Diversity and Distributions, (Diversity Distrib.) (2014) 1–10

Diversity and Distributions A Journal of Conservation Biogeography

BIODIVERSITY RESEARCH

Distinguishing geographical range shifts from artefacts of detectability and sampling effort

Amanda E. Bates[1,2]*, Tomas J. Bird[3,4], Rick D. Stuart-Smith[2], Thomas Wernberg[5], Jennifer M. Sunday[6], Neville S. Barrett[2], Graham J. Edgar[2], Stewart Frusher[2], Alistair J. Hobday[7], Gretta T. Pecl[2], Dan A. Smale[5,8] and Michael McCarthy[3]

[1]*Ocean and Earth Sciences, National Oceanography Centre Southampton, University of Southampton Waterfront Campus, Southampton SO14 3ZH, UK,* [2]*Institute for Marine and Antarctic Studies, University of Tasmania, Hobart, TAS 7001, Australia,* [3]*School of Botany, The University of Melbourne, Parkville, Vic. 3010, Australia,* [4]*Geography and Environment, University of Southampton, Southampton SO17 1BJ, UK,* [5]*UWA Oceans Institute and School of Plant Biology, University of Western Australia, Crawley, WA 6009, Australia,* [6]*Earth to Ocean Research Group, Department of Biological Sciences, Simon Fraser University, Burnaby, BC V5A 1S6, Canada,* [7]*Climate Adaptation Flagship, CSIRO Marine and Atmospheric Research, Hobart, TAS 7001, Australia,* [8]*Marine Biological Association of the United Kingdom, Plymouth PL1 2PB, UK*

*Correspondence: Amanda E. Bates, Ocean and Earth Sciences, National Oceanography Centre Southampton, University of Southampton Waterfront Campus, Southampton SO14 3ZH, UK.
E-mail: A.E.Bates@soton.ac.uk

ABSTRACT

Aim The redistribution of species with climate change is well documented. Even so, the relative contribution of species detectability to the variation in measured range shift rates among species is poorly understood. How can true range shifts be discerned from sampling artefacts?

Location Australia.

Methods We simulate range shifts for species which differ in their abundance for comparison to patterns derived from empirical range shift data from two regional-scale (100s km) empirical studies. We demonstrate the use of spatial occupancy data in a distance-to-edge (DTE) model to assess changes in geographical range edges of fish species within a temperate reef fish community.

Results Simulations identified how sampling design can produce relatively larger error in range shift estimates in less abundant species, patterns that correspond with those observed in real data. Application of the DTE model allowed us to estimate the location of the true range edge with high accuracy in common species. In addition, upper confidence bounds for range edge estimates identified species with range edges that have likely shifted in location.

Conclusions Simulation and modelling approaches used to quantify the level of confidence that can be placed in observed range shifts are particularly valuable for studies of marine species, where observations are typically few and patchy. Given the observed variability in range shift estimates, the inclusion of confidence bounds on estimates of geographical range edges will advance our capacity to disentangle true distributional change from artefacts of sampling design.

Keywords
Climate warming, extreme value statistics, range edge estimation, sampling methodology.

INTRODUCTION

Species are tracking environmental warming by moving towards the poles through range extensions at the leading (poleward) range boundary and range contractions at the trailing (equatorward) range boundary (Chen *et al.*, 2011; Sunday *et al.*, 2012; Poloczanska *et al.*, 2013). Yet even in areas experiencing rapid temperature change, the magnitude of observed biological responses differs among species (Angert *et al.*, 2011; Pinsky *et al.*, 2013; Poloczanska *et al.*, 2013). While some variation can presumably be attributed to differences among species in their responsiveness to temperature, the magnitude and spatio-temporal distribution of sampling effort also has the potential to generate inaccurate range shift estimates. This is a particular problem for species unlikely to be observed due to low population numbers, patchy occupancy patterns or cryptic characteristics (Dorazio & Royle, 2005; Shoo *et al.*, 2006; Blanchard *et al.*, 2008; Hassall & Thompson, 2010; McCarthy *et al.*, 2013). Even when entire assemblages of species are systematically surveyed with the

DOI: 10.1111/ddi.12263
http://wileyonlinelibrary.com/journal/ddi

same effort, error in observed range shifts can be expected due to the occupancy patterns and traits of different species.

Difficulties in measuring species distribution patterns are well known and have been acknowledged in the context of range shifts (reviewed in Tingley & Beissinger, 2009). In particular, the middle of species distribution ranges can be estimated with greater accuracy than range edges, especially for rarer species (Shoo *et al.*, 2006; Hassall & Thompson, 2010; Pinsky *et al.*, 2013). Yet range edges may be of particular interest (Sexton *et al.*, 2009); for example, range extension into new regions and contraction from previously occupied areas lead to species gains and losses that have the potential to change community dynamics (reviewed in Bates *et al.*, 2014). The processes that drive range extensions at the leading edge and contractions at the trailing edge also differ, so mechanistic understanding requires information specific to each edge type (Bates *et al.*, 2014). At the limit of species geographical ranges, local densities are typically lowest, and therefore, detectability, the capacity to detect a species when it is present, is also expected to decline (McCarthy *et al.*, 2013). It is therefore important to understand the type and quantity of error in estimates of range edges.

However, while biases due to variable species detectability have been acknowledged when interpreting observed range shift patterns, biases are generally not accounted for using statistical methods (Tingley & Beissinger, 2009; Brown *et al.*, 2011; Tanadini & Schmidt, 2011; Monk, 2013). For instance, Hassall & Thompson (2010) have suggested that the position of the range edge can be estimated by fitting the gamma frequency distribution to latitudinal occurrence data, rather than the most extreme or averaged observed location of the range edge (e.g. mean of the 10 most extreme range records, as in Hickling *et al.*, 2005; Thomas & Lennon, 1999). Alternatively, to increase the probability that observed range edge shifts represent true distributional change, Jones *et al.* (2010) suggest considering only shifts that are greater than a threshold distance, set for example by the upper 95th percentile of the distances among survey sites. However, both of these examples assume that all species are equally likely to be detected if present and do not provide a means for evaluating the level of confidence that can be placed in a particular range edge observation (Wintle *et al.*, 2013). While quantifying species detectability using distance sampling and employing occupancy modelling techniques would allow for statistical characterization of the range edge, to our knowledge, this technique has only been attempted for butterflies (Isaac *et al.*, 2011), a model taxonomic group for understanding climate-driven range shifts (Parmesan *et al.*, 1999).

Evaluation of range shifts in the ocean, for pragmatic reasons, has generally ignored species detectability, indirectly assuming uniformly high detectability across species, despite the fact that detectability can vary markedly among similar species even when under rigorous sampling protocols (e.g. within reef fish assemblages, MacNeill *et al.*, 2008). This is likely because approaches for quantifying species detectability require a level of replication that is often costly in marine environments, especially at regional scales.

Here, we first assess the extent to which sampling variability contributes to error in estimates of species range shifts for species with varying abundances. We use simulations to establish patterns in range shift estimates that relate to the probability of occupancy and sampling design. We illustrate the use of simulations to understand community-level patterns by providing two examples that demonstrate evidence of detection-related biases and variability in range shift estimates, as predicted by our simulations.

Second, we explore a new method for quantifying uncertainty in range edge estimates. We apply a time-to-extinction model, an optimal linear estimator tool (Solow, 2005), using spatial occurrence data for fishes with a known geographical range edge. Time-to-extinction models are based on the timing of observations leading up to the last sighting, and are commonly used in contexts related to palaeontology and conservation biology for declining processes which can be modelled with a Weibull distribution (as described in Roberts & Solow, 2003 and Solow, 2005). In the same way that the last sighting of an individual from a near-extinct species is unlikely to represent the very last individual of a population (except where the entire population is known), the most extreme location at which a species is observed is unlikely to represent its true range edge (although there may be exceptions, such as where known habitat barriers exist). Exchanging space for time therefore allows us to estimate the true location of the range edge and quantify the upper confidence bound for this estimate. We test the accuracy of this application using real data and introduce the term 'distance-to-edge' model (DTE).

METHODS

In the marine realm, species abundance is one of the more important determinants of both site occupancy (proportion of sites in which a species is present) and detectability (probability of detecting a species if it is present in a site; McCarthy *et al.*, 2013). Abundance is also a convenient descriptor that varies by orders of magnitude within and between species, and can be assessed categorically (i.e. rare versus common) which means that bias or error related to abundance can be evaluated for diverse data. Our simulations characterize variation in leading and trailing range edge estimates due to abundance-related sampling error. However, any other process that would lead to a reduced probability of observing a species near the edge of its distribution, as described above, would result in similar patterns.

Simulations

As range edges were of interest, we simulated the tails of the species abundance distributions. While one or more peaks in abundance can be expected across the range of a species (McGill & Collins, 2003), the shape and number of peaks in the middle of a species range are irrelevant to the present analysis. We therefore simulated abundance distributions with a single peak and tails in which abundance was lower than in the peaks (McGill & Collins, 2003).

For simulations of sampling across a species latitudinal range, we created location-specific probabilities for observing a species given that it is present. To do so, we used R Development Core Team (2011) to first define a standard Gaussian function (although different distributions were tested and produced similar patterns) centred at an arbitrary latitude (+10°) and with a standard deviation of 2° latitude (increasing the standard deviation increases tail length). We multiplied the standard Gaussian function by abundance values of 1 to 500 with a step size of 0.01. This produced, for each level of abundance, a distributional curve representing a probability of occupancy (Fig. 1). Observed presence or absence of a species at each sampled position along this gradient could then be modelled as a Bernoulli trial, with the probability of successfully observing a species defined by the corresponding probability of occupancy specific to each sampling position. Those species occurring with an abundance of 50 or greater at any sampling position were assigned a probability of occupancy of 1. Within the distribution for each simulated species (which we defined as the centre of the Gaussian distribution with the tail length defined by the 2.5 and 97.5 quantiles), we set the minimum observation probability of 0.001 (Appendix S1 in Supporting Information). The probability of occupancy outside the simulated distribution was 0.

We designed two different scenarios (Appendix S1 provides our R code) where range change was simulated 1000 times for each abundance level. In the first scenario, characterization of a species latitudinal distribution was achieved by simulating a stratified sampling design with equivalent effort in time. Fifty positions (sampling effort = 50) were sampled at odd intervals from 1 to 99 to represent a spatial gradient. We repeated the same design twice with no underlying change in distribution, to simulate stability in the range edge location between two sampling time periods. In the second scenario, we displaced the entire distribution for each abundance level by one degree of latitude polewards (Fig. 1). In this second scenario, we manipulated sampling effort in two ways to reflect known biases in a macroalgal data set (Wernberg *et al.*, 2011, described below). First, we randomly excluded up to 10 sampling positions to achieve uneven effort among simulations and across the spatial gradient. Second, we introduced a historical sampling bias. To do so, we biased overall effort by an average of 2 samples by randomly adding −1 to 3 for pre-shift (historical) sampling and −3 to 1 for the post-shift (present) sampling for each simulation. Taken together, total effort for the pre-shift and post-shift sampling ranged from, respectively, 39–53 and 37–51.

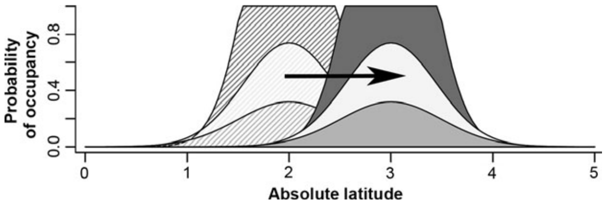

Figure 1 Schematic diagram of a theoretical relationship between probability of occupancy and latitude at two time periods (historical = hatched, present = filled) for a species under three abundance levels (high: dark grey, medium: grey, low: light) across absolute latitude (to represent both the northern and Southern Hemispheres). A ‘peak-and-tail’ pattern is expected if range limits are set by environmental conditions (although this may include multiple peaks). More abundant species have a higher probability of occupancy and are more likely to be observed during sampling. The leading (poleward, high latitude) and trailing (equatorward, low latitude) range edges are predicted to shift towards the poles (black arrow) with increasing climate warming. This leads to an extension (positive shift in latitude or gain) at the leading edge and contraction (negative shift in latitude or loss) at the trailing boundary.

In each range change simulation, the leading edge (for comparison to the fish data set, see below) and trailing edge (for comparison to the macroalgal data set, see below) were recorded, respectively, as the most extreme equatorward and poleward latitudes in which observations were recorded. Simulated range change was quantified as the difference in the pre-shift and post-shift range edge estimates returned for each simulation (Appendix S1).

Comparison of simulated and observed community scenarios

We compared our simulation results with two data sets to determine whether the abundance-related patterns identified in our simulations were present at the community level in the marine environment. Our objective was to ascertain the utility of simulated range shift data for understanding what errors and biases may be present in real data sets.

The first data set was limited to southeast Australia (Tasmania) and comprised range shift responses in shallow reef fishes near their poleward geographical limit. Community data were collected using standardized protocols for underwater visual census methods in which local abundances were recorded for each species (Stuart-Smith *et al.*, 2010). Sampling effort was even in 1994 and 2006, spanned from −43.58°S to −39.21°S ($n = 108$ sites) (Stuart-Smith *et al.*, 2010). The relationship between the measured difference in leading range edge (measured as the most poleward latitude at which each species was recorded) in the two sampling intervals and their mean abundance was determined for 66 fish species.

The second data set comprised 45 seaweed species from the southwest Australian coastline where shifts in the trailing range edge were of interest (Wernberg *et al.*, 2011). Latitudinal displacement was determined from ~1950 (pre-shift) to ~2000 (post-shift) on the basis of opportunistic collections in herbaria where sampling effort was episodic and varied among species (Wernberg *et al.*, 2011). As this data set did not contain information on the abundance of the species, we asked experts to score each species on a qualitative scale of 1 to 5 with respect to their expected relative abundance across the region (see Appendix S2 in Supporting Information: pooled into low (1–2.5) and high (> 2.5) categories for presentation). Based on our simulation results, we expected that abundant species would be over-represented in the pool of species with adequate data to quantify a change in the range edge location (simply because species with a lower probability of occupancy are missed in surveys). We further tested for the influence of abundance on the magnitude of range shifts measured in macroalgae. We used a generalized linear model (link function = log) and included sampling effort (total sample size) and temporal sampling bias (log ratio of effort between the two sampling periods) as covariates. Sampling effort and bias are reported in Wernberg *et al.* (2011) and are known to influence range edge detection (Shoo *et al.*, 2006).

Distance-to-edge (DTE) model of confidence in range edges

To provide an option for evaluating confidence in range edge change estimates, we explore the utility of an optimal linear estimator approach for estimating confidence in range edges that is dependent only on the shape of the tail of the range distribution. Based on the observation that the tails of many probability distributions generally match a Weibull distribution, we first assume that the same applies to the tails of species ranges. A similar assumption is made in the time-to-extinction models described in full in Solow (2005) and made accessible as the R package "sExtinct" in Clements, 2012. Briefly, the time-to-extinction models use an optimal linear endpoint estimate based on the spacing of the H last sightings on record. Here we interpose spatial for temporal measures, to create what we call a distance-to-edge (DTE) model.

Rivadeneira *et al.* (2009) used simulations to evaluate several time-to-extinction models and found that the 'R&S' model, which does not assume even distribution of sampling effort (Roberts & Solow, 2003), was most conservative. Thus, we chose the R&S model to fit the last H observations of a species range (where H ranged from 4 to 9, depending on the total number of samples available) to a Weibull distribution (Roberts & Solow, 2003) to estimate the true range edge and 95% confidence intervals around that edge. In fitting this model, we also assume that occupancy is a linear function of habitat variables that change across space and that detectability (not occupancy) is uniform.

While other sources have described the temporal model, we provide an interpretation of the DTE model terms based on the equations presented in Solow (2005). First, let the $S_1, \ldots, S_n$ represent the locations of observations of a particular species along a spatial gradient. The estimated distance from a species last observed location to its true range edge (S_E) is given by:

$$S_E = \sum_{i=1}^{H} w_i S_{n-i+1}$$

where S_n is the most distant observation (for example the southernmost range observation). Thus, the sequence, $S_{n-H+1,\ldots,}S_n$, represents the H most distant sightings, with H usually chosen to be between 4 and 9 (Rivadeneira *et al.*, 2009; Clements, 2012). We do not need to know n, but assume that it is large. Vector $\boldsymbol{w}$ is used to weight the influence of the H observations according to how they are spaced, and is derived from the following:

$$w = (k'\Lambda^{-1}k)^{-1}\Lambda^{-1}$$

where Λ is a symmetric H by H matrix and k is a vector of H ones, and the notation k' indicates the transpose of matrix k. The elements of Λ are given by:

$$\Lambda_{i,j} = \frac{\Gamma(2\hat{v}+i)\Gamma(\hat{v}+j)}{\Gamma(\hat{v}+i)\Gamma(j)}, j \leq i$$

where Γ is the standard Gamma function and $\hat{v}$ is an estimate of the shape parameter of the Weibull distribution, against which are fitted the last H sightings,

$$\hat{v} = \frac{1}{H-1}\sum_{i=1}^{H-2}\log\frac{S_n - S_n - H + 1}{S_n - S_{i+1}}$$

The upper 95% confidence interval of a species range edge is given by:

$$S_{CiU} = S_n + \frac{S_n - S_n - H + 1}{1 - c(\alpha/2)}$$

where (α)is the significance level (here set to be 0.05 to allow for 95% confidence intervals) and the function

$$c(\alpha/2) = \left[\frac{-\log(\alpha/2)}{H}\right]^{-\hat{v}}$$

scales the confidence intervals according to how many observations are used.

Because we are interested in comparing change in the location of range edges between two time periods, we also need the lower 95% confidence interval for the estimated range edge, which is given by substituting c(1-α/2) for c(α/2) in the equation for S_{CiU}

$$S_{CiL} = S_n + \frac{S_n - S_n - H + 1}{1 - c(1 - \alpha/2)}$$

Calculating the lower bounds for range edges allows us to define a range shift as occurring with high confidence if the 95% confidence intervals for the two time periods do not overlap.

To parameterize the DTE model, occupancy was estimated as the number of sites in which a species was observed within 0.1° of latitude across Tasmania, quantified for 28 species that occurred in both sampling years and were present in at least four locations in each year. The reef fish data set offers the benefit of having a known geographical range limit for shallow marine species at the southern end of Australia (the continental margin).

RESULTS

Simulated range shifts

When simulations did not include a latitudinal displacement in the location of the range edge and sampling was evenly distributed across space and time, simulated estimates of change in the location of the range edges were markedly more variable in rare versus abundant species (Fig. 2a). Greater variability in estimates of range change for rarer species means that a greater proportion of false shifts (range extensions and contractions) are observed if a method such as defining a shift as being greater or less than arbitrary value, such as was used here (± 0.45° of latitude), is selected (Fig. 2a).

Greater variability in estimates of range change at leading edges was also observed in the data set of reef fish abundance, where other sources of variability could conceivably swamp variation due to abundance-related sampling error. The magnitude and direction of range change spanned from 4 to −4° latitude for species with a peak mean abundance of less than three individuals per site, while estimates for species that were more abundant converged on zero. This finding supports the hypothesis that assessments for less abundant species are inaccurate and that large range shifts in both directions will be observed in species that are less abundant due to sampling variability alone (Fig. 2). In fact, observed range changes in the fish data set fell within the 5th and 95th quartiles of simulated range change estimates (Fig. 2b).

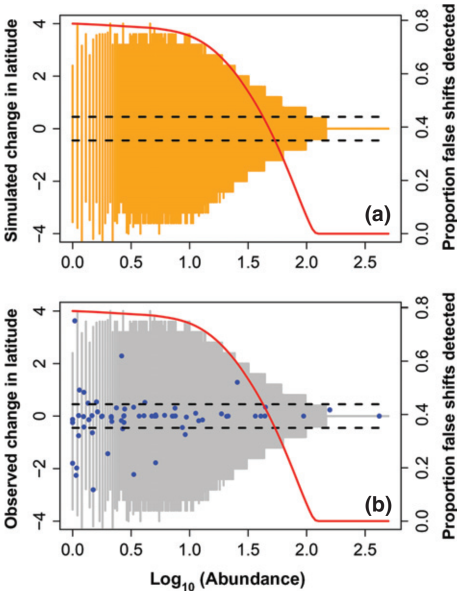

Figure 2 Relationship between abundance, change in latitude and proportion false range extensions detected. (a) 5th and 95th quartiles (orange, left *y*-axis) of simulated estimates of change in the location of leading range edges (sampling effort and species distributions were held constant) for mean peak abundances of 1 to 500 steps of 0.01. Proportion of false range shifts (red line, right *y*-axis) > 0.45° (dotted black lines) in either the positive or negative direction of 1000 simulations at each abundance step. (b) Observed change in the latitude of the leading range edge for 66 fishes from Tasmania (blue, left *y*-axis). Results from the simulations in panel a are shown (grey bars and red line) for direct comparison to the field data.

When we modelled the sampling process with uneven effort among simulations and with greater historical effort, we observed greater overall variability in range change estimates (compare Fig. 2a to Fig. 3a). As we displaced the latitudinal distribution by 1° polewards, we calculated, for each level of abundance, the proportion of true range shifts detected (number of range change estimates with an absolute value of > 0.45° divided by the number of range change estimates returned of 1000 simulations). Under our simulation parameters (Appendix S1), we found that the proportion of true range shifts detected increased with abundance (Fig. 3b). Similarly, in the macroalgal data set, contractions at the trailing range edge > 0.45° were also less evident in species with relatively low abundance (Fig. 3b). However, the magnitude of range contractions was not significantly related to qualitative abundance (generalized linear model $P > 0.1$ Table 1) when methodological covariates that influence detection of true range shifts were included in the model (overall sampling effort and change in sampling effort between the two sampling time intervals).

Our simulations further revealed that a larger proportion of rarer species go unobserved in surveys due to the stochastic nature of the sampling (Fig. 3c). For instance, in 1000 simulations of range edge locations in two time periods for a species with a mean abundance of one, observations in both sampling intervals were detected in only 73 simulations. Similarly, in the macroalgal range shift data set, the proportion of species scored with low abundance was 0.2 (Fig. 3c)

Distance-to-edge (DTE) model of confidence in range edges

In applying the DTE model to the reef fish data, confidence intervals of leading range edge estimates from 2006 and 1994 substantially overlapped for most species (Fig. 4a). There was thus little evidence for change in the leading range limits of the Australian fishes observed during the study period, with the exception of three species. A range edge extension (south) in one species (*Enoplosus armatus*) and contractions (north) in two species (*Atypichthys strigatus* and *Upeneichthys vlamingii*) were supported by non-overlapping confidence intervals. For those species observed at only a few sites (i.e. presence at 3 to 5 sites in each of the two years), the location of estimated range edges was farther south of the last observed presence and the confidence limits were wider in comparison with species with higher occupancy (compare

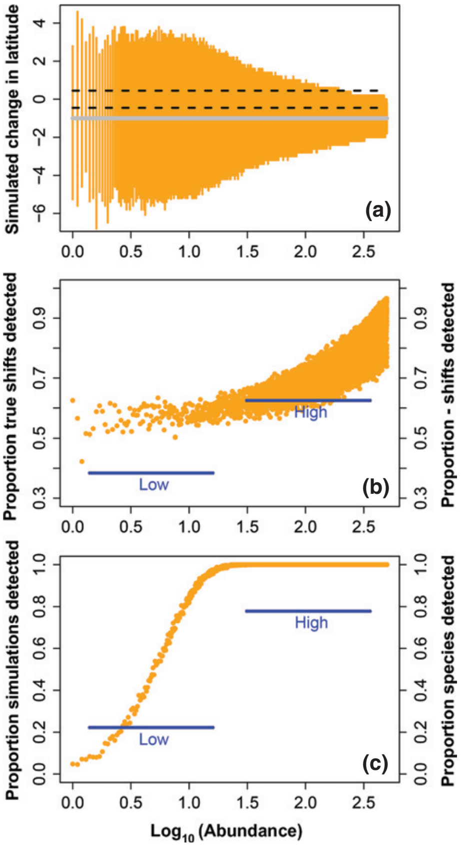

Figure 3 Relationship between abundance, the proportion of range shifts at trailing range edges and species detected. (a) 5th and 95th quartiles of simulated estimates of change in the location of trailing range edges (positive values are equatorward extensions, negative values are poleward contractions) for mean peak abundances of 1 to 500 at 0.1 steps. A range shift of 1° of latitude was modelled (grey line), and sampling effort was varied among simulations with a historical sampling bias in effort. (b) The proportion of simulated trailing range edge shifts (orange, left *y*-axis) showing a polewards contraction of 0.45° (equivalent to values $< -0.45°$ as shown by the bottom dotted black line in panel a) versus abundance. The proportion of macroalgal species (blue, right *y*-axis) scored as having low or high abundance showing trailing range edge contractions with absolute values of $> 0.45°$. (c) Proportion of 1000 simulations with observations in both sampling intervals (orange, left *y*-axis) and proportion of species sampled in herbarium records (blue, right *y*-axis) versus abundance.

Fig. 4b and d). By contrast, for species observed more frequently, the range edge estimates were closer to the last observed location and confidence intervals were narrower, so that we could place higher confidence that a northward range shift occurred in *Upeneichthys vlamingii* (Fig. 4c).

Overall, the DTE model performed well for common species that occurred throughout Tasmania in predicting the range edge at a location close to the true geographical range limit (edge of the continent) with precise confidence limits (Fig. 4d). However, the geographical ranges of some species would likely extend further southwards than the edge of Tasmania if habitat was available.

DISCUSSION

While the issue of variability in detection among species is well known in ecological studies (e.g. Tanadini & Schmidt, 2011), we show here how abundance-related species detectability has the potential to confound our understanding of the true location of range edges. Failure to account for non-detection leads to inaccurate and inconsistent assessment of range shifts among species. Using simulations, we demonstrate how the probability of occupancy and variation in sampling effort influence accuracy when identifying range shifts, emphasizing the importance of long-term monitoring with even sampling effort through time. We further suggest the application of a distance-to-edge (DTE) model to estimate confidence in range edge locations when species distributional limits are of interest. Tools such as simulations and modelling confidence intervals will allow more realistic descriptions of range changes for individual species and entire communities, leading to better understanding of the environmental and ecological factors underpinning range shift dynamics.

Simulations of range edges provide a baseline expectation for the detection of range shifts in species found in varying abundance for comparison to field data and subsequently assist in interpretation of patterns. As expected, estimates of range shifts are more accurate for more abundant species (Shoo *et al.*, 2006; Tanadini & Schmidt, 2011).

While sampling issues associated with detectability (not just those related to abundance) are a well-known problem, the large spatial and temporal scale required to obtain accurate range edge locations presents particular challenges in marine systems (Monk, 2013). Simulations indicate that even under a best case scenario of high-resolution sampling, estimates of range edge locations are highly variable for less abundant species, patterns that are also observed in regional-scale empirical data sets. This issue is likely to be more acute for marine than terrestrial systems due to the sampling effort limitations imposed by logistics of collecting data underwater. Therefore, while solutions such as subsampling data to equalize sampling effort between time periods have been advised on the basis of terrestrial studies (Hill *et al.*, 2002; Hassall & Thompson, 2010), such solutions may be counterproductive for marine studies, where sample sizes can be much lower and maximizing information is important. Moreover, subsampling techniques only exacerbate the fact that rarer species are going undetected or, if measured, are more likely to have inaccurate estimates of change in comparison with more abundant species. These results suggest that the available baseline data in marine systems are insufficient to estimate range movements of rare and inconspicuous species due to quality (e.g. museum collections: Przeslawski *et al.*, 2012) or limited sampling resolution (e.g. spatial positioning of samples may not capture range changes).

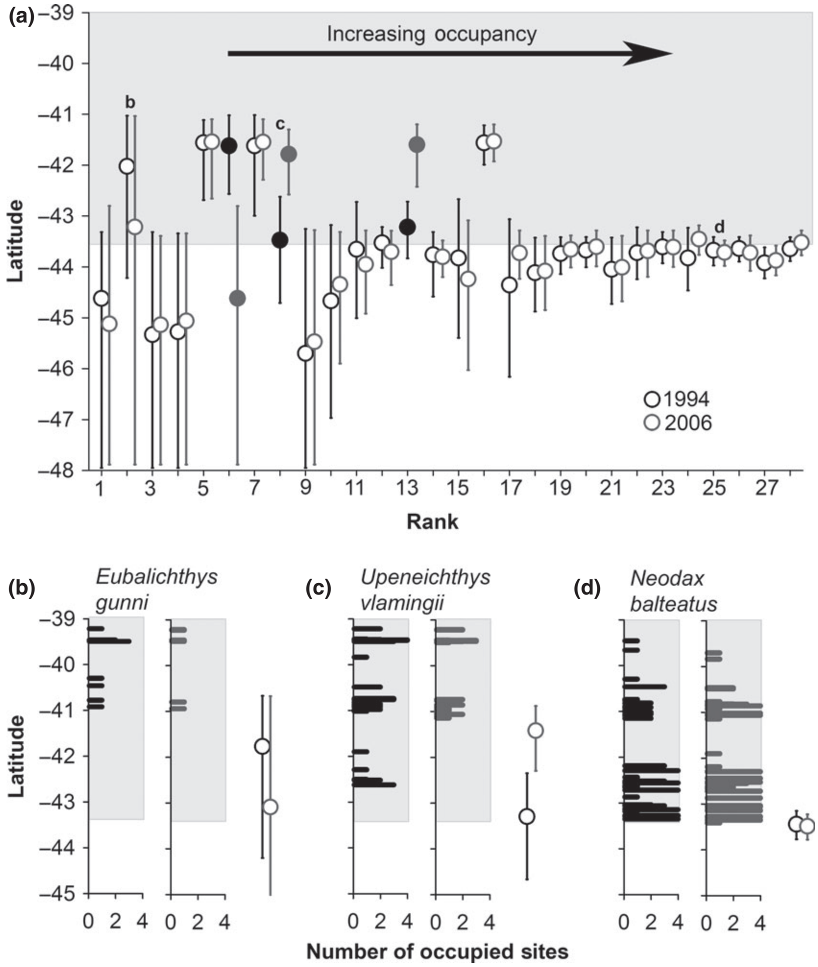

Figure 4 (a) Leading range edges (symbols) and 95% confidence intervals (bars) predicted with a distance-to-edge model for 28 fish species sampled in both 1994 and 2006. Species are ranked by increasing occupancy (i.e. the total number of sites occupied by a species in both time intervals). The majority of the 2006 range edge estimates fell within the 95% confidence interval (CI) of the 1994 estimate; however, in three species, range shifts were supported by non-overlapping confidence intervals (filled symbols). Letters indicate example species highlighted in panels b-d. The grey-shaded area represents the latitudes of southern Australia, and the edge of the grey area is the southern limit of shallow marine habitat on the continent. (b–d) The raw data (number of sites occupied for each 0.1 degrees of latitude) are displayed for comparison with the modelled leading range edge estimates and 95% CIs for 1994 and 2006 (offset to the right of each panel) for three example species.

Table 1 Generalized linear model results for relationships between measured change in the latitudinal location of the range edge in macroalgae versus abundance and sample size. Negative values indicate a declining slope

	Coefficient	Standard error	*t*-value	*P*-value
Intercept	−0.51	0.90	−0.57	0.57
Abundance category	−0.40	0.26	−1.51	0.11
Sample size	0.065	0.024	2.70	0.010
Log(effort ratio)	−1.61	0.69	−2.35	0.024

The signatures of sampling error related to species abundance and sampling effort, both of which influence species detectability, are therefore present in published data sets of marine range shifts. While our simulations support that the Tasmanian reef fish community displayed stable range edges (Stuart-Smith *et al.*, 2010) and that the southern Australian macroalgal community shifted polewards (given that historical sampling effort was greater in this study, Wernberg *et al.*, 2011), attributing confidence in the observed species-specific responses in range edges was not possible at the time of the original studies (nor was this the aim of these studies, e.g. Wernberg *et al.*, 2014).

Our DTE model now provides a tool to estimate confidence in species-specific responses that incorporates information on the number of sites where species were observed and the spatial distribution of occupied sites. For species with higher occupancy, the model estimated the range edge location (i.e. the edge of the Australian continent) with high precision. However, at lower occupancy and when the records of species were less even in space, precision in the range edge position was low, as expected. Thus, confidence intervals allowed us to consider which observed shifts most likely represented real change. We were therefore able to identify species in which the confidence intervals for the two sampled time periods did not overlap and thus support for a range shift with > 95% confidence. Overall, the range edge estimates between the two time periods were generally similar for most species, further supporting the original interpretation of stability in the range edges of reef fish communities. We therefore suggest that further evaluation of additional time-to-extinction models with data that have different sampling resolutions and spatial distribution patterns will be important. Including habitat variables (such as using species distribution modelling approaches) in conjunction with range edge estimates based on spatial occupancy patterns has the potential to build more accurate estimates of where species are located. In addition, trait-based time-to-detection models, such as described by Garrard *et al.* (2012), may be applicable to a spatial setting.

If we are to improve our ability to model and predict both current and future range extensions and contractions, variable detectability of species needs to be considered in analyses and monitoring efforts (Monk, 2013). Although quantifying confidence in the accuracy of range edge estimates is a step forward, the most fundamental issue is that baseline data are either lacking or were not collected at a spatial and temporal resolution designed for rarer species (Maxwell & Simon, 2005). Statistical tools may help to account for the patchy nature of present and historical data and inherent variation in species detectability, such as hidden-Markov or Bayesian hierarchical models (Wintle *et al.*, 2013;). However, monitoring programmes designed to detect future range shifts should be implemented now with the goal of detecting changes in species distributions for species with different detectability, in combination with robust quantitative approaches tailored for climate change ecology (Brown *et al.*, 2011). Well-designed surveys will provide baseline data for comparison to the present, facilitating systematic assessments, so that range shifts in rarer species are not going unnoticed.

CONCLUSION

The capacity to discern true range shifts from those that are simply due to sampling variability will advance our understanding of the mechanisms driving species redistribution. In particular, the lack of data in marine systems means that we need to cautiously apply sampling and statistical approaches developed in terrestrial systems where data sources tend to be richer. However, simulations with simple assumptions can assist in generating baseline expectations for the detection of range shifts and show where incorrect range shift estimates can be expected under different sampling scenarios. The application of tools to estimate confidence in observed shifts also has the potential to indicate range shifts that are most likely to represent real change. Such application will not correct for underestimates in the frequency of range shifts in rare and inconspicuous species due to missed observations; however, such approaches will assist in the appropriate interpretation of observed range shifts.

ACKNOWLEDGEMENTS

K. Dixon and G. Kendrick kindly estimated the abundance of southwest Australia algal species. A.E.B. was supported by the Australian National Network in Marine Science (a collaborative funding initiative between James Cook University, The University of Tasmania and The University of Western Australia) and an Australian Government Fisheries Research and Development Corporation and Department of Climate Change and Energy Efficiency grant. T.J.B. and M.M. were supported in part by the Australian Research Council (ARC) Centre of Excellence for Environmental Decisions, M.M. and T.W. by ARC Future Fellowships, G.J.E. and R.S-S. by an ARC Linkage Grant, and D.S. by a NERC Independent Research Fellowship. This research was also supported by the Marine Biodiversity Hub, a collaborative partnership supported through the Australian Government's National Environmental Research Program.

REFERENCES

Angert, A.L., Crozier, L.G., Rissler, L.J., Gilman, S.E., Tewksbury, J.J. & Chunco, A.J. (2011) Do species' traits predict recent shifts at expanding range edges? *Ecology Letters*, **14**, 677–689.

Bates, A.E., Pecl, G.T., Frusher, S., Hobday, A.J., Wernberg, T., Smale, D.A., Sunday, J.M., Hill, N., Dulvy, N.K., Colwell, R.K., Holbrook, N.J., Fulton, E.A., Slawinski, D., Feng, M., Edgar, G.J., Radford, B.T., Thompson, P.A. & Watson, R.A. (2014) Defining and observing stages of climate-mediated range shifts in marine systems. *Global Environmental Change*, **26**, 27–38.

Blanchard, J.L., Maxwell, D.L. & Jennings, S. (2008) Power of monitoring surveys to detect abundance trends in depleted fish populations: the effects of density-dependent habitat use, patchiness, and climate change. *ICES Journal of Marine Science*, **65**, 111–120.

Brown, C.J., Schoeman, D.S., Sydeman, W.J., Brander, K., Buckley, L.B., Burrows, M.T., Duarte, C.M., Moore, P.J., Pandolfi, J.M., Poloczanksa, E.S., Venables, W.N. & Richardson, A.J. (2011) Quantitative approaches in climate change ecology. *Global Change Biology*, **17**, 3697–3713.

Chen, I.C., Hill, J.K., Ohlemüller, R., Roy, D.B. & Thomas, C.D. (2011) Rapid range shifts of species associated with high levels of climate warming. *Science*, **333**, 1024–1026.

Clements, C. (2012) sExtinct: Calculates the historic date of extinction give a series of sighting events. R package version 1.0.0 (Available at: http://CRAN.R-project.org/package=sExtinct).

Dorazio, R. & Royle, J.A. (2005) Estimating size and composition of biological communities by modelling the occurrence of species. *Journal of the American Statistical Association*, **100**, 389–398.

Garrard, G.E., McCarthy, M.A., Williams, N.S.G., Bekessy, S.A. & Wintle, B.A. (2012) A general model of detectability using species traits. *Methods in Ecology and Evolution*, **4**, 45–52.

Hassall, C. & Thompson, D.J. (2010) Accounting for recorder effort in the detection of range shifts from historical data. *Methods in Ecology and Evolution*, **1**, 343–350.

Hill, J.K., Thomas, C.D., Fox, R., Telfer, M.G., Willis, S.G., Asher, J. & Huntley, B. (2002) Responses of butterflies to twentieth century climate warming: implications for future ranges. *Methods*, **269**, 2163–2171.

Hickling, R., Roy, D.B., Hill, J.K. & Thomas, C.D. (2005) A northward shift of range margins in British Odonata. *Global Change Biol.*, **11**, 502–6.

Isaac, N.J.B., Cruickshanks, K.L., Weddle, A.M., Rowcliffe, J.M., Brereton, T.M., Dennis, R.L.H., Shuker, D.M. & Thomas, C.D. (2011) Distance sampling and the challenge of monitoring butterfly populations. *Methods in Ecology and Evolution*, **2**, 585–594.

Jones, S.J., Lima, F.P. & Wethey, D.S. (2010) Rising environmental temperatures and biogeography: poleward range contraction of the blue mussel, *Mytilus edulis* L., in the western Atlantic. *Journal of Biogeography*, **37**, 2243–2259.

MacNeill, M.A., Tyler, E.H.M., Fonnesbeck, C.J., Rushton, S.P., Polunin, N.V.C. & Convoy, N.J. (2008) Accounting for detectability in reef-fish biodiversity estimates. *Marine Ecology Progress Series*, **367**, 249–260.

Maxwell, D. & Simon, J. (2005) Power of monitoring programmes to detect decline and recovery of rare and vulnerable fish. *Journal of Applied Ecology*, **42**, 25–37.

McCarthy, M.A., Moore, J.L., Morris, W.K., Parris, K.M., Garrard, G.E., Vesk, P.A., Rumpff, L., Giljohann, K.M., Camac, J.S., Bau, S.S., Friend, T., Harrison, B. & Yue, B. (2013) The influence of abundance on detectability. *Oikos*, **122**, 717–726.

McGill, B. & Collins, C. (2003) A unified theory for macroecology based on spatial patterns of abundance. *Evolutionary Ecology Research*, **5**, 469–492.

Monk, J. (2013) How long should we ignore imperfect detection of species in the marine environment when modelling their distribution? *Fish Fisher*, **15**, 352–358. doi:10.1111/faf.12039.

Parmesan, C., Ryrholm, N., Stefanescu, C. *et al.* (1999) Poleward shifts in geographical ranges of butterfly species associated with regional warming. *Nature*, **399**, 579–583.

Pinsky, M.L., Worm, B., Fogarty, M.J., Sarmiento, J.L. & Levin, S.A. (2013) Marine taxa track local climate velocities. *Science*, **341**, 1239–1242.

Poloczanska, E.S., Brown, C.J., Sydeman, W.J., Kiessling, W., Schoeman, D.S., Moore, P.J., Brander, K., Bruno, J.F., Buckley, L.B., Burrows, M.T., Duarte, C.M., Halpern, B.S., Holding, J., Kappel, C.V., O'Connor, M.I., Pandolfi, J.M., Parmesan, C., Schwing, F., Thompson, S.A. & Richardson, A.J. (2013) Global imprint of climate change on marine life. *Nature Climate Change*, **3**, 919–925.

Przeslawski, R., Falkner, I., Ashcroft, M.B. & Hutchings, P. (2012) Using rigorous selection criteria to investigate marine range shifts. *Estuarine, Coastal and Shelf Science*, **113**, 205–212.

R Development Core Team (2011) *R: A language and environment for statistical computing*. R Foundation for Statistical Computing, Vienna, Australia.

Rivadeneira, M.M., Hunt, K. & Roy, K. (2009) The use of sighting records to infer species extinctions: an evaluation of different methods. *Ecology*, **90**, 1291–1300.

Roberts, D.L. & Solow, A.R. (2003) Flightless birds: when did the dodo become extinct? *Nature*, **426**, 245–245.

Sexton, J.P., McIntyre, P.J., Angert, A.L. & Rice, K.J. (2009) Evolution and ecology of species range limits. *Annual Reviews in Ecology, Evolution and Systematics*, **40**, 415–436.

Shoo, L.P., Williams, S.E. & Hero, J.M. (2006) Detecting climate change induced range shifts: where and how should we be looking? *Austral Ecology*, **31**, 22–29.

Solow, A.R. (2005) Inferring extinction from a sighting record. *Mathematical Biosciences*, **195**, 47–55.

Stuart-Smith, R.D., Barrett, N.S., Stevenson, D.G. & Edgar, G.J. (2010) Stability in temperate reef communities over a decadal time scale despite concurrent ocean warming. *Global Change Biology*, **16**, 122–134.

Sunday, J.M., Bates, A.E. & Dulvy, N.K. (2012) Thermal tolerance and the global redistribution of animals. *Nature Climate Change*, **2**, 686–690.

Tanadini, L.G. & Schmidt, B.R. (2011) Population size influences amphibian detection probability: implications for biodiversity monitoring programs. *PLOS ONE*, **6**, e28244.

Thomas, C.D. & Lennon, J.J. (1999) Birds extend their ranges northward. *Nature*, **399**, 213.

Tingley, M.W. & Beissinger, S.R. (2009) Detecting range shifts from historical species occurrences: new perspectives on old data. *Trends in Ecology and Evolution*, **24**, 625–633.

Wernberg, T., Russell, B.D., Thomsen, M.S., Gurgel, C.F.D., Bradshaw, C.J.A., Poloczanska, E.S. & Connell, S.D. (2011) Seaweed communities in retreat from ocean warming. *Current Biology*, **21**, 1828–1832.

Wernberg, T., Russell, B.D., Bradshaw, C.J.A., Gurgel, C.F.D., Thomsen, M.S., Poloczanska, E.S. & Connell, S.D. (2014) Misconceptions about analyses of Australian seaweed collections. *Phycologia*, **53**, 215–220.

Wintle, B.A., Walshe, T.V., Parris, K.M. & McCarthy, M.A. (2013) Making sense of species occupancy data when detection is imperfect. *Diversity and Distributions*, **18**, 417–424.

SUPPORTING INFORMATION

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

Appendix S1 R code for simulations.

Appendix S2 Qualitative abundance of the macroalgal species estimated by experts.

BIOSKETCH

Amanda Bates is a Lecturer in Macroecology at the National Oceanography Centre Southampton, University of Southampton. Her research links physiological thresholds of organisms to the environment they experience to quantify changes in animal behaviour, the outcome of species interactions and distribution patterns.

Author contributions: A.E.B., T.J.B and M.M. conceived the ideas with input from all authors. A.E.B. and T.J.B. co-wrote the manuscript and conducted the analyses with advice from M.M. R.D.S-S., N.B., G.E. and T.W. provided raw data and assisted with interpreting these data. All authors contributed comments to draft versions of the manuscript.

Editor: David M. Richardson