

REVISED ASSESSMENT FOR OCEANIA HUMPBACKS: SURVEY DESIGN CONSIDERATIONS AND POWER ANALYSIS

Emma. L. Carroll¹, Lyndon Brooks², Dan Burns^{2,4}, Jennifer A. Jackson⁴, C. Scott Baker^{1,5}, Phillip Clapham⁶, Rochelle Constantine¹, Claire Garrigue⁷, Nan Hauser⁸, David Mattila⁹, Marc Oremus⁷, Michael Poole¹⁰, Jooke Robbins¹¹, Debbie Steel⁶ and Rachel M. Fewster¹²

¹*School of Biological Sciences, University of Auckland, Private Bag 92019, Auckland, New Zealand*

²*Marine Ecology Research Centre, Southern Cross University, PO Box 157 Lismore, New South Wales 2480 Australia*

³*Blue Planet Marine, PO Box 919 Jamison Centre, 2614, ACT Australia*

⁴*British Antarctic Survey, High Cross, Madingley Road, Cambridge CB30ET, UK*

⁵*Hatfield Marine Science Center, Oregon State University, 2030 SE Marine Science Drive, Newport, OR 97365 USA*

⁶*National Marine Mammal Laboratory, Alaska Fisheries Science Center, National Marine Fisheries Service, NOAA, 7600 Sand Point Way NE, Seattle, WA 98115*

⁷*Opération Cétacés, BP 12827, 98802 Nouméa, Nouvelle-Calédonie*

⁸*Cook Islands Whale Research, P.O. Box 3069, Avarua, Rarotonga, Cook Islands*

⁹*Hawaiian Islands Humpback Whale National Marine Sanctuary, Kihei, Hawaii 96753, USA*

¹⁰*Centre de Recherches Insulaires et Observatoire de l'Environnement, BP 1013, Moorea, Polynésie Française*

¹¹*Provincetown Center for Coastal Studies, Provincetown, Massachusetts 02657, USA*

¹²*Department of Statistics, University of Auckland, Private Bag 92019, Auckland, New Zealand*

ABSTRACT

The Oceania population of humpback whales, comprising sub-stocks E2, E3, and F, was heavily depleted by whaling operations in the 20th century and is now classed as Endangered by the IUCN. A mark-recapture study of individuals identified by genotypes and photo-identification found the Oceania population had a superpopulation size of 4,329 whales (95% CL 3345, 5315) from 1999 to 2005 and an annual population growth (λ) for this period of 1.03 (95% CL 0.90-1.18). This low and imprecise estimate of growth contrasts with the high rate of growth for east Australia (10.9% per annum, corresponding to $\lambda = 1.109$; 95% CL 10.5%-11.3%) but the wide confidence intervals do not allow a firm trend to be detected. At a recent meeting of the South Pacific Whale Research Consortium (SPWRC), it was recommended that a future assessment of the Oceania humpback stock through mark-recapture methodology should be able to meet 3 objectives: (1) determine population size with a coefficient of variation (CV) of <20% and (2) detect if λ is significantly different from 1 (i.e. the population is increasing or decreasing) and (3) detect if λ is significantly different from that of east Australia. Here we make recommendations for the sampling design of future surveys and conduct a power analysis to see if the proposed surveys will meet the defined objectives. Simulations were undertaken for the combined regions of Oceania, in addition to individual wintering grounds of New Caledonia (NC: E2) and Tonga (TG: E3). The proposed survey designs include a capture probability of $p=0.10$ for new surveys and target the three core regions of New Caledonia, Tonga and French Polynesia. We recommend conducting surveys that span the wintering period to reduce heterogeneity in capture probability due to the difference in migratory timing between demographic classes. Under the simulated scenarios for Oceania, incorporating data from the previous genotype surveys from 1999-2005 with three new survey years will give sufficient power to meet objective 1, detect if the growth rate is significantly lower than that of east Australia if the true $\lambda \leq 1.05$ and detect if the growth rate is significantly >1 if the true $\lambda=1.05$. The simulations also suggest that the power to meet the objectives on a regional basis varies with the survey design and simulated scenario. However, in general, the biennial survey design was able to detect with $\geq 90\%$ power if the growth rate is significantly lower than that of east Australia if the true $\lambda \leq 1.03$ for both NC and TG. Therefore, it seems the proposed surveys would allow us to test whether population growth rates in these different regions are significantly different: this appears to be important given the recent rapid increase in the New Caledonian calving ground.

INTRODUCTION

The Oceania population of humpback whales was heavily depleted by whaling operations in the 20th century and is now classed as Endangered by the IUCN (Childerhouse *et al.* 2008). The superpopulation (N_S) estimate for Oceania from 1999 to 2005 was 4,329 (95% CL 3345, 5315) whales and the population growth for this period was 1.03 (95% CL 0.90-1.18) based on a genotype mark-recapture study (Constantine *et al.* 2012). The estimated rate of increase is lower than that of east Australia of 10.9% per annum, corresponding to $\lambda = 1.109$, (95%CL 10.5%-11.3%: Noad *et al.* 2011), although the wide confidence interval (CI) from Oceania includes the CI from east Australia. Given the apparent low rate of recovery, there is a strong need to estimate abundance and trend information for this population. Using data from several South Pacific breeding grounds, mark-recapture data provide one method of obtaining this information. At a recent meeting of the South Pacific Whale Research Consortium (SPWRC), it was recommended that a future assessment of the Oceania humpback stock through mark-recapture methodology should be able to meet 3 objectives: (1) determine population size with a coefficient of variation (CV) of <20% and (2) detect if the population growth rate (λ) is significantly different from 1 (i.e. the population is increasing or decreasing) and (3) detect if λ significantly different from that of east Australia. Here we discuss survey design considerations and conduct a power analysis to determine if our proposed sampling design is likely to meet the objectives, under various scenarios for the true demographics of the population.

SURVEY DESIGN CONSIDERATIONS

Heterogeneity in capture probability: Heterogeneity in capture probability is typically thought of as an intrinsic characteristic of an individual that will create a bias in abundance estimates if unaccounted for (Seber 1982). In order to estimate population growth rate, we will need to use an open mark-recapture model, as these account for removals from (deaths/emigration) and additions to (births/immigration) the population. However, despite some advances (Pledger 1998, 2000), open models cannot account for heterogeneity in capture probability, particularly when capture probability is low. Therefore, we need to account for heterogeneity by careful sampling design or with suitable covariates. Two major sources of heterogeneity between individuals are (1) timing of migratory arrivals on the Oceania wintering ground and (2) duration of stay on Oceania wintering grounds, both of which are associated with age/sex class and reproductive status.

Timing of migratory arrivals: Humpback whales show heterogeneity based on sex and demographic class in migratory behaviours, whereby different age classes of whales leave high latitude summer feeding areas and arrive at low latitude winter breeding areas at different times. This occurs in a generally set sequence: the first whales to leave Antarctica are the lactating females and yearlings, followed approximately 12 days later by immature whales of both sexes, with mature males, non-pregnant females and pregnant females leaving approximately 20, 23 and 31 days, respectively, after the start of the migration. In general, the earliest and last whales in each category travel approximately six weeks before and after the central peak for that category (Dawbin 1966). On the return journey south, newly pregnant or resting females and immature whales leave tropical waters first, followed 10 days later by mature males, and finally mothers with calves a further 6 days later (Dawbin 1966, 1997). This migratory behaviour creates heterogeneity in capture probability linked to sex and demographic class on the wintering grounds.

Some individuals also show consistent timing in their migration from year to year (Burns *et al.* 2012). Burns *et al.* (2012) found that 15 of 22 interseason resights of whales photographed during their southern migration off the eastern coast of Australia passed the sampling location within 7 days of their previous sighting day of the year. If this pattern holds true for Oceania breeding grounds then heterogeneity in capture probability would also be evident at the individual whale level, depending on timing of migration relative to survey effort.

In addition, there is evidence for heterogeneity in capture probability linked to female reproductive status. Across-season resightings suggest that females tend to arrive at and depart from the breeding grounds earlier in years when they do not have a calf than in years when they give birth (Burns *et al.* 2012; Craig *et al.* 2003; Glockner & Venus 1983). As humpback whales have a modal reproductive cycle of two or three years (although it can vary between one and five years), the migratory timing of individual females, and therefore availability for capture on wintering grounds, appears to vary across years depending on reproductive status (Baker *et al.* 1987; Craig *et al.* 2003).

Duration of stay: Recent evidence suggests that whales probably remain on winter breeding areas for around 3-5 weeks (Burns *et al.* In Review; Craig *et al.* 2001; Dulau-Drouot *et al.* 2012). Although Dawbin (1966) found no strong evidence that different age classes spend different lengths of time in tropical waters, more recent authors report that males have longer residency times in breeding areas than females (Burns *et al.* In Review; Clapham 2002; Craig *et al.* 2003; Craig *et al.* 2001; Dulau-Drouot *et al.* 2012). In particular, females are likely to return

quickly to the feeding grounds after becoming pregnant, while mature males remain on the breeding grounds longer in an attempt to increase mating opportunities (Burns *et al.* In Review; Clapham 2002; Craig *et al.* 2003; Craig *et al.* 2001; Dulau-Drouot *et al.* 2012). Therefore, males are likely to have a higher capture probability on wintering grounds. This is reflected in the male bias in genetic samples from the Oceania wintering grounds (Constantine *et al.* 2012), as well as in sighting histories of whales from Hawaii (Herman *et al.* 2011).

Recommendations to reduce heterogeneity: The variation in the timing of migratory arrivals and duration of stay between demographic classes underlies the variation in availability for capture and capture probabilities between the sexes and demographic classes. In order to account for this, the duration of sampling needs to be long enough to capture all classes within a season. The duration of sampling also needs to be consistent between seasons to provide the opportunity for recaptures. **Therefore we recommend sampling across the entire 10-12 week wintering period to decrease heterogeneity in capture probability and ensure all demographic classes are captured.**

Females are less available for capture on wintering grounds and show heterogeneity in capture probability linked to reproductive status. Identifying the sex of the whale will therefore be important to account for this heterogeneity. In addition, males have a higher sampling and capture probability on wintering grounds. **Therefore we recommend sampling in a way that allows for the identification of sex, implying a preference for biopsy sampling.** Mark-recapture modelling based on males only could also be more effective, due to their higher capture probability and lack of heterogeneity linked to reproductive status (although there is likely heterogeneity due to age or size). Recently, a model was developed to incorporate this type of heterogeneity among females into open population models (Carroll *et al.* In press).

Sampling areas: Previous surveys included three core regions (New Caledonia: NC; Tonga: TG; and French Polynesia: FP) with nearly annual surveys and a number of secondary regions surveyed once or twice (Constantine *et al.* 2012). There was little difference between the estimates of abundance for Oceania derived from all available regions and the three core regions. New data from the three core sites will be directly comparable with existing data and should allow us to investigate overall, and in some cases regional, λ with adequate precision (see below). This is particularly useful given the recent evidence of increases in New Caledonia but not other Oceania wintering grounds (Garrigue *et al.* 2012).

Additionally, if there are different capture probabilities across the three core regions, the estimate of abundance will be biased low. Fitting separate models to data from each region, or sampling regions in proportion to their abundance to yield approximately equal capture probabilities, should address this issue. **Therefore, we recommend focussing sampling on the three core sites and sampling these in proportion to their estimated abundance.**

Maximising capture probability: Efforts should be made to maximise capture probabilities, such as sampling for the entire wintering period. The previous study showed the capture probability varied between 0.02 and 0.10 per year (Constantine *et al.* 2012). We believe that by increasing the duration of survey and concentrating on three core sites a capture probability of 0.1 per year would be achievable (for estimated samples per year for this see results). **We recommend the new surveys aim for a capture probability of 0.1 by targeting sample sizes reflecting 10-15% of the regional population size** (see below).

Model: The λ -POPAN model provides estimates of N_S size (regional estimate), number of animals in the population each year (N_t), and population rate of change (Carroll *et al.* In press). The model is able to directly estimate λ , which is distinct from the standard POPAN model available in program MARK, as it does not estimate the probability of entry as free parameters. Instead, the probability of entry parameters are constrained to follow a trajectory determined by a single value of λ .

POWER ANALYSIS

We undertook simulations to determine how effective this sampling design would be at estimating abundance and detecting trends in the Oceania population. These simulations are based on the following assumptions about the survey design and population, as well as the standard assumptions for open capture-recapture models:

1. Correct mark-recapture identification of male humpback whales (i.e. genotypes) in Oceania.
2. The proposed surveys have a capture probability of $p=0.10$ for males. We believe this is achievable given previous estimates of capture probability and proposed sampling intensities.
3. A constant λ and Φ over the time period, where Φ denotes annual apparent survival.

Methods: We fitted the λ -POPAN model to the male genotype capture-history data for New Caledonia (1999-2005, NC), Tonga (2000-2003 & 2005, TG) and the synoptic Oceania region (1999-2005) from Constantine *et*

al. (2012). In cases where not all model parameters were estimable, or estimates at boundary values were returned, some parameters were constrained. Apparent survival could be fixed at 0.95 ($\Phi=0.95$; as found in Constantine *et al.* 2012) or years with similar sample sizes (taken as a proxy for survey effort) were constrained to have the same capture probability. We then used the best fitting λ -POPAN model(s) to simulate future scenarios assuming three new annual or biennial surveys will take place from 2014.

To determine the appropriate superpopulation size N_S for the simulated periods 1999-2016 or 1999-2018, we note that the true N_S depends upon the true values of λ and Φ as well as the time period examined. For this reason we selected a baseline absolute abundance for 1999 to be common to all simulation scenarios for each region. The estimated number of males in the population in 1999 was derived from the demographic parameters in the best-fitting model from the 1999-2005 data. This common N_{1999} value was used to produce scenario-specific N_S values for time periods 1999-2016 or 1999-2018, for $\lambda = 0.98, 1.03$ or 1.05 , and for Φ fixed at 0.95 . The capture probabilities from the best-fitting 1999-2005 model were used to simulate inclusion of the 1999-2005 data, and capture probabilities of $p=0.1$ were used for data from 2014 onwards.

Under each simulation scenario, 1000 replicate populations were simulated using the scenario values of N_S , λ (annual growth rate) and Φ (apparent survival). The mark-recapture study was simulated for each replicate with capture probability p for k survey years. Simulated capture histories were then fitted with λ -POPAN to estimate N_S , λ , Φ and p . For each scenario, the mean estimate of N_S and N_{2014} (number in the population in 2014) and mean estimated coefficient of variation (CV) of N_S and N_{2014} were calculated from the 1000 simulations.

To determine the power of each scenario to determine if λ was significantly different from 1, we calculated the percentage of simulations where the 95% CL of λ were entirely <1 and >1 , for simulations where $\lambda = 0.98$ and 1.03 or 1.05 , respectively. We also investigated the power of each scenario to detect if λ was significantly different from the east Australian growth rate of 10.9% per annum (95% CL 10.5%-11.3%), corresponding to $\lambda = 1.109$.

We simulated three 'regions': NC (Stock E2), TG (Stock E3) and Oceania (NC, TG and FP). FP (Stock F) was explored but the low recapture rate in existing data prevented a region-specific power analysis. Simulations were undertaken using custom R code (available from R. Fewster on request).

RESULTS

Oceania: The best fitting model for the male genotype dataset for Oceania was $\lambda(.)p(t)\Phi(0.95)$ (Supplementary Table 1). This model gave the following capture probabilities: $p_{1999}=0.02$, $p_{2000}=0.04$, $p_{2001}=0.12$, $p_{2002}=0.08$, $p_{2003}=0.03$, $p_{2004}=0.04$, $p_{2005}=0.09$. The simulations were run using these capture probabilities and $N_{1999}=1019$, and annual (2014:2016) and biennial surveys (2014, 2016, 2018) with constant $p=0.1$ for the new, proposed surveys. The model fitted to the simulated data allowed this capture probability to vary by capture occasion according to the pattern in the 1999-2005 data but had a constant capture probability for the proposed new surveys. Where models did not converge adequately we implemented model constraints, as described above.

The results of simulations are found in Table 1: scenarios O1-O3 simulated annual surveys from 2014-2016 while scenarios O4-O6 simulated three biennial surveys from 2014. When $\lambda = 1.05$, both annual and biennial survey designs have significant ($>95\%$) power to detect that $\lambda > 1$ and also that the true $\lambda < 1.10$. When $\lambda = 1.03$, annual and biennial surveys had 68% and 76% power, respectively, to detect if $\lambda > 1$ and both annual and biennial surveys have strong power (100%) to detect that the true λ is < 1.10 . At these higher rates of increase, the survey design appears sufficient both to detect that the population is increasing and also to detect that the rate of increase is not as high as east Australia.

Under both survey configurations (annual and biennial new surveys), simulations suggest there is significant power to detect that λ is < 1.05 , when the true $\lambda = 0.98$. However, when the actual growth rate is 0.98, biennial surveys provide a 5-10% increase in power to detect if the actual growth rate is lower than all values of 1 and above. The simulations suggest both proposed annual and biennial survey designs will allow us to detect if λ is significantly different from east Australia when the actual growth rate is low.

Both N_S and N_{2014} were estimated with a CV $< 20\%$ in all simulated scenarios.

New Caledonia: There were three models that fit the NC data within ΔAIC of 3 (see Supplementary Table 2). Therefore we used two N_{1999} values: a high case of 326 whales and a conservative case of 210 whales. We used the capture probabilities from the model that produced a value of λ (1.03, 95% CL 0.96, 1.11) and Φ (fixed at 0.95) most similar to the overall findings for Oceania in Constantine *et al.* (2012). This model had the following capture probabilities: $p_{1999}=0.05$, $p_{2000}=0.09$, $p_{2001}=0.13$, $p_{2002}=0.05$, $p_{2003}=0.13$, $p_{2004}=0.05$, $p_{2005}=0.13$. The conservative case (N1-N6) scenarios and high case (N7-N12) were both simulated for annual (2014:2016) and biennial surveys (2014, 2016, 2018) with constant $p=0.1$.

The simulation results are found in Table 2. When the true $\lambda = 1.05$, simulations suggest that the proposed survey designs have ~70% power to detect if the true $\lambda > 1$ under the conservative case, which increases to ~90% power under the high case. When the true $\lambda = 1.05$, annual surveys have 66% and 83% power to detect if $\lambda < 1.10$, under the conservative and high cases, respectively. This power is increased 5-10% when the new surveys are biennial rather than annual.

When the true $\lambda = 1.03$, there is weak power (<50%) to detect that $\lambda > 1$ under any scenario. However, when the true $\lambda = 1.03$ there is moderate power to detect if $\lambda < 1.10$ under the conservative case (87-90%) and high power to detect if $\lambda < 1.10$ under the high case (>97% power). This is also true when the true $\lambda = 0.98$, although there is more power to detect if $\lambda < 1.10$ under all scenarios (99-100%). Therefore, at lower rates of growth, the proposed survey design is sufficient to detect that the population growth rate of New Caledonia is different from that of east Australia. When the true $\lambda = 1.05$, power to detect if λ is significantly >1 or significantly different from east Australia is enhanced by conducting biennial surveys from 2014.

Under all simulated scenarios, N_S is estimated with a CV<20%. However, a CV<20% for N_{2014} is only obtained for the NC dataset under the high case simulations. Otherwise, the CV ranges from 20-30% for the conservative case.

Tonga: The historical data used for Tonga was limited to 2000:2003 and 2005 as there were no data for Tonga from 2004 and the sample size for Tonga for 1999 was <10. There were three models that fit the Tongan data within $3\Delta AIC$ value (see Supplementary Table 3). Therefore we used two N_{1999} values: a high case of 1397 whales and a conservative case of 639 whales. We note that the high case is larger than the estimate used for Oceania. This high variability in N_S is due to the large impact that small variations in very low capture probabilities have on the overall estimate of population size. Again, the sparseness and short time frame of the data meant that we used model constraints in some cases. We used the capture probabilities from the best-fitting model: $p_{2000}=0.02$, $p_{2001}=0.04$, $p_{2002}=0.04$, $p_{2003}=0.04$, $p_{2005}=0.02$. The conservative case (T1-T6) scenarios and high case (T7-T12) were both simulated for proposed annual (2014:2016) and biennial surveys (2014, 2016, 2018) with constant $p=0.1$.

The simulation results are found in Table 3. When the true $\lambda = 1.05$, simulations indicate that the biennial survey design has ~70% power to detect if the true $\lambda > 1$ under the conservative case, which increases to ~90% power under the high case. In contrast, the annual survey design has low power (<50%) under the conservative case and moderate power (79%) under the high case to detect if the true $\lambda > 1$. When the true $\lambda = 1.05$, the high case has strong (>95%) power to detect if $\lambda < 1.10$ under the biennial survey design, but not the annual survey design (power decreases to 71%).

When the true $\lambda = 1.03$, there is generally weak power (<50%) to detect that $\lambda > 1$ under any scenario. However, when the true $\lambda = 1.03$ or 0.98 there is strong power (>95%) to detect if $\lambda < 1.10$ under the high case scenarios and biennial surveys under the conservative case.

The proposed survey design is sufficient to detect that the population growth rate of Tonga is different from that of east Australia when the true growth rate is ≤ 1.03 , except when annual surveys are conducted under the conservative case (power decreases to 74%). Biennial surveys will also increase the power to detect if $\lambda > 1$ or if λ is significantly different from that of east Australia when the true $\lambda = 1.05$.

Under all simulated scenarios, N_S is estimated with a CV<20%. However, a CV>20% for N_{2014} was obtained when $\lambda=0.98$ for the conservative case (both annual and biennial surveys) and when $\lambda=1.03$ (biennial surveys only). We caution that these results are only applicable to the simulated scenarios. Due to the low recapture rate and high variance in the models fitted to the Tongan dataset, it could be that the simulated scenario is not an accurate representation of the population.

CONCLUSIONS

Under the simulated scenarios for Oceania, incorporating data from the previous genotype surveys from 1999-2005 with three new survey years will give sufficient power to (1) detect if the growth rate is significantly lower than that of east Australia if the true $\lambda \leq 1.05$ and (2) detect if the growth rate is significantly >1 if the true $\lambda=1.05$. The simulations also suggest that the power to meet the objectives on a regional basis varies with the survey design and simulated scenario. However, in general, the biennial survey design was able to detect with $\geq 90\%$ power if the growth rate is significantly lower than that of east Australia if the true $\lambda \leq 1.03$ for both NC and TG. If the true $\lambda=1.05$, biennial surveys will modestly increase the power of survey design to meet objectives 2 and 3. Therefore, it seems the proposed surveys would allow us to test whether population growth rates in these different regions are significantly different: this appears to be important given the recent rapid increase in the New Caledonian calving ground (Garrigue *et al.* 2012). Restricting analyses to only the proposed three new annual or biennial surveys, it was found that the precision around λ was very poor (data not shown). Inclusion of the full time series is therefore necessary to obtain the required precision around λ .

The modest (5-10%) increase in power to detect whether the true λ is different from that of east Australia was found when the proposed surveys are biennial rather than annual in the Oceania, TG and NC high case simulations. However, the impact of biennial surveys on the capture probability of females was not explored, as we only simulated the male half of the population. It could be that a cohort of females that calves every two years will be captured during biennial surveys, and the impact of the reproductive cycle on female capture probability would need to be considered in this context.

Under all scenarios investigated for Oceania, the desired precision on both N_S and N_{2014} were obtained. However, while the simulations indicated the target of $CV < 20\%$ was also met for N_S for both TG and NC, the same could not be said for precision around N_{2014} . The conservative cases, simulated with low rates of growth, were most likely to have decreased precision in estimates of abundance. However, the CV value never exceeded 30% and there are more existing data available that could be incorporated into the future assessment to increase the precision of estimated abundance.

The sample sizes for each of the simulated scenarios are presented in Tables 1-3. These numbers represent unique males identified in each year, and need to be adjusted for females (up to 50%) and within-year recaptures (typically 10-20%). Even with these adjustments, these sample sizes are reasonable given that the proposed survey design includes 10-12 week concurrent surveys in New Caledonia, Tonga and French Polynesia and capture probabilities obtained from previous surveys. If New Caledonia is growing at a rate comparable with east Australia (10% per annum), then the sample sizes increase to 80-130 unique males per year under the conservative case and 130-200 unique males per year under the high case.

Other uses of the data: We assessed the power of a mark-recapture study based on males identified through genotypes to meet the objectives put forward by the SPWRC on a new assessment. However, as part of this and other ongoing research, data on female humpback whales will also be collected and we can extract information on the relative capture probabilities of males and females (with and without calf). While we did not simulate and model this explicitly, models that incorporate heterogeneity in capture probability would be able to use these data to model abundance and trends from the female dataset and could provide information on calving intervals and other reproductive parameters (e.g. Carroll *et al.* In press; Cooke *et al.* 2003).

Caveats: These simulations are based on the standard assumptions of capture-recapture models and the assumptions laid out in the Methods section. Furthermore, the Oceania level simulations did not attempt to model the process as a multi-site mark-recapture study, rather, the proportional distribution of survey effort should account for differences between sites.

We did not attempt to account for movements between east Australia and Oceania. There is evidence for low-level movements between New Caledonia and east Australia based on both photo-identification records (Garrigue *et al.* 2011) and genotype records. However, the absolute number of matches between Oceania and east Australia was five times lower than the number of matches within Oceania (Garrigue *et al.* 2011). In addition, movements between individual breeding grounds are much less common than recaptures within Oceania breeding grounds (Garrigue *et al.* 2007). These findings are consistent with significant differences in mtDNA haplotype frequencies between the breeding grounds (Olavarria *et al.* 2007) and led to the Oceania stock structure hypothesis wherein each breeding ground is considered a sub-stock (E1, E2, E3, F).

However, the recent rapid increase in abundance in New Caledonia suggests this situation may not be static, at least in part of Oceania (Garrigue *et al.* 2012). In addition, a multistate movement model, based on microsatellite genotype data collected from 1999 to 2004, suggests that movements between east Australia and New Caledonia are not significantly different from movements between New Caledonia and Tonga. This indicates that population connectivity levels between east Australia, New Caledonia and Tonga are of similar but low magnitude, despite a strong difference in population trend between the east Australia and Oceania regions (Jackson *et al.* 2012).

We also assumed a constant Φ and λ for the simulations. As noted above, the recent rapid increase in abundance in New Caledonia indicates that a constant λ is likely to be unrealistic in this breeding ground. Furthermore, emigration would also invalidate the assumption that apparent survival or Φ is constant. However, making these assumptions will allow us to test whether there is a difference in λ between Tonga and New Caledonia. Additionally, other models can be used in future: for example, the standard POPAN model that models probability of entry as a time-varying, independent parameter might be more appropriate for New Caledonia.

Table 1: Parameters used to simulate a mark-recapture study of Oceania humpback whales and the results from 1000 simulations of each scenario. The parameters for each simulated scenario O1-O6 are listed: superpopulation size (N_S), population growth rate (λ), apparent survival (Φ). The capture probabilities for each survey year are as follows: p1999=0.02, p2000=0.04, p2001=0.12, p2002=0.08, p2003=0.03, p2004=0.04, p2005=0.09, and p=0.1 for all new surveys. Survey years indicate the whether the proposed new set of surveys is annual (2004:2016) or biennial (2014, 2016, 2018) and under the simulated scenario, the unique male sample size for each new survey year is shown for p=0.1. Results from simulations show the mean estimate and mean CV for N_S and N_{2014} from the 1000 simulations. The λ : power row also shows the proportion of simulations with 95% CI that were entirely <1 or >1, depending on the scenario. Finally, for each simulation, the table shows the percentage of simulations that had confidence intervals (CIs) that did not overlap with each λ value listed on the left-hand column (greater or less than the given λ). For example, 98% of simulations of scenario O1 had CIs below or above 1.05.

Oceania						
Simulated Scenario	O1	O2	O3	O4	O5	O6
Survey years	1999:2005 2014:2016	1999:2005 2014:2016	1999:2005 2014:2016	1999:2005 2014,2016, 2018	1999:2005 2014,2016, 2018	1999:2005 2014,2016, 2018
N_S	1436	2793	3652	1506	3066	4130
λ	0.98	1.03	1.05	0.98	1.03	1.05
Φ	0.95	0.95	0.95	0.95	0.95	0.95
Year: sample size (unique males)	2014: 75 2015: 74 2016: 72	2014: 159 2015: 164 2016: 168	2014: 212 2015: 222 2016: 236	2014: 75 2016: 72 2018: 69	2014: 159 2016: 169 2018: 179	2014: 212 2016: 236 2018: 258
Results from Simulations						
Mean N_S	1446	2816	3685	1509	3091	4166
N_S CV	0.08	0.07	0.07	0.08	0.07	0.07
Mean N_{2014}	759	1613	2151	770	1612	2152
N_{2014} CV	0.18	0.14	0.13	0.19	0.14	0.12
λ : power	<1: 22%	>1: 68%	>1: 98%	<1: 28%	>1: 76%	>1: 100%
λ	Overall Power					
0.99	10	90	100	10	94	100
1.00	22	68	98	28	76	100
1.01	44	34	91	52	41	96
1.02	67	10	68	75	15	81
1.03	85	5	38	89	5	48
1.04	94	11	12	97	14	15
1.05	98	31	5	99	40	6
1.06	99	61	14	100	71	16
1.07	100	86	38	100	90	46
1.08	100	96	66	100	98	78
1.09	100	99	90	100	100	95
1.10	100	100	98	100	100	99
1.11	100	100	100	100	100	100

Table 2: Parameters used to simulate a mark-recapture study of the New Caledonian humpback whale wintering ground and the results from 1000 simulations of each scenario. The parameters for each simulated scenario N1-N12 are listed: superpopulation size (N_S), population growth rate (λ), apparent survival (Φ). The capture probabilities for each survey year are as follows: p1999=0.05, p2000=0.09, p2001=0.13, p2002=0.05, p2003=0.13, p2004=0.05, p2005=0.13, and p=0.1 for all new surveys. Survey years indicate the whether the proposed new set of surveys is annual (2004:2016) or biennial (2014, 2016, 2018) and under the simulated scenario, the unique male sample size for each new survey year is shown. The results show the mean estimate and mean CV for N_S and N_{2014} from 1000 simulations. The λ : power row also shows the proportion of simulations that were <1 or >1, depending on the scenario. Finally, for each simulation, the table shows the percentage of simulations that had confidence intervals (CIs) that did not overlap with the lambda value listed on the left-hand column (greater or less than the given λ). For example, 99% of simulations of scenario N1 had CIs below 1.10. Scenarios N1 to N6 (conservative case) had a baseline of 210 whales in 1999, whereas scenarios N7 to N12 (high case) had a baseline of 326 whales in 1999.

New Caledonia		Conservative Case					High Case					
	N1	N2	N3	N4	N5	N6	N7	N8	N9	N10	N11	N12
Survey years	1999:2005 2014:2016	1999:2005 2014:2016	1999:2005 2014:2016	1999:2005 2014,2016, 2018	1999:2005 2014,2016, 2018	1999:2005 2014,2016, 2018	1999:2005 2014:2016	1999:2005 2014:2016	1999:2005 2014:2016	1999:2005 2014,2016, 2018	1999:2005 2014,2016, 2018	1999:2005 2014,2016, 2018
N_S	302	576	752	310	632	851	468	894	1168	482	981	1322
lambda	0.98	1.03	1.05	0.98	1.03	1.05	0.98	1.03	1.05	0.98	1.03	1.05
Φ	0.95	0.95	0.95	0.95	0.95	0.95	0.95	0.95	0.95	0.95	0.95	0.95
Year: Sample size	2104 to 2016: 15 each	2014: 33 2015: 34 2016: 35	2014:44 2015:46 2016:48	2014: 16 2016: 15 2018: 14	2014: 33 2016: 35 2018: 37	2014: 44 2015: 46 2018: 53	2014: 24 2015: 24 2016: 23	2014: 51, 2015: 52 2016: 54	2014: 68 2015: 71 2016: 75	2014: 24 2016: 23 2018: 22	2014: 51 2016: 54 2018: 57	2014: 68 2016: 75 2018: 82
Summary statistics from Simulations												
Mean N_S	305	591	767	313	650	874	469	905	1188	485	997	1342
N_S CV	0.14	0.13	0.13	0.14	0.14	0.14	0.11	0.10	0.10	0.12	0.11	0.11
Mean N_{2014}	159	340	451	161	339	449	246	519	701	247	523	697
N_{2014} CV	0.30	0.26	0.23	0.31	0.25	0.23	0.25	0.21	0.20	0.26	0.21	0.19
λ : power	<1: 14%	>1:26%	>1: 72%	<1: 14%	>1: 26%	>1: 78%	<1:15%	>1: 37%	>1: 92%	<1: 20%	>1: 43%	>1: 95%
λ	Overall Power											
1.00	14	26	72	14	26	78	15	37	92	20	43	95
1.05	74	18	4	75	20	4	88	23	4	92	25	4
1.10	99	87	66	99	90	74	100	97	83	100	98	88

Table 3: Parameters used to simulate a mark-recapture study of the Tongan humpback whale wintering ground and the results from 1000 simulations of each scenario. The parameters for each simulated scenario T1-T12 are listed: superpopulation size (N_S), population growth rate (λ), apparent survival (Φ). The capture probabilities for each survey year are as follows: p2000=0.02, p2001=0.04, p2002=0.04, p2003=0.04, p2005=0.02, and p=0.1 for all new surveys. Survey years indicate the whether the proposed new set of surveys is annual (2004:2016) or biennial (2014, 2016, 2018). The results show the unique male sample size for each new survey year, and the mean N_S from 1000 simulations and the coefficient of variation (CV) around the mean. The lambda: power row also shows the proportion of simulations that were <1 or >1, depending on the scenario. Finally, for each simulation, the table shows the percentage of simulations that had confidence intervals (CIs) that did not overlap with the lambda value listed on the left-hand column. For example, 98% of simulations of scenario T1 had CIs below 1.10.

Tonga	Conservative Case						High Case					
	T1	T2	T3	T4	T5	T6	T7	T8	T9	T10	T11	T12
Survey years	1999:2005 2014:2016	1999:2005 2014:2016	1999:2005 2014:2016	1999:2005 2014,2016, 2018	1999:2005 2014,2016 , 2018	1999:2005 2014,2016 , 2018	1999:2005 2014:2016	1999:2005 2014:2016	1999:2005 2014:2016	1999:2005 2014,2016 , 2018	1999:2005 2014,2016 , 2018	1999:2005 2014,2016, 2018
N_S	905	1669	2151	931	1836	2437	1976	3650	4702	2036	4014	5327
λ	0.98	1.03	1.05	0.98	1.03	1.05	0.98	1.03	1.05	0.98	1.03	1.05
Φ	0.95	0.95	0.95	0.95	0.95	0.95	0.95	0.95	0.95	0.95	0.95	0.95
Year: sample size (unique males)	2014:48 2015:47 2016:46	2014:97 2015:100 2016:103	2014:127 2015:133 2016:140	2014:48 2016:46 2018:44	2014: 97 2016:103 2018:109	2014:127 2016:140 2018:154	2014:105 2015:103 2016:101	2014:211 2015:218 2016:225	2014:277 2015:290 2016:305	2014:105 2016:101 2018:97	2014:211 2016:224 2018:238	2014:277 2016:305 2018:336
Simulation Results												
Mean N_S	931	1712	2205	950	1883	2468	2014	3710	4759	2059	4053	5376
N_S CV	0.18	0.14	0.13	0.18	0.14	0.12	0.13	0.09	0.08	0.12	0.09	0.08
Mean N_{2014}	477	963	1271	462	979	1288	1028	2107	2765	1023	2127	2771
N_{2014} CV	0.24	0.19	0.18	0.27	0.22	0.20	0.17	0.14	0.12	0.18	0.15	0.14
λ : power	<1: 6%	>1: 21%	>1: 47%	<1: 10%	>1: 29%	>1: 73%	<1: 12%	>1: 38%	>1: 79%	<1:16%	>1: 56%	>1:96%
λ	Overall Power											
1.00	5	21	47	10	29	73	12	38	79	16	56	96
1.05	59	8	5	77	16	4	93	20	4	99	30	4
1.10	99	74	47	100	92	71	100	98	82	100	100	98

Literature cited

- Baker CS, Perry A, Herman L (1987) Reproductive histories of female humpback whales *Megaptera novaeangliae* in the North Pacific. *Marine Ecology Progress Series* **41**, 103-114.
- Burns D, Brooks L, Clapham P, Harrison P (2012) Between-year synchrony in migratory timing of individual humpback whales, *Megaptera novaeangliae*. *Marine Mammal Science* **29**, 228-235.
- Burns D, Brooks L, Harrison P, *et al.* (In Review) Migratory movements of individual humpback whales photographed off the eastern coast of Australia. *Marine Mammal Science*.
- Carroll EL, Childerhouse S, Fewster R, *et al.* (In press) Accounting for female reproductive cycles in a superpopulation capture recapture framework. *Ecological Applications* doi.org/10.1890/12-1657.1.
- Childerhouse S, Jackson JA, Baker CS, *et al.* (2008) IUCN Proposal for separate listing for Oceania sub-population of humpbacks (ed. Evaluators Rtl), p. 10 pp.
- Clapham P (2002) Humpback whale *Megaptera novaeangliae*. In: *Encyclopedia of Marine Mammals* (eds. Perrin WF, Wursig B, Thewissen JGM), pp. 589-592. Academic Press, San Diego.
- Constantine R, Jackson JA, Steel D, *et al.* (2012) Abundance of humpback whales in Oceania using photo-identification and microsatellite genotyping. *Marine Ecology Progress Series* **453**, 249-261.
- Cooke J, Rowntree V, Payne R (2003) Analysis of inter-annual variation in reproductive success of South Atlantic right whales (*Eubalaena australis*) from photo-identification of calving females observed off Peninsula Valdés. *Unpublished report (SC/55/O23) presented to the Scientific Committee of the International Whaling Commission, Cambridge, UK*.
- Craig A, Herman L, Gabriele C, Pack A (2003) Migratory timing of humpback whales (*Megaptera novaeangliae*) in central North Pacific varies with sex, age and reproductive status. *Behavior* **140**, 981-1001.
- Craig A, Herman L, Pack A (2001) Estimating residence times of humpback whales in Hawaii. *Report prepared for the Hawaiian Islands Humpback Whale National Marine Sanctuary, Office of National Marine Sanctuaries, National Oceanic and Atmospheric Administration, U.S. Department of Commerce, and the Department of Land and Natural Resources, State of Hawaii*. .
- Dawbin W (1966) The seasonal migratory cycle of humpback whales. In: *Whales, dolphins and porpoises* (ed. Norris K). University of California Press, Berkeley.
- Dawbin W (1997) Temporal segregation of humpback whales during migration in Southern Hemisphere waters. *Memoirs of the Queensland Museum* **42**, 105.
- Dulau-Drouot V, Fayan J, Mouysset L, Boucad V (2012) Occurrence and residency patterns of humpback whales off Reunion Island during 2004-2010. *Journal of Cetacean Research and Management* **12**, 255-263.
- Garrigue C, Albertson R, Jackson JA (2012) An anomalous increase in the New Caledonia humpback whales breeding sub-stock E2. *Unpublished report (SC/64/SH6) presented to the Scientific Committee of the International Whaling Commission, Cambridge, UK*.
- Garrigue C, Baker CS, Constantine R, *et al.* (2007) Interchange of humpback whales in Oceania (South Pacific) 1999 to 2004 (revised SC/A06/HW55). *Unpublished report (SC/59/SH14) presented to the Scientific Committee of the International Whaling Commission, Cambridge, UK*.
- Garrigue C, Franklin T, Constantine R, *et al.* (2011) First assessment of interchange of humpback whales between Oceania and the east coast of Australia. *Journal of Cetacean Research and Management Special Issue* **3**, 269-274.
- Glockner DA, Venus SC (1983) Identification, growth rate and behaviour of humpback whales (*Megaptera novaeangliae*) cows and calves in the waters off Maui, Hawaii, 1977-1979. In: *Communication and behaviour of whales* (ed. Payne R), pp. 223-258. Westview Press Inc., Boulder, CO, USA.
- Herman L, Pack A, Rose K, *et al.* (2011) Resightings of humpback whales in Hawaiian waters over spans of 10-32 years: site fidelity, sex ratios, calving rates, female demographics and the dynamics of social and behavioural roles of individuals. *Marine Mammal Science* **27**, 736-768.
- Jackson JA, Anderson M, Steel D, *et al.* (2012) Multistate measurements of interchange between east Australia and Oceania (IWC breeding sub-stocks E1, E2, E3 and F) between 1999 and 2004. *Unpublished report (SC/64/SH22) presented to the Scientific Committee of the International Whaling Commission, Cambridge, UK*.
- Noad M, Dunlop R, Paton D, Kneist HE (2011) Abundance estimates of the east Australian humpback whale population: 2010 survey and update. *Unpublished report (SC/63/SH22) presented to the Scientific Committee of the International Whaling Commission, Cambridge, UK*.
- Olavarria C, Baker CS, Garrigue M, *et al.* (2007) Population structure of South Pacific humpback whales and the origins of the East Polynesian breeding grounds. *Marine Ecology Progress Series* **330**, 257-268.

Pledger SA (1998) Finite mixture in closed capture-recapture models. *Research Report 98-18, School of Mathematical and Computing Sciences, Victoria University of Wellington, New Zealand.*

Pledger SA (2000) Unified maximum likelihood estimates for closed capture-recapture models using mixtures. *Biometrics* **56**, 434-442.

Seber G (1982) *The estimation of animal abundance and related parameters* Edward Arnold, London.

Supplementary Table 1: Estimates of male superpopulation size (N_S), apparent survival (Φ) and population growth rate (λ) from the λ -POPAN model fitted to the Oceania male genotype mark-recapture dataset from Constantine et al (2012). The model parameters are shown on the left and are population growth rate (λ), apparent survival (Φ), and capture probability (p) and follow this notation: (.) denotes the model parameter is time-invariant, (t) denotes the model parameter varies with capture occasion and (0.95) denotes the model parameter is fixed at 0.95. Where the model did not return variance estimates due to non-convergence or boundary estimates, CIs are marked as NA. Model fit is shown by Δ AIC. It should be noted that the sparseness of the data (as highlighted by Constantine *et al.* 2012) and short time span of the 1999-2005 surveys meant these data can be problematic for mark-recapture modelling.

Model	Δ AIC	N_S	95% CL	λ	95% CL	Φ	95% CL
$\lambda(.)p(t)\Phi(0.95)$	0.00	1257	1005, 1573	0.99	0.86, 1.11	0.95	fixed
$\lambda(.)p(t), \Phi=1=\lambda$	0.35	1122	918, 1370	1.00	fixed	1	fixed
$\lambda(.)p(t)\Phi(.)$	4.28	1153	NA	1.02	NA	1	NA
$\lambda(.)p(.)\Phi(0.95)$	112.31	1342	1088, 1657	1.02	0.97, 1.02	0.95	fixed
$\lambda(.)p(.)\Phi(.)$	113.37	1457	1154, 1840	1.01	0.97, 1.02	0.88	0.74, 1.00

Supplementary Table 2: Estimates of male superpopulation size (N_S), apparent survival (Φ) and population growth rate (λ) from the λ -POPAN model fitted to the New Caledonian male genotype mark-recapture dataset from Constantine et al (2012). The model parameters are shown on the left and are population growth rate (λ), apparent survival (Φ), and capture probability (p) and follow this notation: (.) denotes the model parameter is time-invariant, (t) denotes the model parameter varies with capture occasion and (0.95) denotes the model parameter is fixed at 0.95. In some models sample size is used as a proxy for survey effort and years with similar survey effort are constrained to have the same capture probability, as shown by '=' , for example, 1999=2002=2004 indicates these three years are constrained to have the same capture probability. Where the model did not return variance estimates due to non-convergence or boundary estimates, CIs are marked as NA. Model fit is shown by Δ AIC.

	Δ AIC	N_S	95% CL	λ	95% CL	Φ	95% CL
$\lambda(.)p(1999=2002=2004, 2001=2003=2005, 2000), \Phi(.)$	0.00	529	406, 690	1.02	0.94, 1.09	0.78	0.62, 0.93
$\lambda(.)p(1999=2002=2004, t), \Phi(0.95)$	1.16	531	401, 704	1.12	1.00, 1.24	0.95	fixed
$\lambda(.)p(1999=2002=2004, 2001=2003=2005, 2000), \Phi(0.95)$	2.25	495	383, 639	1.03	0.96, 1.11	0.95	fixed
$\lambda(.)p(1999=2002, t)\Phi(0.95)$	3.06	523	390, 702	1.11	0.95, 1.26	0.95	fixed
$\lambda(.)p(t)\Phi(0.95)$	4.80	538	389, 743	1.13	0.94, 1.32	0.95	fixed
$\lambda(.)p(1999=2002, t)\Phi(.)$	7.08	508	NA	1.12	NA	1.00	NA
$\lambda(.)p(1999=2002=2004, t), \Phi(0.95)$	19.94	953	NA	1.18	NA	1.00	NA
$\lambda(.)p(.)\Phi(0.95)$	40.77	536	411, 698	1.09	1.02, 1.16	0.95	fixed
$\lambda(.)p(.)\Phi(.)$	40.86	560	429, 730	1.09	1.02, 1.16	0.83	0.67, 0.99
$\lambda(.)p(t)\Phi(.)$	354.38	205	193, 219	1.07	0.85, 1.28	1	NA

Supplementary Table 3: Estimates of male superpopulation size (N_S), apparent survival (Φ) and population growth rate (λ) from the lambda-POPAN model fitted to the Tongan male genotype mark-recapture dataset from Constantine et al (2012). The model parameters are shown on the left and are population growth rate (λ), apparent survival (Φ), and capture probability (p) and follow this notation: (.) denotes the model parameter is time-invariant, (t) denotes the model parameter varies with capture occasion and (0.95) denotes the model parameter is fixed at 0.95. In some models sample size is used as a proxy for survey effort and years with similar survey effort are constrained to have the same capture probability, as shown by '=', for example, 1999=2002=2004 indicates these three years are constrained to have the same capture probability. Where the model did not return variance estimates due to non-convergence or boundary estimates, CIs are marked as NA. Model fit is shown by Δ AIC.

	Δ AIC	N_S	95% CL	λ	95% CL	Φ	95% CL
$\lambda(.)p(2000=2005, 2001=2002=2003)\Phi(.95)$	0	1549	852, 2815	0.98	0.89,1.08	0.95	fixed
$\lambda(.)p(2001=2002, 2000=2005, 2003)\Phi(0.95)$	1.05	1130	793, 1609	0.95	0.85,1.05	0.95	fixed
$\lambda(.)p(2000=2005, 2001=2002=2003)\Phi(.)$	1.93	1466	699, 3074	0.99	0.89,1.09	0.98	0.79, 1.00
$\lambda(.)p(t)\Phi(0.95)$	3.92	1702	NA	0.95	NA	0.95	fixed
$\lambda(.)p(2001=2002,t)\Phi(0.95)$	7.91	817	413, 1616	0.95	0.61,1.29	0.95	fixed
$\lambda(.)p(.)\Phi(.)$	19.82	1519	745, 3099	0.96	0.88,1.04	0.96	0.77, 1.00
$\lambda(.)p(.)\Phi(0.95)$	22.08	890	594, 1332	0.95	0.88,1.02	0.95	fixed