. (2005). Conservation Biogeography: assessment and prospect. *Diversity and Distributions*, 11(1), 3-23. doi:10.1111/j.1366-9516.2005.00143.x
- Wilson, A. J., Kruuk, L. E., & Coltman, D. W. (2005). Ontogenetic patterns in heritable variation for body size: using random regression models in a wild ungulate population. *American Naturalist*, 166(6), E177-192. doi:10.1086/497441

Yoccoz, N. G., Nichols, J. D., & Boulinier, T. (2001). Monitoring of biological diversity in space and time. *Trends in Ecology & Evolution*, 16(8), 446-453. doi:10.1016/S0169-5347(01)02205-4

**Data Accessibility Statement:**

All output maps are present in netcdf format in Appendix I and can be accessed using GIS software (R, QGIS, ArcGIS etc.). Spotlight data used for the analysis was provided by Emilie Roy-Dufresne (University of Adelaide). Code for the hierarchical N-Mixture model is available on request from the corresponding author.

**Tables:**

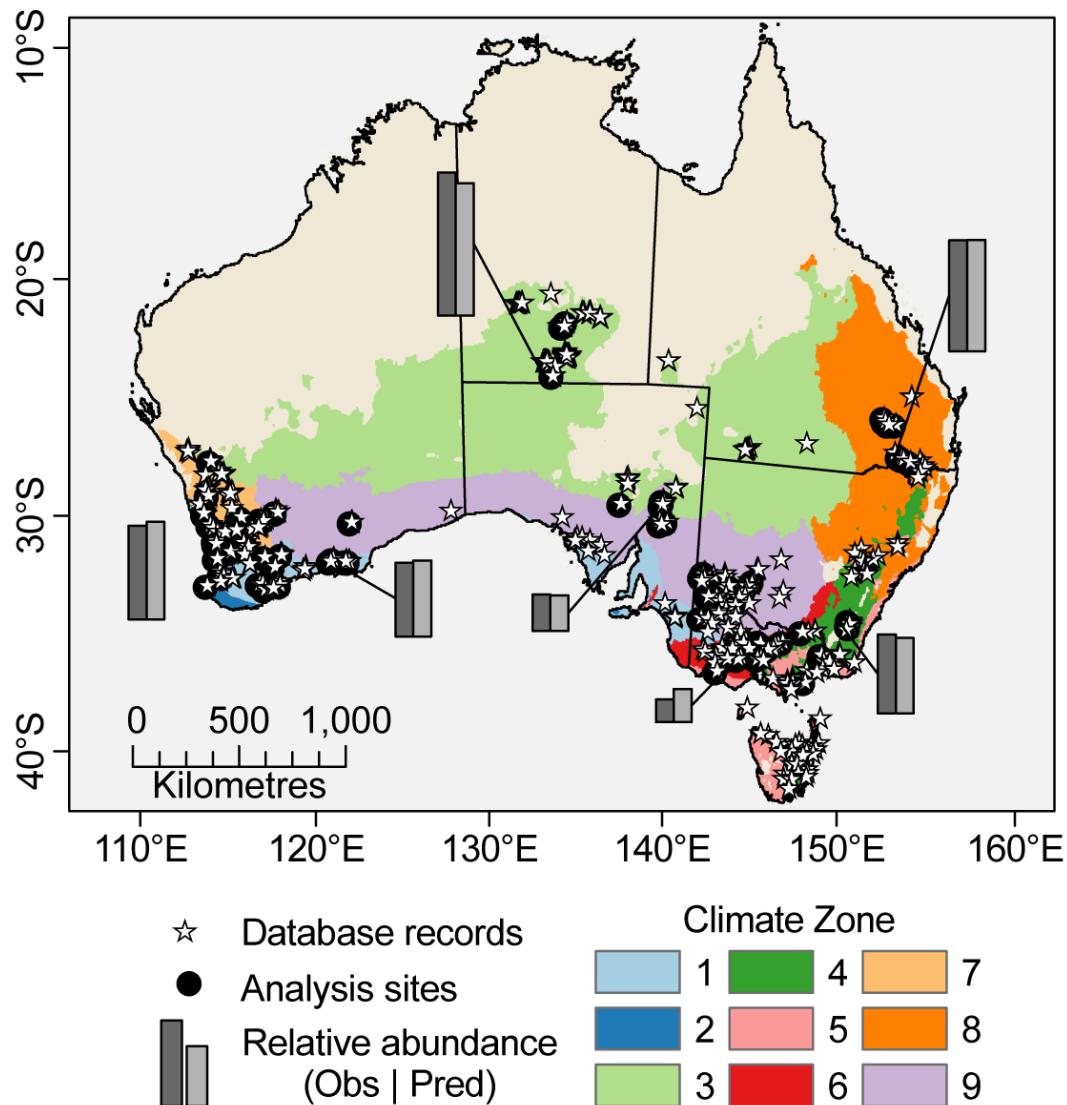
**Table 1:** Coefficient estimates ( $\pm$  S.D) for the climatic covariates used to model detection probability of rabbits during spotlight surveys in different climate zones. Descriptions of the climate zones are provided in Table S2.

<i>Climate Zone</i>	<i>Daily maximum temperature</i>	<i>Daily minimum temperature</i>	<i>Daily Precipitation</i>
1	-0.02 (0.10)	0.08 (0.11)	0.05 (0.05)
2	-0.34 (0.15)	0.04 (0.11)	0.03 (0.05)
3	0.52 (0.17)	0.00 (0.11)	0.07 (0.13)
4	-0.30 (0.21)	-0.02 (0.14)	0.04 (0.12)
5	0.02 (0.41)	0.03 (0.17)	0.05 (0.14)
6	-0.28 (0.23)	0.00 (0.10)	0.05 (0.09)
7	-0.20 (0.17)	0.01 (0.13)	0.01 (0.10)
8	0.26 (0.17)	0.09 (0.12)	0.08 (0.11)
9	0.30 (0.13)	0.07 (0.10)	0.07 (0.09)

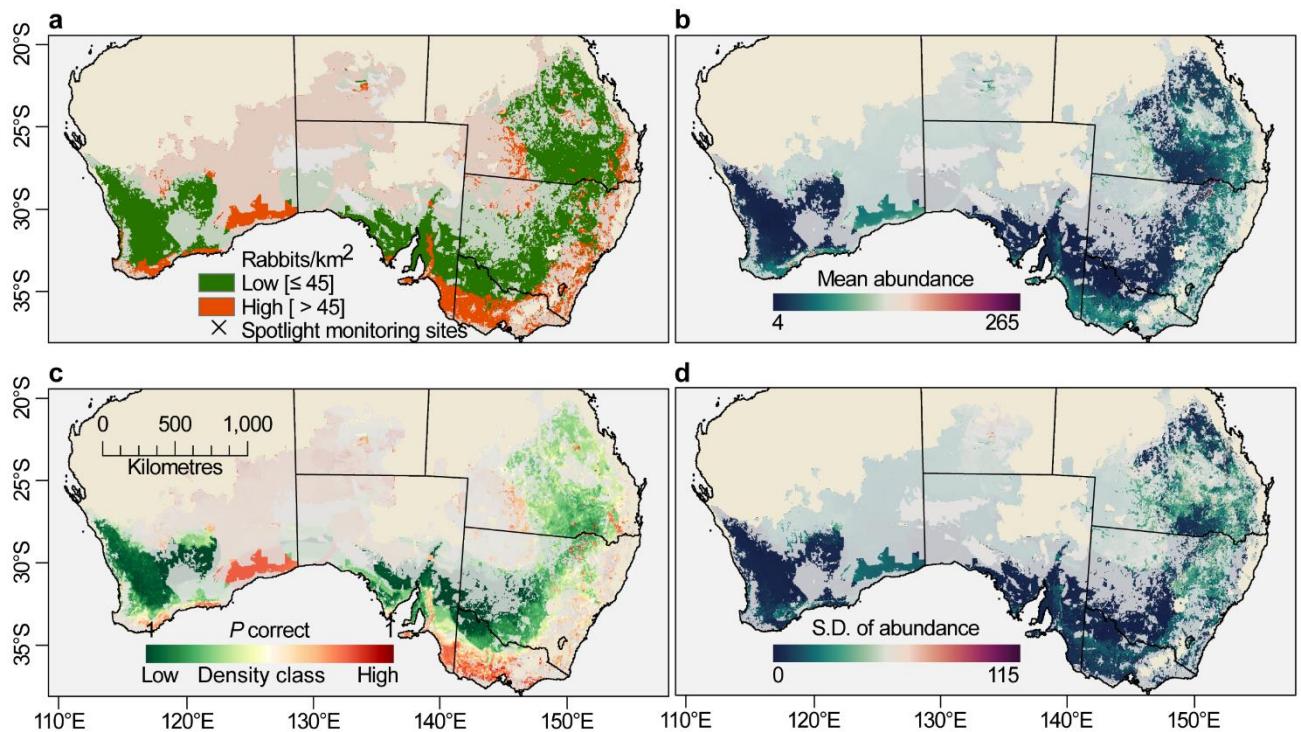
**Table 2:** Performance of the random forest regression and classification models. Cross validated model performance was assessed using 1000 repeats of Monte-Carlo cross validation with an 80/20 (train/test) split. Accuracy of both the regression and classification models was excellent across climate zone and in most cases fair to excellent within climate zones. For climate zones 5 and 7, mean balanced accuracy and true skill statistics could not be calculated as the classification model predicted all samples as belonging to the same class.

		Classification			Regression		
<i>Overall Climate Zone</i>	Sensitivity	Specificity	Mean Balanced Accuracy	True Skill Statistic	Mean observed	Mean predicted	Mean absolute error
	0.86	0.85	0.86	0.71	6.4 ( $\pm 11.9$ )	4.9 ( $\pm 6.4$ )	1.5
1	0.68	0.76	0.72	0.44	8.7 ( $\pm 13.9$ )	6.2 ( $\pm 6.6$ )	2.6
2	0.87	0.50	0.68	0.37	5.7 ( $\pm 9.9$ )	4 ( $\pm 5.8$ )	1.7
3	0.60	1.00	0.80	0.60	14.7 ( $\pm 20.1$ )	10 ( $\pm 8.2$ )	4.7
4	0.49	0.87	0.68	0.36	11.5 ( $\pm 16.3$ )	8.6 ( $\pm 6.4$ )	2.9
5	NA	1.00	NA	NA	7.4 ( $\pm 7.3$ )	6.2 ( $\pm 1.9$ )	1.2
6	0.67	0.90	0.78	0.57	5.3 ( $\pm 6.9$ )	5.2 ( $\pm 4.3$ )	0.1
7	1.00	NA	NA	NA	1.6 ( $\pm 2.5$ )	1.4 ( $\pm 1.1$ )	0.1
8	0.88	0.68	0.78	0.56	12.9 ( $\pm 15.4$ )	11 ( $\pm 11$ )	2.0
9	0.99	0.50	0.75	0.49	2.1 ( $\pm 4.1$ )	1.6 ( $\pm 1.4$ )	0.6

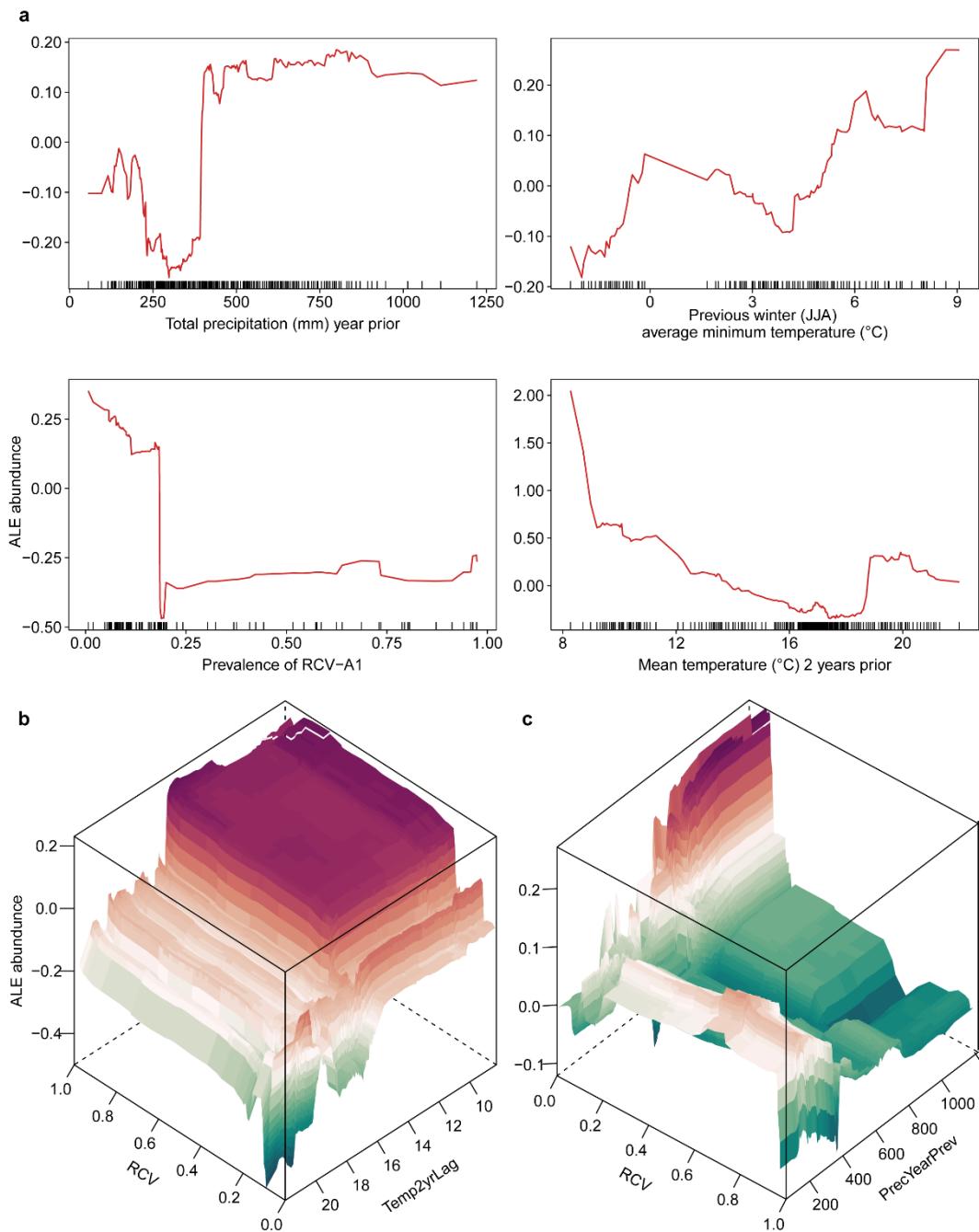
**Figures:**



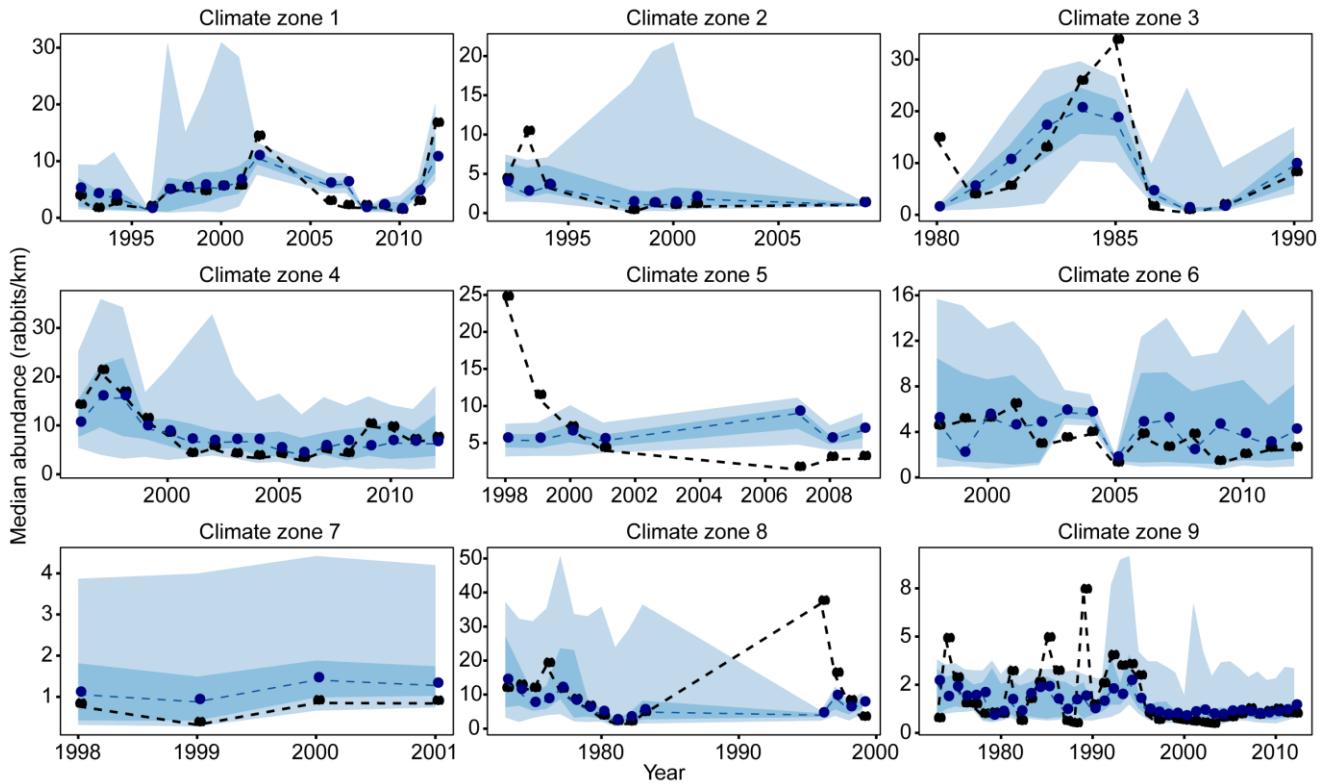
**Figure 1: Map showing the location of all spotlight rabbit abundance records (1971-2012) and the sites used in the final analysis.** Column plots show the relative magnitudes for observed (dark grey) and Bayesian model corrected abundance (light grey) for a selection of sites. The different colours represent the 9 climatic zones that were used in our analysis. Descriptions of the climate zones are in Table S2. The beige colour represents areas that are outside of the climatic envelope of our sampling sites. Map is projected in GDA 1994 Australian Albers (EPSG:3577).



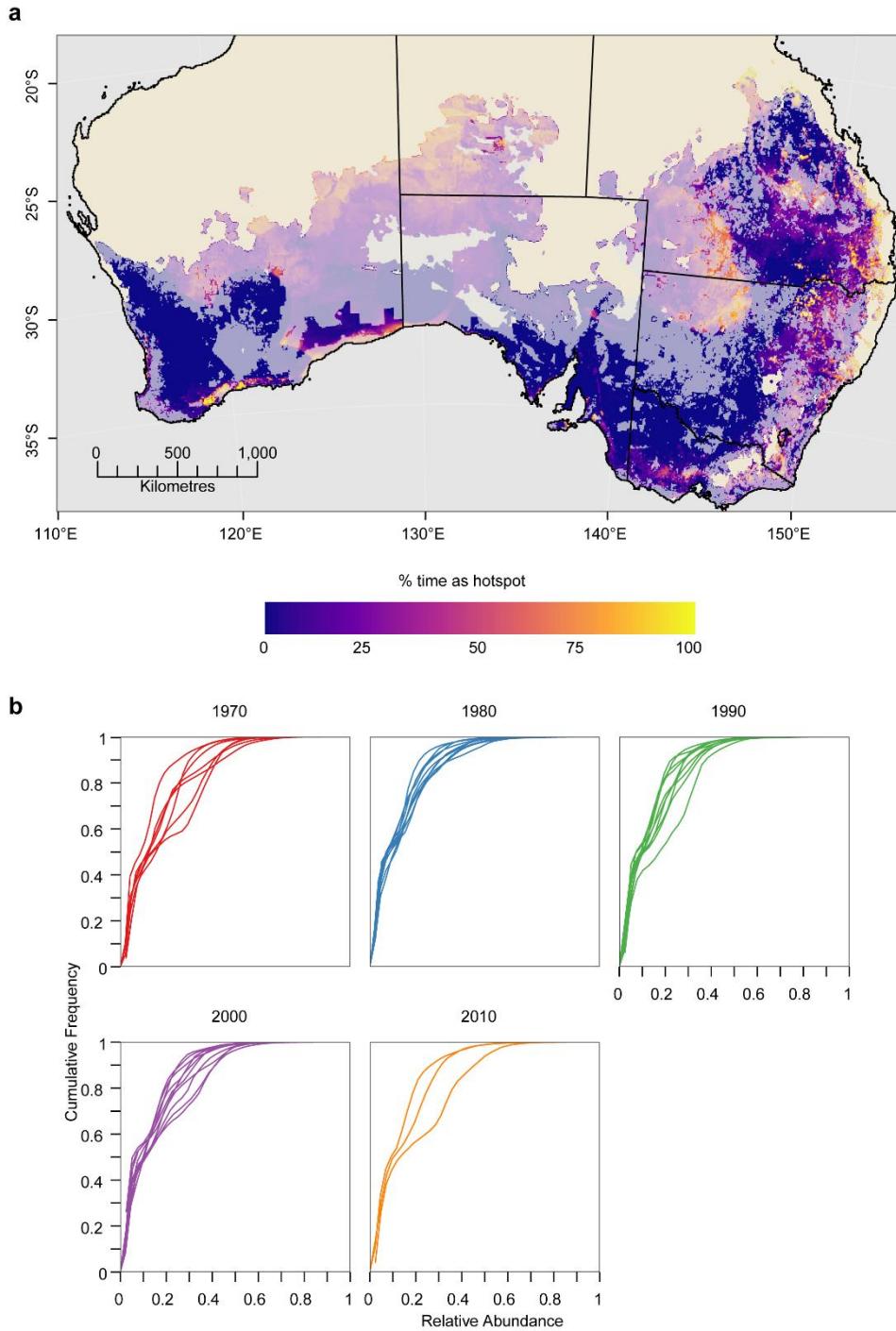
**Figure 2: Spatial predictions of rabbit densities and abundances inferred from the random forest models.** (a) Spatial variation in the mode of predicted corrected densities between 1972 and 2012. (b) Mean corrected abundance (rabbits/km of transect) between 1973 and 2012. (c) The probability of a correct model classification (low/high), with darker colours (dark green/dark red) indicative of high confidence in the predicted class. (d) The standard deviation of the predicted corrected abundances between 1973 and 2012. All maps are masked to the climatic envelope of the climate zones shown in Fig. 1. The muted colours represent areas with low prediction confidence i.e., pixels where the model is forced to extrapolate beyond the climate and environmental envelope used to calibrate the model. Maps are projected in GDA 1994 Australian Albers (EPSG:3577).



**Figure 3. Accumulated Local Effects (ALE) plots of four important variables depicting the variation in rabbit abundance across environmental gradients.** Values on the y-axes represent changes in abundance from the average sample conditional on the values on the x-axes (**a**). The rug plots on the inside of the x-axis show the distribution of sites across that variable. The 3-d plots (**b**, **c**) show the second-order effects of benign RCV-A1 calicivirus with two climatic variables; average temperature 2 years prior to the survey (**b**), and total precipitation the year prior to the survey (**c**), after removing the main effects of the features. In other words, panels **b** and **c** only show the additional interaction effect of the two features. The colours correspond to the values on the z-axis, with reds indicating a more positive effect on abundances



**Figure 4: Empirically corrected abundance estimates (i.e. corrected; black line) and projected (blue line) abundance estimates of rabbits for nine selected climate zones in Australia.** Points display the median estimates of corrected abundance estimates and predicted abundances (based on the random forest regression model) across all sites within a climate zone for a given year. The lines are for connecting points between sampling years, but do not represent the estimates between years, and show the discontinuous nature of the sampling within some climate zones. The blue bandings represent the 95% (light blue) and 50% (dark blue) predictions intervals for the regression model predictions within each zone.



**Figure 5: Percent time spent in hotspot conditions for high rabbit abundances.** Values for each cell (a) represent the proportion of time (between 1973 and 2012) that cells were considered “hotspots” of rabbit abundance. The map is masked to the climatic envelope of the climate zones shown in Fig. 1. Muted colours represent areas with low prediction confidence i.e., pixels where the RF model was forced to extrapolate beyond the climate and environmental envelope used to calibrate the model. Map is projected in GDA 1994 Australian Albers (EPSG:3577). The cumulative relative distribution frequency plots (b) show the inter-annual variability in rabbit populations for the RF predicted abundances that were used to define hotspot thresholds.