# Report of the Workshop on Hector's and Māui Dolphins in New Zealand: Consideration of Spatial Risk Assessment of Threats

Bled, Slovenia, 22-23 April 2023

# Report of the Workshop on Hector's and Māui Dolphins in New Zealand: Consideration of Spatial Risk Assessment of Threats[1]

## 1. INTRODUCTORY ITEMS

The Scientific Committee pre-meeting Workshop on '*Hector's and Māui Dolphins in New Zealand: Consideration of Spatial Risk Assessment of Threats*' was held at the Rikli Balance Hotel, Bled, Slovenia from 22-23 April 2023. The list of participants is given as Annex A.

Fortuna welcomed the participants, thanking them for their time and written contributions, especially the reviewers and the experts from the New Zealand Government. Zerbini (Scientific Committee Chair) and Staniland (IWC Secretariat Head of Science) also expressed their gratitude to the New Zealand government for financial support and availability throughout the organisational phase.

### 1.1 Appointment of rapporteur/s
Cipriano acted as rapporteur.

### 1.2 Adoption of agenda
The adopted agenda is given as Annex B.

### 1.3 Review of available documents
The list of documents is given as Annex C.

### 1.4 Objectives of the pre-meeting
A report (Roberts *et al.*, 2019) on the development of a spatial risk assessment model (SEFRA[2]) to examine threats to Hector's and Māui dolphins (*Cephalorhynchus hectori*) was presented to the 2019 Scientific Committee meeting. There was insufficient time for a thorough review of this large and detailed report (IWC, 2020). A review process was agreed, and the Committee recommended that several subjects should be reviewed by independent reviewers intersessionally, co-ordinated by a steering group under Staniland, and review topics should include Māui and Hector's dolphins' life history parameters; spatial distributions; estimates of bycatch rates and vulnerability; the potential of a toxoplasmosis effect on the modelling; and the risk model outputs (IWC, 2020). Due to the disruptions related to the COVID19 pandemic, progress with the review was stalled until the government of New Zealand offered in 2022 to sponsor it (IWC, 2023). The Scientific Committee developed the following terms of reference for the review and pre-meeting, to be held immediately prior to SC69A (IWC, 2023):

(1)  evaluate the design and structure of the multi-threat risk assessment model;
(2)  evaluate the overall sensitivity to model choices, data selection, uncertainties or potential biases identified in the review papers; and
(3)  make recommendations to reduce key uncertainties and improve the utility of the model to inform management decisions.

Although the primary focus was Roberts *et al.* (2019), any relevant new data and analyses should also be carefully reviewed intersessionally prior to the pre-meeting (IWC, 2023). Reviewers were provided with written responses to their comments from the New Zealand government experts in advance of the pre-meeting (Appendix 8).

The HIM sub-committee work plan will also include an examination of whether the risk assessment work undertaken in New Zealand might assist with the work of the IWC Bycatch Mitigation Initiative (BMI), as there are several topics included in the review that will also apply to bycatch risk assessments more generally (IWC, 2023).

### 1.5 Overview
The SEFRA approach involves several independent but related steps: (a) estimation of seasonal spatial dolphin density surfaces based on habitat models fitted to Hector's dolphin aerial survey data and opportunistic sightings of Māui dolphins; (b) estimation of parameters related to the rate of dolphin interaction and death in commercial set net and inshore trawl fisheries (each fishery group is considered independent of the other fishery group); (c) estimation of total dolphin captures (non-lethal interactions and deaths) in commercial fisheries (inshore trawl and commercial set net) by overlaying the

---

[1]Presented to the SC meeting as SC/69A/REP/04.
[2]Spatially Explicit Fisheries Risk Assessment.

dolphin density surfaces with commercial fisheries intensity and accounting for risk of interaction and death; (d) calculation of dolphin deaths due to all causes other than commercial fishing (including due to infection with *Toxoplasma gondii*) and allocation of non-fishing deaths to various causes; and (e) comparison of the deaths to a Population Sustainability Threshold (PST) related to population recovery.

The SEFRA model uses a Bayesian framework in which a prior probability distribution is defined for each parameter, and information from available data (defined by a likelihood function) is used to update prior distributions, resulting in a posterior probability distribution for each parameter. In the case of Hector's and Māui dolphins, some parameters were not informed by the data and hence were based entirely on prior distributions (see the Appendices to Roberts *et al.*, 2019). Model estimated parameters (see Appendix 1 to this report) included: (i) vulnerability, which gives the probability of a capture event; and (ii) the probability of a dolphin being alive when captured, for which the fisheries observer data were informative. Variables whose priors were not updated given the available data included: (i) the probability a capture event is observable, given that a fisheries observer is on watch; (ii) the probability of survival on release; and (iii) population size by subspecies. The values for these variables were therefore determined from their assumed prior distributions. Presently, the SEFRA approach does not propagate all known sources of uncertainty through each stage of the analysis; notably, the uncertainty in the dolphin density surface models is not propagated into the final estimates of dolphin deaths and annual risk ratios.

All model parameters, priors and the derivation of priors can be found in Tables 4 and 5 and Appendix 10 of Roberts *et al.* (2019), respectively; model diagnostic plots are given in Appendix 13 (*ibid.*).

The implementation of the SEFRA model can estimate annual deaths and risk for other predefined lethal threats. Briefly, this involves fitting a Dirichlet probability distribution to proportional causes of death in the necropsy sample of stranded dolphins, including toxoplasmosis, predation by sharks, and all other causes grouped together. The resulting posterior probability distributions (see also p.144 of Roberts *et al.*, 2019) are used to partition the estimated annual deaths for each subspecies (derived from an assumed prior on annual survival rate applied to population size), once estimated commercial fisheries deaths had been removed. A sensitivity run explored an alternative scenario in which the relative detection probability for shark predation was modified (Appendix 14 of Roberts *et al.* (2019)). Estimates of annual deaths and the annual risk ratio for non-fisheries causes of death were also plotted and tabulated in section 5.3 of Roberts *et al.* (2019), p.48.

## 2. REVIEW OF THE REPORTS FROM INDEPENDENT EXPERTS

The pre-meeting followed the following approach:

(1) New Zealand scientists briefly introduced each of the relevant sections of Roberts *et al.* (2019) under each agenda item;
(2) each independent reviewer presented: (a) their original reviews with a focus on recommendations for improvements to the original model; and (b) any additional comments/recommendations in response to the written comments provided by New Zealand scientists prior to the pre-meeting (Appendix 8 to this report); and
(3) general discussion focussed on the reviews and responses above.

The overarching goal of the pre-meeting was to develop conclusions on the applicability and robustness of the SEFRA modelling approach to provide advice related to the management of Hector's and Māui dolphins, given the available information, and to provide recommendations for further work where applicable. Management implications will be discussed during the HIM sub-committee.

### 2.1 Estimation of $r_{max}$

Roberts summarised the methods used to estimate $r_{max}$ (see definition in Appendix 1 to this report) for Hector's and Māui dolphins, in the implementation of the SEFRA risk assessment model. This followed the method proposed by Dillingham *et al.* (2016) and implemented by Moore (2015), which generates distributions of λ (see definition in Appendix 1) given distributions for the 'optimal' values of life history parameters, and then filters out biologically implausible samples using an allometric invariant observed across all vertebrates between optimal generation time ($T_{opt}$) and $r_{max}$. The pre-meeting was informed that Dillingham reviewed the implementation of his method by Edwards *et al.* (2019) and concluded that it was broadly appropriate. Appendix 1 of Roberts *et al.* (2019) describes supplementary ageing and maturity information used in this estimate. Appendix 2 of Roberts *et al.* (2019) describes the estimation of $r_{max}$ for Hector's dolphin, while Appendix 3 of Roberts *et al.* (2019) describes how this distribution is adjusted for Māui dolphin, based on individual-based models representing some Allee mechanisms. The resulting $r_{max}$ posterior for Māui dolphin, as used by the SEFRA model, had a mean value (0.045). This mean value was much higher than was used previously for this subspecies (0.018), and slightly above the default value used in applications of the Potential Biological Removal (PBR) approach (0.04), although is consistent with allometric relationships. Roberts *et al.* (2019) [p.79] acknowledged that some other Allee mechanisms and low or variable prey availability were not accounted for in the derivation of the distribution for $r_{max}$. These additional factors could potentially reduce the distribution for $r_{max}$ for Māui dolphins to lower values.

*2.1.1 Review demographic rates, estimation of $r_{max}$ and its use in the assessment*

The approach of Appendix 1 of Roberts *et al.* (2019) for deriving a distribution for $r_{max}$ relies on distributions for survival, age-at-maturity, maximum age and fecundity, as the Authors believed that these were likely to include an appropriate range for when the population is growing at $r_{max}$. Extensive discussions within Scientific Committee over many years have emphasised the difficulties in deriving estimates of $r_{max}$ by consensus, recognising that different approaches have different strengths and weaknesses (IWC, 2023). DeMaster provided an initial set of comments on the different approaches used to estimate $r_{max}$ and how this was handled by Roberts *et al.* (2019) (see Appendix 2 to this report).

There was general agreement that the approach used by Roberts *et al.* (2019) to derive a distribution for broad $r_{max}$ was reasonable given the lack of direct information on the rate of increase at low population size. Discussion focussed more on the appropriate method(s)/data in this case, how to account for the inevitable uncertainty, understanding how the allometric invariant modifies the distributions for the life history parameters, and the robustness of any management-related conclusions to uncertainty. There is a reasonable amount of long-term demographic information (e.g., survival, age-at-maturity, maximum age and fecundity) for this species. The data are consistent with previous data (see Table 1 in Appendix 2), which makes the change in the range of values for $r_{max}$ surprising, implying that the allometric invariance is consequential. Recent information on the proportion of calves in Hector's and Māui dolphin populations should be compared with the implication of the distribution of $r_{max}$ in Roberts *et al.* (2019).

Several general factors that need to be taken into account when deciding $r_{max}$ values to be used in the Roberts *et al.* approach and how to incorporate uncertainty in a modelling context were highlighted including:

(a) the use of 'real' data on demographics and trends requires a good understanding of the nature of the data sources at several levels including the 'status' (with respect to carrying capacity) of the populations for which the data are sampled (including consideration of $r_0$ vs $r_{max}$);

(b) the lack of detail on how the allometric invariant impacts the distribution for $r_{max}$ inferred from the input distributions for the life history parameters, which should be quantified by plots of how the input distributions for the demographic parameters are modified by the allometric invariant which appears to be substantial;

(c) the applicability and implications of using data/obtaining estimates of $r_{max}$ at interspecific, species and population levels;

(d) the need to take both environmental and demographic stochasticity into account, especially for populations that are 'small' due to either intrinsic factors or reduction from a historically large population size; and

(e) the need to consider the implications of different degrees of habitat suitability between populations, including large differences in prey availability.

As noted in Appendix 3, with respect to the latter point and Hector's/Māui dolphins, it was **agreed** that, given the circumstances (e.g., see Cooke *et al.* (2019); Cooke *et al.* (2018), an $r_{max}$ estimate for Māui similar to that for Hector's is hard to reconcile with the available information, because Māui dolphins are either naturally rare and/or severely depleted, with either scenario implying a low $r_{max}$. Indeed, it is confusing to use the term $r_{max}$ for both species and populations. It would be better to distinguish, as suggested, for example, by Cooke *et al.* (2019), between $r_0$, the mean growth rate of a population at low population size, which will differ between core and marginal populations, and $r_{max}$, the maximum growth rate of a species under good conditions.

The pre-meeting **agreed**, however, that Appendix 2 of Roberts *et al.* (2019) is insufficiently detailed and **strongly recommended** that application of the Dillingham *et al.* (2016) method be more fully described, including documentation of how the joint prior for the life history parameters is modified owing to the constraint imposed by the allometric invariant.

## 2.2 Spatial distribution of Hector's and Māui dolphins

The SEFRA approach requires estimating summer and winter spatial density surfaces for Hector's and Māui dolphins, which are then used by the SEFRA model to calculate spatial overlap with commercial fishing (inshore trawl and commercial set net), to adjust for potential biases in overlap comparing observed and non-observed fishing effort, when estimating total annual fisheries deaths and risk. The habitat variables that were considered to be potential predictors of dolphin density included [sea surface turbidity, sea surface chlorophyll-a concentration, KPAR (light attenuation), bathymetric slope, depth, distance to shore, and the probability of presence of the six main prey species in terms of dietary mass contribution]. Gradients in some of the non-prey layers were also considered, although were rejected early in the model exploration process. The best predictors of spatial distribution were found to be surface turbidity levels and prey presence (Roberts *et al.*, 2019 p.29-41). These and other candidate habitat variables were plotted spatially in Appendices 5 and 6 of Roberts *et al.* (2019). The final spatial distribution layers are plotted in Appendix 7 of Roberts *et al.* (2019), where they are overplotted with other sources of spatial information not used for model fitting. Subsequent research has updated the spatial distribution modelling for the Hector's dolphin subspecies (Roberts and Webber, 2023) and addressed some of the more critical reviewer recommendations made in Appendix 4 to this report.

### 2.2.1 Review estimation of spatial distribution of Hector's and Māui dolphins and its use in the assessment

Ferguson provided a thorough review of the density surface models (DSMs) for Hector's and Māui dolphins and the spatial methods used to infer risk to the dolphins by commercial fishing. Ferguson's initial review included a set of recommendations (see Appendix 4), which were used as the basis for further analyses prior to the pre-meeting and subsequent discussion during the pre-meeting.

Ferguson's review focussed on: (i) identifying whether 'best' practices were used in Roberts *et al.* (2019) dolphin and bycatch spatial analyses, providing references and resources for cases in which alternate preferred methods are available; and (ii) evaluating whether the authors' conclusions were supported by the results of their spatial models and by other available information about the ecology and fisheries that is relevant to the issue of Hector's and Māui dolphin management in New Zealand. Two key themes were: (a) the spatial and temporal scales (resolution and extent) in the data, model construction, and inference; and (b) the identification and evaluation of uncertainty, including both precision and bias.

After a detailed discussion regarding the model design and structure, the pre-meeting **agreed** that consideration of the points (1)-(5) below would assist in improving estimates of summer and winter spatial density surfaces for Hector's and Māui dolphins. However, it also **agreed** that addressing these considerations does not prevent the results from presently being used in a management context. These are:

(1) model prey biomass or abundance instead of occurrence, including seasonal changes in prey availability, if such data are available;
(2) evaluate a complementary log-log link function for the prey occurrence models (Appendix 4);
(3) investigate the use of anisotropic splines and barrier-respecting splines in the prey and dolphin spatial models (Appendix 4);
(4) try relaxing the restriction on the basis dimension ($k$) defining the smoothing splines in the GAMs (Appendix 4); and
(5) investigate the availability and utility of remotely sensed data that provide information about mesoscale oceanographic activity such wind direction and strength, upwelling intensity, ocean surface fronts, and eddy kinetic energy, which may affect prey availability and, hence, dolphin density (Appendix 4).

The pre-meeting made several **recommendations** to improve the statistical methodology used in the dolphin and bycatch spatial analyses, to reduce key uncertainties and to improve the utility of the SEFRA approach to inform management decisions. It noted that while implementation of these recommendation might not change the results that are relevant for making management decisions considerably, they allow for greater confidence in the robustness of the results and help to better understand uncertainty and sensitivity (and should be considered in any future applications of this approach for Hector's and Māui dolphins and for other species/areas). These were:

(1) examine the effect of different spatial resolutions (e.g., 5km x 5km or 10km x 10km) for the dolphin DSMs (Appendix 4);
(2) incorporate survey effort and detectability as an offset term on the right-hand side of the dolphin DSM model equations (Appendix 4);
(3) compare the fit of the negative binomial and Tweedie probability distributions for the coastal and harbour dolphin DSMs using QQ plots and residual plots rather than rely solely on AIC values (Appendix 4);
(4) consider a hierarchical GAM structure or factor-smooth interactions to incorporate seasonal differences in dolphin distribution with respect to distance offshore, depth, and slope in the coastal dolphin DSM (Appendix 4);
(5) revise the harbour dolphin DSM (Appendix 4, Section D), taking into account: the nature of public Māui dolphin sightings; whether the recreational aerial survey data can provide a relatively unbiased estimate of public observation effort; the sensitivity of model results to deletion of sightings whose midpoints were on land; known or hypothesized effects of anthropogenic activities on dolphin distribution and density in harbours; and performance of a purely spatial DSM for dolphins in harbours.
(6) expand the investigation of the consequences of using only summer prey data[3] - this should be reconsidered to see if seasonal prey data fit the model better than seasonal turbidity vs. alternative covariates (Appendix 4);
(7) revise estimated CVs in the dolphin density predictions to incorporate uncertainty in the estimates of detection probabilities for the dolphin aerial line-transect surveys, and propagate that uncertainty through the risk assessment model (Appendix 4); and
(8) undertake sensitivity analysis to investigate the effects of making predictions from the DSM based on climatological vs. contemporaneous covariates (Appendix 4).

In the context of (4) above, the pre-meeting welcomed the fact that in response to one recommendation of the independent review (i.e., background densities; Appendix 4), Roberts and Webber (2023) provided a new analysis in which regional dolphin densities were allowed to vary between the summer and winter seasons due to a purely spatial smooth (not explainable by the habitat covariates). This revised model inferred much lower densities occurred in the south and

---

[3]Winter prey information was absent in some regions and was only used for the East Coast South Island, where most dolphins are caught.

north of the South Island relative to the east and west coasts. In addition, Roberts and Webber (2023) explored alternative error distributions when updating the South Island Hector's dolphin modelling, but only used AIC to compare them. In discussion, it was noted that additional aerial survey data is available from three summers and three winters (Rayment *et al.*, 2010), which would help address (4) above or could provide the base for model validation.

The pre-meeting **recommended** that a spatial cross-validation of the dolphin DSM be conducted to evaluate the assumption that Hector's and Māui dolphin habitat preferences are constant throughout New Zealand's waters (Appendix 4). Roberts and Webber (2023) trialled two alternative cross-validation approaches while updating the spatial distribution modelling for South Island Hector's dolphins: (i) a random k-fold approach, which favoured the most complex candidate model, and (ii) a source-based cross validation approach, iteratively withholding different sources of information as the test data, which favoured models with only a couple of habitat terms. The pre-meeting deemed the second of these approaches to be best for guiding the selection of models with good predictive power. Based on the information provided in Roberts and Webber (2023), the 'data source' method of cross-validation confounded the effects of spatial region and data type in the validation process. Seasonal prey data are available to explore the relationship between dolphin distribution and the availability of their prey but only at some locations (e.g. Miller *et al.*, 2013).

### 2.3 Estimates of bycatch rates and vulnerability of Hector's and Māui dolphins relative to a Population Sustainability Threshold (PST)

Roberts reiterated that the risk assessment of threats to Hector's/Māui dolphins used the spatially-explicit fisheries risk assessment (SEFRA) approach for estimating annual commercial fisheries deaths and risk relative to a Population Sustainability Threshold (PST). He summarised the fishery data inputs used by the SEFRA model to estimate vulnerability to capture and annual deaths - namely the fisheries observer data fitted to by a model. The relevant SEFRA model methods are described in Roberts *et al.* (2019), observed captures are tabulated in Table 11 (*ibid.*) and plotted spatially in Appendix 15 (*ibid.*), along with spatial plots of observed effort and observer coverage rate.

Roberts noted that the SEFRA model estimates of deaths were greatly influenced by the prior on the extent of cryptic mortality rate[4], which was created based on a review in Appendix 10 in Roberts *et al.* (2019). Roberts *et al.* (2019) concluded that better information is needed to inform cryptic mortality rate and the pre-meeting **concurred** with this view, noting that inclusion of cryptic mortality is a strength of this approach.

Furthermore, it is plausible and indeed likely that vulnerability to capture varies across different sectors of the commercial set net fleet Roberts *et al.* (2019), but this is not modelled in Roberts *et al.* (2019). Coarse-level target species vulnerability groupings were explored in subsequent modelling (Roberts and Webber, 2023), resulting in a small reduction in estimated deaths and risk for Hector's dolphins in the assessed locations.

### 2.3.1 Review of estimates of bycatch rates and vulnerability of Hector's and Māui dolphins relative to a Population Sustainability Threshold (PST)

Ferguson provided detailed comments on the issue of cryptic mortality rate and an initial set of conclusions and recommendations (see Appendix 4 to this report). Peltier provided several comments on bycatch rates and design of monitoring programmes (Appendix 5 to this report). These were used as basis for the Panel's discussion.

Various issues were debated, including implications of low and non-random observation effort, whether spatial overlap between species density and effort is an adequate measure of bycatch risk, various biases in bycatch rate as a result of observer presence, sensitivity of estimated bycatch to the choice of priors, and implications of temporal changes in abundance or distribution. It was noted that the bycatch model accounts for the location of observed hauls and sets and hence partially addresses of the effects of fishers changing behaviour in response to the presence of observers. This part of the model could be better understood by plotting the spatial distribution of observed and unobserved fishing effort.

New Zealand experts acknowledged the low and highly variable observer coverage that has been achieved in New Zealand's inshore set net and trawl fisheries. They provided information on New Zealand's observer programme and examples of targeted observer coverage areas (Appendix 8 to this report).

### 2.4 Risk model outputs

The key outputs of the SEFRA are probability distributions for the parameters used to define annual commercial fisheries deaths (arising from entanglement with each type of fishing gear) and the risk ratios relative to a PST. The inshore trawl and commercial set net are the only gear type for which risk ratios were computed, with separate vulnerability parameters estimated for each fishery group. There are other fishery and aquaculture sectors, e.g., recreational, rock lobster and 'customary' for which these ratios are not computed. It was noted that these sectors include those for which entanglements have been anecdotally reported but not observed (e.g. red rock lobster pot fishery); recreational fishing, for which there are historically reported captures, but little information on reporting rate.

---

[4]Mortality due to pre-catch losses, ghost fishing, and post-release mortality.

Given the respective overlap of each commercial fishing event with the dolphin density surface, the SEFRA model posteriors were used to infer the annual number of dolphin deaths across all fishing events in the commercial gillnet and inshore trawl fisheries (observed and unobserved). For each assessed fishery group, estimates of annual deaths and annual risk ratio (deaths divided by the PST) were averaged across the final three years of fishing effort included in the model (see section 5.3 of Roberts *et al.*, 2019).

### 2.4.1 Review of risk model outputs
The pre-meeting noted the complexity of the modelling approach and that there was limited information available to inform some stages of the Hector's and Māui dolphin risk assessment. In particular, the model-based estimates of dolphin deaths and risk ratios associated with the commercial fisheries are sensitive to the estimated values of key model parameters ($v_g$, $P_g^{obs}$, $\psi_g$, and $\omega_g$). This has an effect on both precision and biases of model outputs and ability to validate the model. This is addressed further under Item 3. Inshore gillnet fisheries include a variety of metiers[5], each of them likely differs in terms of risk to Hector's and Māui dolphins; however, the data available do not allow a risk analysis to distinguish this.

The pre-meeting **recommended** that the PST be computed accounting for the correlation between survival and $r_{max}$ in the calculation of the PST.

## 2.5 Other anthropogenic threats
Roberts summarised methods used to consider other anthropogenic threats in Roberts *et al.* (2019). A base run assumed that the proportional non-fisheries causes of death across each subspecies were the same as in the sample of stranded animals that had been necropsied, with a sensitivity run exploring a scenario where the relative detection probabilities of shark predation differed. The assessment also estimated relative spatial overlap for other threats, e.g., from recreational netting, and other anthropogenic threats, which provided estimates of the relative threat intensity by subpopulation (p.47 of Roberts *et al.* (2019).

The necropsy data are summarised in Appendix 4 of Roberts *et al.* (2019). The necropsy methods are described in pp.11-13, the relevant risk assessment methods (pp.28-29), and methods for generating spatial threat layers are given in Appendix 8 (shark predation risk) and Appendix 9 (*T. gondii* oocyst densities). The risk assessment results for other anthropogenic threats are described in pp.42-56 (*ibid.*), with the main conclusions on pp.73-75. Roberts commented that whilst acknowledging the many potential biases associated with stranded animals, if the necropsy samples are in any way representative of the wider proportional causes of death, toxoplasmosis is likely to pose a significant anthropogenic population risk to Māui and Hector's dolphin.

### 2.5.1 Identify data or research priorities to improve understanding of Toxoplasmosis
Grigg provided an overview of issues related to Toxoplasmosis detection and mitigation, as this pathogen may affect the values for the parameters used in most population dynamics models (Appendix 6 to this report).

Roberts *et al.* (2019) provided multiple tables and figures (see table 14, fig. 19, figs A12-2 of their report) that listed several anthropogenic stressors and other variables potentially affecting the health and resilience of New Zealand dolphins. The variables were listed as frequency estimates, with 1.0 indicating greatest risk, and 0.0 lowest. However, the rankings were not based on the absolute number of cases assessed for each variable, which seemed problematic as it can inflate the perceived threat. In absolute terms, the likelihood of the threat from, for example, an oil spill, or aquaculture, is negligible, compared to other factors such as infectious diseases or fishery bycatch. Grigg recommended the absolute numbers should be reported for each source, to better inform relative risk and prioritise research on the most relevant factors that impact dolphin recovery efforts.

Death is not the only outcome from infection with *Toxoplasma gondii.* This parasite alters reproductive fitness, can affect the survival of neonates, impact the behaviour of positive individuals, and can significantly impact the health and population recovery of marine mammal populations, as has been reported for other marine species (e.g., Barbieri *et al.*, 2016; Gibson *et al.*, 2011; and Miller *et al.*, 2004). Grigg noted that one approach may be to specifically quantify population level risks associated with *Toxoplasma* by comparing the model estimates of deaths due to *Toxoplasma gondii* by incorporating a PRP (Potential Removal due to Protozoan Infection) variable with that of the Potential Biological Removal Factor (PBR) (Wade, 1998).

In terms of monitoring and possibly mitigating the issue, several aspects were debated. Coupe (2021) has identified seroprevalence rates of >60% in companion cats, established that 1.6% of feral/stray cats assayed were actively shedding oocysts and produced estimated oocyst load calculations of between 190-240 oocysts/sq. meter, which is significant, and in line with studies pursued in California, where documented land-to-sea transmission of *Toxoplasma* infection into threatened marine mammal species has been established. Coupe (2021) also showed that 13 of 104 mussels were PCR positive for *Toxoplasma*, and that within field sites in key Māui dolphin habitats, *ca.*2% of mussels in the river terminal tested positive for *Toxoplasma*.

---

[5]A 'metier' is a group of fishing operations targeting a specific assemblage of species, using a specific gear, during a precise period of the year and/or within the specific area.

The possibility of Live Capture Initiatives to assess health and serostatus to profile the risk of exposure to infectious disease agents that may impact the health and resilience of the New Zealand dolphin species was discussed. So far, there have been no published serology studies, nor screening for the prevalence of infection in the Hector's or Māui dolphin populations. A representative sample of the population, regardless of disease state, is necessary to generate baseline datasets. Nor have there been any assessments on the ability of the parasite to cause congenital infection and alter reproductive fitness. Finally, Grigg inquired at the pre-meeting whether vaccination is being considered, since ongoing discussions among caretakers of the critically endangered Hawaiian monk seal have expressed an interest in utilizing the S48 ToxoVax vaccine, which is routinely used in New Zealand among livestock populations. While it is one possible prevention initiative to consider, his preference was to consider transmission blocking vaccines, such as T-263, that block oocyst formation in cats and thereby reduce both the environmental load of oocysts and the potential exposure risk to the Hector's or Māui dolphin populations at river catchments predicted to be hotspots for infection.

After receiving this information, the pre-meeting **endorsed** the following recommendation by Grigg:

(1)  necropsy every cetacean, including bycaught dolphins to estimate the prevalence of *Toxoplasma* infection in Hector's and Māui dolphin populations;

(2)  produce cat habitat suitability models with field-based surveys of stray and feral cat density;

(3)  although to date Type II *Toxoplasma* has not been detected in New Zealand cats to draw any reasonable conclusions – the number of cats examined needs to be increased substantially to assess the true genetic diversity of *Toxoplasma* strains[6] infecting cats;

(4)  estimate coastal *Toxoplasma* oocyst density by screening tissues of filter-feeding species collected from target at-risk coastal locations;

(5)  assess whether parasite genotype impacts disease - to reliably ascertain this, all stranded and bycaught dolphins, regardless of age and disease presentation, should be tested using PCR and/or serology to determine *Toxoplasma* exposure, and genotype the infecting parasite to identify all strains of *Toxoplasma* capable of infecting New Zealand dolphins. Further, prey animals and cats upstream of river terminals frequented by New Zealand dolphins should be assessed for *Toxoplasma* infection status and genotype, to ascertain population genetic diversity of the parasite species in New Zealand, in general; and

(6)  determine whether diet preference increases the risk of infection and the development of acute toxoplasmosis - this is critical information to generate.

The New Zealand government experts thanked Grigg for these suggestions and informed the pre-meeting that they are setting up a *Toxoplasma* programme and all these recommendations are timely to improve their plans.

### 2.6 Other information
#### 2.6.1 Feedback from ASG on Constantine et al. (2021)
The pre-meeting received information on a new abundance estimate of Māui dolphins by Constantine *et al.* (2021) based on microsatellite genotypes. The estimate is 48 Māui dolphins in 2021, with 95% confidence interval (40, 57)[7]. It was endorsed as Category 3, meaning that the estimate is informative for providing a general indication of abundance, but subject to various possible concerns that preclude a higher categorization. Such an estimate might be used for fitting population models but is not suitable for use in implementations of IWC management procedures.

## 3. CONCLUSIONS AND RECOMMENDATIONS TO THE SCIENTIFIC COMMITTEE
## SMALL CETACEANS SUB-COMMITTEE

SEFRA provides a way to integrate data in a probabilistic manner and capture parameter uncertainty using probability distributions. Although information for some critical inputs to the model is limited, the ability to identify risks spatially is a strength of the approach, allowing it to be used to assess spatial management strategies.

The pre-meeting had extensively considered various aspects on the SEFRA model with the intent of providing the best possible advice to the New Zealand government on potential aspects of the SEFRA model that needed to be improved. Most of the analyses/parameter values are based on data for Hector's dolphins because of the extremely small size of the population of Māui's dolphins. However, given the results in Appendix 3 of Roberts *et al.* (2019) and the discussions in this report, Māui dolphins are likely to be less (perhaps substantially) resilient than Hector's dolphins.

Previous sections of this report have identified research recommendation that will improve the implementation of SEFRA for Hector's and Māui dolphins. The pre-meeting did not identify obvious flaws in the SEFRA and its application that would preclude its application to support management. However, considerable additional work should be conducted to better

---

[6]Currently, only Type II toxoplasma has been identified from Hector's and Maui dolphins - Type II was not detected in a small sample of New Zealand cats tested so far.
[7]Note: this estimate was later formally endorsed by the full Committee (IWC, this volume).

explore uncertainty and understand which parameters are both uncertain and consequential in terms of management-related outcomes. This section highlights those issues that are likely to affect the results of the SEFRA model if they are not considered, but at an unknown level. These could have implications for New Zealand government management decisions based on Roberts *et al.* (2019). **Recommendations** on these more important aspects are summarized below. They focus on the overall sensitivity to model choices, data selection, uncertainties or potential biases identified in the review papers and aim to improve the utility of the model to inform management decisions.

(1)  It is essential that the uncertainty in the species distribution model is fully propagated into the risk assessment.

(2)  An insufficient exploration of sensitivity is provided in Roberts *et al.* (2019). A more complete exploration should be undertaken that should involve: (a) identifying the key model outputs (e.g. the estimates of deaths due to various sources and the annual risk ratio); (b) re-running the model fixing each key parameter including $v_g$, $P_g^{obs}$, $\psi_g$, and $\omega_g$ in turn to a range of values, and changing structural assumptions within the model, (c) plotting the change in the key model outputs as a function of the values for these parameters. Sensitivity analyses should investigate whether incorporation of vessel-reported captures and (if feasible) necropsy data in the spatial threat intensity model affects the SEFRA model results, particularly the estimated number of deaths and the risk ratio (Appendix 4). The results of these sensitivity can help to identify areas where additional research is needed, particularly when the distributions for the parameters(s) concerned are based primarily on expert judgement.

(3)  Simulation is commonly used by the Committee to evaluate methods. A fuller quantification of uncertainty would involve using simulation to explore the accuracy and precision of model outputs and the ability of the SEFRA contribute to achieving clearly specified management goals.

## 4. ADOPTION OF THE REPORT

The report was adopted by correspondence on the 1st of May 10:00.

Fortuna again thanked New Zealand for supporting the workshop, and the Secretariat for hosting the workshop in such an efficient manner, the rapporteur and the participants for their hard work in completing the review expeditiously. The participants thanked the Chair for her usual fair and efficient conduct of the pre-meeting.

## 5. CITED LITERATURE

Barbieri, M.M., Kashinsky, L., Rotstein, D.S., Colegrove, K.M., Haman, K.H., Magargal, S.L., Sweeny, A.R., Kaufman, A.C., Grigg, M.E., Littnan, C.L. 2016. Protozoal-related mortalities in endangered Hawaiian monk seals *Neomonachus schauinslandi*. *Dis Aquat Organ.* 121(2):85-95. [Available at: https://doi.org/10.3354/dao03047. PMID: 27667806].

Constantine, R., Steel, D., Carroll, E., Hamner, R.M., Hansen, C., Hickman, G., Hillock, K., Ogle, M., Tukua, P. and Baker, C.S. 2021. Estimating the abundance and effective population size of Maui dolphins (*Cephalorhynchus hectori maui*) in 2020-2021 using microsatellite genotypes, with retrospective matching to 2021. Creative Services Team *NZ Department of Conservation*: 66pp. ISBN 9780473600280.

Cooke, J., Constantine, R., Hamner, R.M., Steel, D. and Baker, C.S. 2019. Population dynamic modeling of the Māui dolphin based on genotype capture-recapture with projections involving bycatch and disease risk. New Zealand Aquatic Environment and Biodiversity Report No. 216. 38 pp. [Available at: *www.mpi.govt.nz/news-and-resources/publications*].

Cooke, J.G., Steel, D., Hamner, R.M., Constantine, R. and Baker, C.S. 2018. Population estimates and projections of Maui dolphin (*Cephalorhynchus hectori maui*) based on genotype capture-recapture, with implications for management of mortality risk. Paper SC/67b/ASI05 presented to the IWC Scientific Committee, April-May 2018, Bled, Slovenia (unpublished). 16pp. [Paper available from the Office of this Journal].

Coupe, A. 2021. Investigating *Toxoplasma gondii* in the marine environment in New Zealand: from cats to kai moana (shellfish). PhD thesis, Massey University.

Dillingham, P.W., Moore, J.E., Fletcher, D., Cortés, E., Curtis, K.A., James, K.C. and Lewison, R.L. 2016. Improved estimation of intrinsic growth rmax for long-lived species: integrating matrix models and allometry. *Ecological Applications* 26(1): 322-33

Gibson, A.K., Raverty, S., Lambourn, D.M., Huggins, J., Magargal, S.L., Grigg, M.E. 2011. Polyparasitism is associated with increased disease severity in T*oxoplasma gondii*-infected marine sentinel species. *PLoS Negl Trop Dis.* 5(5):e1142. [Available at: https://doi.org/10.1371/journal.pntd.0001142]. Epub 2011 May 24. PMID: 21629726; PMCID: PMC3101184

International Whaling Commission. 2020. Report of the Scientific Committee. *J. Cetacean Res. Manage. (Suppl.)* 21:1-65.

International Whaling Commission. 2023. Report of the Scientific Committee. *J. Cetacean Res. Manage. (Suppl.)* 24:1-190.

Miller, M.A., Grigg, M.E., Kreuder, C., James, E.R., Melli, A.C., Crosbie, P.R., Jessup, D.A., Boothroyd, J.C., Brownstein, D., Conrad, P.A. 2004. An unusual genotype of *Toxoplasma gondii* is common in California sea otters (*Enhydra lutris nereis*) and is a cause of mortality. *Int J Parasitol.* 34(3):275-84. [Available at: https://doi.org/10.1016/j.ijpara.2003.12.008]. PMID: 15003489.

Miller, E., Lalas, C., Dawson, S., Ratz, H.and Slooten, E. 2013. Hector's dolphin diet: The species, sizes and relative importance of prey eaten by *Cephalorhynchus hectori*, investigated using stomach content analysis. *Mar. Mamm. Sci., 29*: 606-628.

Moore, J.E. 2015. Intrinsic growth (rmax) and generation time (T) estimates for the genera *Sousa*, *Orcaella* and *Neophocaena*. *NOAA Tech. Mem.* NOAA-TM-NMFS-SWFSC-550: 14pp.

Rayment, W., Dawson, S.M. and Slooten, E. 2010. Seasonal changes in distribution of Hector's dolphin at Banks Peninsula, New Zealand: implications for protected area design. *Aquatic Conservation Marine and Freshwater Ecosystems* 20: 106-16.

Roberts, J.O. and Webber, D.N. 2023. Local-scale spatial risk assessment of inshore commercial fisheries on Hector's dolphins. *New Zealand Aquatic Environment and Biodiversity Report* 308(April 2023): 104pp.

Roberts, J.O., Webber, D.N., Roe, W.D., Edwards, C.T.T. and Doonan, I.J. 2019. Spatial risk assessment of threats to Hector's and Māui dolphins (*Cephalorhynchus hectori*). *New Zealand Aquatic Environment and Biodiversity Report No.* 214: 175pp.

Wade, P.R. 1998. Calculating limits to the allowable human-caused mortality of cetaceans and pinnipeds. *Mar. Mamm. Sci.* 14(1): 1-37.

# Annex A

# List of Participants (in-person and virtual)

**Reviewers**
Brost, B.
DeMaster, D.
Ferguson, M.C. (virtual)
Grigg, M.E. (virtual)
Peltier, H.

**New Zealand**
Gibson, W.
Lundquist, D. (Head of Delegation)
Roberts, J.

**Invited Participants**
Baker, C.S.
Butterworth, D.S.
Cipriano, F. (rapporteur)
Cooke, J. (virtual)
Donovan, G.D.
Fortuna, C.M. (pre-meeting Chair)
Goetz, K.
Heinemann, D.
Punt, A.E.
Rose, N.
Slooten, L.

**Scientific Committee Chairs and Convenors**
Leaper, R. (HIM)
Porter, L. (SM)
Zerbini, Z. (SC)

**Secretariat**
Staniland, I.

# Annex B

# Agenda

1. Introductory items
    1.1  Appointment of rapporteur/s
    1.2  Adoption of agenda
    1.3  Review of available documents
    1.4  Objectives of the pre-meeting and *modus operandi*
2. Review of the reports from independent experts
    2.1  Estimation of $r_{max}$
    2.2  Spatial distribution of Hector's and Māui dolphins
    2.3  Estimates of bycatch rates and vulnerability of Hector's and Māui dolphins
    2.4  Risk model outputs
    2.5  Other anthropogenic threats
    2.6  Other information
3. Recommendations to the Scientific Committee small cetaceans sub-committee
    3.1  Recommendations and conclusions on the design and structure of the multi-threat risk assessment model
    3.2  Recommendations and conclusions on the overall sensitivity to model choices, data selection, uncertainties or potential biases identified in the review papers
    3.3  Recommendations to reduce key uncertainties and improve the utility of the model to inform management decisions
4. Adoption of the Report

# Annex C

# List of Documents

Working papers produced specifically for use at the 2023 pre-meeting.

**SC/69A/MD/WP/**
01.  Dr Doug DeMaster. Considerations regarding an Estimate of $r_{max}$ for Hector's and Māui dolphin (20 March 2023).
02.  Dr Brian Brost. Spatial risk assessment of threats to Hector's and Māui dolphins.
03.  Dr Hélène Peltier. Spatial risk assessment of threats to Hector's and Māui dolphins.
04.  Dr Megan Ferguson. Review of Spatial Risk Assessment for Hector's and Māui Dolphins in New Zealand: Spatial Distribution.
05.  New Zealand Government. IWC review of $r_{max}$ - additional information from the New Zealand government [to SC/69A/MD/WP/01] – DRAFT.
06.  New Zealand Government. IWC review of spatial distribution modelling - a response by the New Zealand government [to SC/69A/MD/WP/04].
07.  New Zealand Government. Various maps.
08.  No paper.
09.  New Zealand Government. [Dr Peter Dillingham's] Review of Edwards *et al.* [implementation of Dillingham's methodology]
10.  No paper.
11.  Dr Michael E. Grigg. Spatial Risk Assessment of Threats to Hector's and Māui dolphins in New Zealand.
12.  New Zealand Government. IWC review of SEFRA modelling - additional information and suggestions from the New Zealand government [to SC/69A/MD/WP/02].
13.  New Zealand Government. IWC review of fisheries information for SEFRA modelling - additional information from the New Zealand government [to SC/69A/MD/WP/03].

# Appendix 1

# Summary of the Parameters and Derived Quantities in the SEFRA Model

| Category | Symbol | Description |
|---|---|---|
| *Indices* | | |
| | i | A fishing event (e.g., a net is set or a trawl tow begins) that occurs at a time and location |
| | s | Population – Hector's dolphin or Māui dolphin |
| | g | Commercial fishery group – set net (SN) or inshore trawl |
| | k | Necropsy type – toxoplasmosis, predation, or other |
| | $l$ | A necropsy event |
| | c | A cell in a map |
| *Data* | | |
| | $\left(c^{live}\right)'$ | Number of observed live captures |
| | $\left(c^{dead}\right)'$ | Number of observed dead captures |
| | $\eta_l$ | Necropsy observations |
| *Covariates* | | |
| | $a'_{gi}, a_{gi}$ | Observed fishing intensity and fishing intensity (km of net for set net and number of tows for inshore trawl) |
| | $p^s_{si}, p^w_{si}$ | Relative density during summer and winter |
| | $O'_{sg}, O_{sg}$ | Observed overlap and overlap |
| *Estimated parameters* | | |
| | $v_g$ | Vulnerability |
| | $p^{obs}_g$ | Probability that an event is observable |
| | $\psi_g$ | Probability of an individual being alive given that it is caught |
| | $N_s$ | Population size for each sub-species |
| | $\rho_k$ | Proportion of non-fishery deaths for each necropsy type |
| *Random variables* | | |
| | $r^{max}_s$ | Intrinsic population growth rate (referred to as $r_{max}$ in this report) |
| | $\omega_g$ | Live release survival rate |
| | $S^{1+}_s$ | Non-calf annual survival rate |
| | $\phi$ | Calibration coefficient |
| | $\partial_k$ | Necropsy detection scalar |
| *Derived quantities* | | |
| | $C^{live}_{sg}, C^{dead}_{sg}$ | Live and dead captures |
| | $D_{sg}$ | Deaths |
| | $PST_s$ | Population Sustainability Threshold |
| | $R_{pg}$ | Risk ratio |
| | $k_g$ | Cryptic mortality |
| *Relevant equations* | | |
| | $r_{max} * T_{opt} = a_{rT}$, | $T_{opt}$ is the associated generation time ("opt" implies "optimal" environmental conditions that allow for the population to grow at a rate of $r_{max}$), and $a_{rT}$ is an allometric constant, also referred to as a demographic invariant (e.g., Niel and Lebreton 2005). |
| | $r_{max} = \ln(\lambda_{max})$ | Maximum growth rate of a population under ideal conditions or intrinsic growth rate. |
| | $R_{max} = \lambda_{max} - 1$ | $R_{max}$ is a theoretical concept represented by a population model parameter that describes the per capita population growth rate at low population size (specifically, at a population level where neither depensatory nor compensatory density dependence is occurring). $R_{max}$ describes the upper limit on the per capita population growth rate in deterministic models, and the average maximum per capita growth rate in stochastic models (IWC 2023). |
| *Other useful definitions* | | |
| | Allometric invariant | Relationship between optimal generation time (the average age of a breeder during optimal growth) and $r_{max}$ observed across vertebrate species. |
| | λ | Rate of increase or annual per capita population growth rate. |

## References

International Whaling Commission. 2023. Report of the Scientific Committee. *J. Cetacean Res. Manage. (Suppl.)* 24:1-190.

Niel, C. and Lebreton, J.D. 2005. Using demographic invariants to detect overharvested bird populations from incomplete data. *Cons. Biol.* 19(3). [Available at: https://doi.org/10.1111/j.1523-1739.2005.00310.x].

# Appendix 2

# Considerations Regarding an Estimate of $r_{max}$ for Hector's and Māui Dolphin

Reviewer: Doug DeMaster

## 1. INTRODUCTION

Many populations of marine mammals are currently depleted (e.g., below 50% of the population size that would exist in the absence of anthropogenic effects). In some cases, the cause of depletion was related to directed takes by humans for the purpose of deriving products of value, such as blubber, fur, oil, meat, etc. In some cases, the cause of depletion was indirect mortality caused by commercial or subsistence fisheries or other human activities. Populations can also become depleted due to epizootics or natural disasters that affect habitat quality or cause mortality.

From a management perspective, it is generally accepted that depleted populations of marine mammals should be allowed to recover to what are considered healthy or optimal levels (although the definition of this threshold varies by country and by international organization). In theory, a severely depleted population that is no longer experiencing mortality associated with anthropogenic activities should recover at a rate, often referred to as the maximum rate of increase ($r_{max}$). This rate of increase is species-specific and is a function of species-specific life history parameters, such as age of first reproduction, interval between births, and age-specific survival.

A working definition of $r_{max}$ was proposed during last year's meeting of the Scientific Committee:

*'$R_{max}$ is a theoretical concept represented by a population model parameter that describes the per capita population growth rate at low population size (specifically, at a population level where neither depensatory nor compensatory density dependence is occurring). $r_{max}$ describes the upper limit on the per capita population growth rate in deterministic models, and the average maximum per capita growth rate in stochastic models.' (developed by the Small Cetacean Working Group for $r_{max}$, April 2022).*

The objective of this document is to provide a summary of approaches that could be used to estimate $r_{max}$ in the management of the Hector's and Māui dolphin, and to provide a recommendation as to which approach or approaches are most appropriate from the perspective of stewardship towards this population. More specifically, this review will pay special attention to the estimation procedure (i.e., the demographic rates used) and application of $r_{max}$ in the Roberts *et al.* (2019) publication.

## 2. APPROACHES TO CONSIDER

Six primary approaches were described by the $r_{max}$ Working Group of the IWC Scientific Committee. Based on published literature (including Roberts *et al.*, 2019), I have collapsed the six approaches described by the Working Group into 4 approaches that can be used to provide a reasonable estimate of $r_{max}$ for a given population (i.e., in this case, a population of Hector's or Māui dolphin):

2.1 A direct estimate of the maximum rate of increase ($R$) of the population based on a series of abundance estimates from a population that is severely depleted, and is recovering in the absence of anthropogenic mortality and the absence of density dependent factors limiting the rate of recovery. Referred to as R1. See (Wade, 1998), Appendix 1) for examples of this approach applied to marine mammal species. These data are not available for the population of Hector's or Māui dolphin.

2.2 A direct estimate of the maximum rate of increase of the population based on estimates of life history parameters (i.e., age-specific survival rates, age-specific reproductive rates) from a population that is severely depleted, and is recovering in the absence of anthropogenic mortality and the absence of density dependent factors limiting the rate of recovery (e.g., use of Lotka's equation or Leslie model). See (Wade, 1998), Appendix 1) for examples of this approach applied to marine mammal species. Referred to as R2. These data are not available for the population of Māui and Hector's dolphin.

2.3 A direct estimate of the maximum rate of increase of the population based on an estimate of $r_{max}$ from a species with similar life history characteristics or partial data from the species of interest and partial data from species believed to have similar life history characteristics (e.g., use of Lotka's equation or a Leslie model). For example, in implementing mandates of the US Marine Mammal Protection Act, the National Marine Fisheries Service uses a default value of 1.040 (which is defined in general as $N(t+1)/N(t)$) for all cetacean species in the regime to govern marine mammal-commercial fishery interactions. See Taylor *et al.* (2000) for the rationale behind this approach. Similarly, the approach described in

Edwards *et al.* (2018) (i.e., invariant method of estimating $r_{max}$) would also be included in this category.

2.4 An indirect estimate of the maximum rate of increase of the population based on output from population models, typically fit with input data on abundance over time, anthropogenic mortality over time, and life history data.

As noted above, combinations of the above approaches are possible. For example, Slooten and Lad (1991) used life history information from a field study of Hector's dolphin, which were considered depleted, in combination with age-specific survival data from other species to estimate a specific maximum rate of increase for the Hector's dolphin, as well as the 'maximum possible' estimate of $r_{max}$ based on survivorship data from human life tables and the highest possible rate of age-specific reproduction for this species. The reported range of $r_{max}$ values was 1.018-1.049. The lower value reported by Slooten and Ladd is frequently referenced by authors who have published results related to the conservation status of the Māui dolphin (see Table 1) as a lower limit to $r_{max}$ for this species.

## 3. ESTIMATES OF $R_{MAX}$ FROM THE LITERATURE

Over the last 30 or more years, concern over the conservation status and the best way to recover Hector's or Māui dolphin populations are discussed in many publications. Some of the publications most often referenced regarding the impacts of anthropogenic removals of Māui dolphins are provided in Table 1. Table 2 presents a list of critiques of several papers listed in Table 1. Table 3 presents a summary of results from Moore (2015), relating adult survival, $r_{max}$ (continues estimate of maximum rate of increase), and generation time (both optimal generation time from a low status relative to *K*, and generation time from a population status close to *K*). Adjunct 1 presents a summary of a subset of combinations of life history parameters and associated $r_{max}$ values from either the published literature or from application of a deterministic Lotka's equation.

Table 1

Summary of estimates of $r_{max}$ for the Hector's/Māui dolphin or other small cetacean species from the published literature.

| Author(s) | Estimate of $r_{max}$ | Notes |
|---|---|---|
| Slooten and Lad (1991) | Best Estimate: 1.018; Maximum Value: 1.049 | Combination of R2 and R3 |
| Martien *et al.* (1999) | 1.044 | From Slooten and Ladd |
| Davies *et al.* (2008) | Maximum value of $r_{max}$=1.02 | R3 |
| Burkhart and Slooten (2003) | Range – 1.018, 1.023, 1.044, 1.049 | |
| Slooten and Davies (2011) | Bayesian posterior: 1.016-1.052 <br> 3 values used: 1.016, 1.03, 1.04 | Combination of R2 and R3 |
| Lebreton (2012) | 1.018 (from Slooten and Ladd) | |
| Hamner *et al.* (2014) and Currey *et al.* (2012) | 1.018 (from Slooten and Ladd) and 1.04 (default value) | |
| Edwards *et al.* (2018) | N(0.0449, 0.0107) | Combination of R2 and R3 - Based on inter-specific life history data |
| Roberts *et al.* (2019) | From Edwards *et al.* (2018) Priors - Hector's N(0.05, 0.011²), Māui (0.045, 0.011²). Posterior | R4. Used updated age data and survival data, and Edwards *et al.* (2018). Max age of breeding increased to 30 (from 20). |
| Cooke *et al.* (2019) | Base case – $r_{max}$ was a free parameter, being determined by other parameters in the model. <br> Other cases – from Edwards *et al.* (2018). | R4 |
| Parra *et al.* (2021) | For common dolphins in waters off Australia - 1.02 and 1.04 | R2 |
| Mannocci *et al.* (2012) | For common dolphins in eastern North Atlantic – 1.045 (<u>+ 0.009</u>) | R2 |
| Lowry *et al.* (2008) | For Bristol Bay beluga whale – 1.048 | R1 |
| Reilly and Barlow (1986) | Theoretical maximum for odontocete – 1.06. Eastern spinner dolphin – 1.02 | R2 |
| Wade (1998) | Lack of evidence for $r_{max}$ for odontocete>1.04. Worst case for odontocete is 1.02 or lower. | Combination |
| Moore (2015) | Range of $r_{max}$ values for 5 odontocete species – 1.031-1.071 | R3 |

Table 2

Published critiques of published estimates of $r_{max}$ for Hector's/Māui dolphin.

| Critique author(s) | Author(s) being critiqued | Notes |
|---|---|---|
| Slooten and Dawson (2021) | Roberts *et al.* (2019) and Edwards *et al.* (2018) | |
| Cooke *et al.* (2019) | Roberts *et al.* (2019) | Used different cases to investigate different approaches reported in Roberts *et al.* (2019) |
| Taylor *et al.* (2018) | Roberts *et al.* (2019) | |

Table 3

Parameter estimates from Moore (2015): mean, and sd (in parentheses).

| Species | $S_{max}$ | $r_{max}$ | $T_{opt}$ | $T_0$ |
|---|---|---|---|---|
| *O. brevirostris* | 0.96 (0.01) | 0.038 (0.01) | 21 (2.7) | 20 (1.9) |
| *O. heinsohni* | 0.96 (0.01) | 0.037 (0.01) | 21 (2.7) | 20 (1.7) |
| *N. asiaeorientalis* | 0.93 (0.01) | 0.071 (0.02) | 13 (1.2) | 15 (0.9) |
| *S. chinensis* | 0.97 (0.01) | 0.032 (0.01) | 26 (1.1) | 25 (1.2) |
| *S. plumbea* | 0.97 (0.01) | 0.031 (0.01) | 26 (1.8) | 25 (1.0) |

## 4. SPECIFIC COMMENTS ON THE ESTIMATE OF $R_{MAX}$ AND ITS USE REPORTED IN ROBERTS *ET AL.* (2019)

From Roberts *et al.* (2019): Appendix 2, p.93:

*'The revised spatial risk assessment of threats to Hector's/Māui dolphins used individual size at age and maturity stage information to derive an $r_{max}$ for the species (Edwards et al., 2018). This assessment followed the method proposed by Dillingham et al. (2016) and implemented by Moore (2015), which uses an allometric invariant between optimal generation time (the average age of a breeder during optimal growth) and $r_{max}$ observed across vertebrate species. See Edwards et al. (2018) for a detailed description of methods and sensitivity runs.*

*'The assessment by Edwards et al. (2018) obtained base case Monte Carlo distributions of age at maturity of 6.91 (95% CI = 5.82-8.24) and $r_{max}$ of 0.050 (95% CI=0.029-0.071) for Hector's dolphin. This updated the previous base case $r_{max}$ of 0.018 assumed by the most recent Māui dolphin multi-threat assessment (Currey et al., 2012), although this was based on a maximum longevity of 20 (Slooten and Lad, 1991), which is now known to be an underestimate for this species (Gormley, 2009).*

*'Here, the assessment by Edwards et al. (2018) was updated with the supplementary ageing and maturity information detailed in Appendix 1 (new data are displayed in Tables A1-1 and A1-2).*

*'In addition, a sensitivity model run was undertaken, in which the sensitivity of the $r_{max}$ posterior to assuming a maximum breeding age of 30 was assessed (previously this was infinite).'*

And from p.94:

*'Update using supplementary age and maturity information*

*'Updating the assessment by Edwards et al. (2018) with supplementary ageing produced an identical median and 95% CI for $r_{max}$ to three decimal places, i.e. 0.050 (95% CI=0.02-0.071). The updated posterior is displayed in Figure A2-1.'*

In Appendix 3 (Roberts *et al.*, 2019), the authors recommend use of a lower value for $r_{max}$ for the Māui dolphin population due to 'low population size', including demographic stochasticity, inbreeding depression, and social Allee mechanisms. The mean estimate of $r_{max}$ for the Māui dolphin population was 1.046.

From Roberts *et al.* (2019), p.31) the following distributions for $r_{max}$ were used:

*'Priors for the intrinsic population growth rate ($rs_{max}$) random variable were developed by fitting normal distributions to posterior samples produced by the analysis of Edwards et al. (2018), updated in Appendix 2 (for Hector's dolphin) and from an analysis adjusting for small population size in Appendix 3 (for Māui dolphin). The prior for Hector's dolphin is simulated from:*

$$rs=HDO_{max}\sim normal(0.050, 0.0112) \text{ and for Māui dolphin } rs=HDM_{max}\sim normal (0.045, 0.0112).'$$

The mean life history parameters that resulted in an $r_{max}$ estimate of 1.051 for Hector's dolphin (or $r_{max}$ of 0.05 in notation of Roberts *et al.* (2019)) or an $r_{max}$ estimate of 1.046 for Māui dolphin (or $r_{max}$ of 0.045) are shown in Appendix 1 of this report. The $r_{max}$ value in Edwards *et al.* (2018): $r_{max}$ =1.055) uses a mean age of first reproduction of 6.9 years, mean non-calf survival rate of 0.95, mean calf survival rate equal to the square of non-calf survival, a mean birth interval of 2 years on average, and a maximum age of 26 years. The $r_{max}$ value for Hector's dolphin in Roberts *et al.* (2019): $r_{max}$ = 1.051) uses a mean age of first reproduction of 8 years, mean non-calf survival rate of 0.958, mean calf survival rate of 0.795, a mean birth interval of approximately 2.2 years, and a maximum age of 30 years.

It should be noted that use of the allometric 'invariant' method, as part of the overall SEFRA approach, truncated certain combinations of life history parameters and associate estimates of $r_{max}$, that produced estimates of generation time inconsistent with the invariant method. For example, use of the mean life history values in Roberts *et al.* (2019) without the truncation step would result in a point estimate of $r_{max}$ of approximately 1.033. With the truncation step, that same value becomes approximately 1.05. This is a significant difference in $r_{max}$ values, and the underlying merits of using a relationship among a wide variety of mammalian taxa to truncate life history data from Hector's dolphin or closely related species needs further exploration. As presented in the paper by Roberts *et al.*, and associated papers, there was insufficient information to adequately support using allometric data to truncate the results from life history-based estimates of $r_{max}$ for Hector's dolphin or closely related taxa.

Given the lack of availability of life history data for Hector's or Māui dolphins, it is not possible to evaluate the degree to which existing $r_{max}$ estimates are credible or appropriate for use in this management context (i.e., use in SEFRA protocol). To a large extent, the robustness of estimates and the associated uncertainty in estimates of $r_{max}$ from an allometric analysis that uses life history data from other mammalian taxa needs to be more fully evaluated. Ideally, a suite of simulations would be performed to address whether the performance of a management approach based on allometric derived estimates of $r_{max}$ would be carried out to determine the likelihood of such an approach providing adequate management outcomes. It can be said that the estimates of $r_{max}$ used in Roberts *et al.* (2019) for assessing the risk of anthropogenic interactions to Hector's and Māui dolphin populations are the highest values proposed in the literature for this application.

## 5. RECOMMENDATIONS

There is no clear rationale for selecting a specific $r_{max}$ value from those values listed in Table 1 for either Hector's or Māui dolphin populations. However, there are advantages and disadvantages in using a given $r_{max}$ for the purpose of evaluating the status of Hector's and Māui dolphin populations and in choosing a way forward regarding stewardship options. The following $r_{max}$ values are considered: (1) the most conservative value for $r_{max}$ is 1.018 (from Slooten and Lad (1991; (2) a default value for $r_{max}$ — 1.04, which was used for cetaceans in implementing a regime to govern marine mammal-commercial fishery interactions in US waters; (3) $r_{max}$ value for other odontocetes ($R_{max}$ of 1.045; Mannocci *et al.*, 2012) or 1.048 (Lowry *et al.*, 2008); (4) maximum value from combination of R2 and R3 approach — 1.049 (Slooten and Lad, 1991); and (5) $r_{max}$ value from Edwards *et al.* (2018), which formed the basis for the $r_{max}$ value used in Roberts *et al.* (2019) of 1.051 for Hector's dolphin and 1.046 for Māui dolphin.

The most conservative approach (i.e., $r_{max}$ value of 1.018 from Slooten and Lad) would provide for the greatest likelihood that management actions would be effective in recovering this population, assuming anthropogenic removals were driving the dynamics of these populations. However, this approach would also have the greatest adverse impact on human activities that overlap with the range of Māui dolphins. Further, several authors have noted that this estimate of $r_{max}$ may be negatively biased (see Peter Dillingham report, pers. comm, Edwards *et al.*, 2018, and Roberts *et al.*, 2019). Using an $r_{max}$ of 1.018 requires assumptions along the lines made by Slooten and Lad concerning the life history of the species (Hector's dolphin), as well as status of the Hector's dolphin population at the time the life history data were collected. Regarding the latter point, use of the $R^2$ method for estimating $r_{max}$ for species -specific life history parameters requires the assumption that the status of the population was well below its carrying capacity at the time the life history data were collected (or that the value of a specific life history parameter is independent of density). Information on the feasibility of the carrying capacity for this population being 10 times greater than the population size at the time the research was conducted would be extremely useful in this decision process (note: an estimated growth rate of 1.018 for a population at 10% of its carrying capacity, assuming linear density dependence, would correspond to an $r_{max}$ estimate of 1.02).

The merits of using an $r_{max}$ value of 1.04 are described in Appendix 1 of Wade (1998) and in Taylor *et al.* (2000). Wade (1998) noted that a lack of evidence of higher rates suggests that 4% is probably a suitable default value for odontocetes and that 2% represents a reasonable worst-case scenario. However, some caution is required, as so few data exist on observed rates of increase of odontocetes. Also, although several odontocete populations have apparently declined from human-caused mortality, none have been observed to recover. Although this may be due to the difficulty in monitoring odontocete populations, it also suggests that maximum rates of increase for some odontocetes could be even lower than 2%. Similar concerns regarding $r_{max}$ values for the Māui dolphin being above 1.02 are raised by Cooke *et al.* (2019) and others.

The merits of using an $r_{max}$ value for Hector's or Māui dolphin populations based on published $r_{max}$ values for other species of odontocetes is difficult to evaluate absent additional information on the life history of Hector's or Māui dolphin. Estimates of $r_{max}$ for a beluga whale population in Alaska (i.e., 1.048 from Lowry *et al.*, 2008) or common dolphin population from the eastern Atlantic (i.e., 1.045 from Mannocci *et al.*, 2012) may or may not be representative of an $r_{max}$ value for Hector's or Māui dolphin populations. Based on available life history data for Hector's dolphin, it appears these values could serve as some sort of upper limit.

Using an $r_{max}$ value of 1.049 or higher is not recommended by this author for an $r_{max}$ value for evaluating the risk posed by anthropogenic activities to Hector's or Māui dolphin populations. As noted in Slooten and Lad (1991), such a value requires aggressive estimates of age-specific reproductive and survival rates or the truncation of estimates of $r_{max}$ based on allometric analyses that place limits on generation time. Not using such a high value for the rate of increase for the Māui dolphin subspecies is supported by the following rationale provided by Cooke *et al.* (2019):

> '*However, there are reasons to doubt whether a high r0 is likely for Māui's dolphin. The value of $r_0$ is a function not only of the species but also of its habitat. In principle, the edge of the natural range of a species is defined by the $r_0 = 0$ contour (Caughley et al., 1988). Thus, populations near the edge of the range tend to have a lower $r_0$, and lower population density at K (carrying capacity) than populations within the core range. The actual picture can be more complex, because of environmental variability and the movement of animals, but the basic principle applies. Māui dolphins are the northernmost extant population of C.*

*hectori, in a location where the abundance of suitable fish prey is estimated to be up to an order of magnitude lower than occurs in the core range for Hector's dolphins (Roberts et al., 2019). If this represents the extreme of the natural range of the species, then $r_0$ for Māui is likely to be below the average for populations of this species. Consequently, as a sensitivity test, runs were also conducted with $r_0$ set to 0.02. This particular value is arbitrary, but is close to the value of 0.018 estimated by Slooten and Lad (1991).'*

Based on discussions during the pre-meeting which focused on a review of the SEFRA protocol and the Roberts *et al.* (2019) paper, a clear recommendation for the authors of the Roberts *et al.* paper would be to run the spatial habitat model using fixed values for $r_{max}$ (i.e., 1.03, 1.04 and 1.05). Absent such an analysis, it is very difficult to tease apart the results of integrating uncertainty associated with the $r_{max}$ estimate from the interactions among the various elements of the SEFRA protocol.

## CONCLUSION

The following points are key in recommending an $r_{max}$ value between 1.03 and 1.04 for the purpose of evaluating conservation strategies to recover populations of Hector's or Māui dolphin.

(1) There are no suitable time series data for abundance for populations of Hector's or Māui dolphin to support an R1 approach.

(2) There are no adequate life history data (i.e., age-specific reproduction and age-specific survival) for populations of Hector's or Māui dolphin to support an R2 approach.

(3) an $r_{max}$ value of 1.018 presumes the limited life history data were collected from a severely depleted population with a very small carrying capacity, which may not be the case. That is, an $r_{max}$ value estimated at 1.018, if within 10% of the true value, would require the population from which the life history data were collected to be something like 10% of its carrying capacity. From the available literature on Hector's dolphin, that does not appear to be the case.

(4) $R_{max}$ values of between 1.03 and 1.04 are plausible given the life history of this species and other odontocete species (Moore, 2015). It should be noted that the one example of an $r_{max}$ greater than 1.05 in Moore (2015) was for a species that has an interbirth interval less than 2 years. Such a life history is not consistent with the available data on interbirth interval for Hector's dolphin and should not be used in evaluating $r_{max}$ values appropriate for use in managing Hector's and Māui dolphin.

(5) an $r_{max}$ value of 1.04 is near the middle of the range of $r_{max}$ values for other species of odontocetes.

(6) an $r_{max}$ value as high as 1.04, if used in setting management practices along the lines described by Wade (1998) and Taylor *et al.* (2000), would still provide for adequately precautionary management.

### Literature Cited

Burkhart, S.M. and Slooten, E. 2003. Population viability analysis for Hector's dolphin (*Cephalorhynchus hectori*): a stochastic population model for local populations. *New Zealand Journal of Marine and Freshwater Research* 37(3): 553-66.

Caughley, G., Grice, D., Barker, R. and Brown, B. 1988. The edge of the range. *J. Animal Ecology* 57(3): 771-85.

Cooke, J., Constantine, R., Hamner, R.M., Steel, D. and Baker, C.S. 2019. Population dynamic modeling of the Māui dolphin based on genotype capture-recapture with projections involving bycatch and disease risk. New Zealand Aquatic Environment and Biodiversity Report No. 216. 38 pp. [Available at: *www.mpi.govt.nz/news-and-resources/publications*].

Currey, R., Boren, L., Sharp, B.R. and Peterson, D. 2012. *A risk assessment of threats to Maui's dolphins*. Ministry for Primary Industries and Department of Conservation, New Zealand. 51pp.

Davies, N.M., Bian, R., Starr, P., Lallemand, P., Gilbert, D.A. and McKenzie, J. 2008. *Risk analysis for Hector's dolphin and Maui's dolphin subpopulations to commercial set net fishing using a temporal-spatial age-structured model*. Ministry of Fisheries, Wellington, New Zealand. [Available at: *www.fish.govt.nz/en-nz/Consultations/Hector+new/default.htm*].

Dillingham, P.W., Moore, J.E., Fletcher, D., Cortés, E., Curtis, K.A., James, K.C. and Lewison, R.L. 2016. Improved estimation of intrinsic growth rmax for long-lived species: integrating matrix models and allometry. *Ecol. Appl.* 26(1): 322-33

Edwards, C.T.T., Roberts, J.O. and Doonan, I.J. 2018. Estimation of the maximum rate of intrinsic growth for Hector's dolphin. Final Research Report available from Fisheries New Zealand. 22pp.

Gormley, A.M. 2009. Population modelling of Hector's dolphins. PhD thesis, University of Otago, Dunedin, New Zealand. [Available at: *www.otago.ac.nz/*].

Hamner, R.M., Wade, P., Oremus, M., Stanley, M., Brown, P., Constantine, R. and Baker, C.S. 2014. Critically low abundance and limits to human-related mortality for the Maui's dolphin. *Endangered Species Research* 26: 87-92.

Lebreton, L.C.M. 2012. A quantitative analysis of fishing management proposals based on the risk assessment of threats to Māui's dolphins. Unpublished paper. 13pp.

Lowry, L.F., Frost, K.J., Zerbini, A., DeMaster, D. and Reeves, R.R. 2008. Trend in aerial counts of beluga or white whales (*Delphinapterus leucas*) in Bristol Bay, Alaska, 1993-2005. *J. Cetacean Res. Manage.* 10(3): 201-08.

Mannocci, L., Dabin, W., Augeraud-Véron, E., Dupuy, J.F., Barbraud, C. and Ridoux, V. 2012. Assessing the impact of bycatch on dolphin populations: the case of the common dolphin in the eastern North Atlantic. *PLoS ONE* 7(2): e32615.

Martien, K.K., Taylor, B.L., Slooten, E. and Dawson, S.M. 1999. A sensitivity analysis to guide research and management for Hector's dolphin. *Biol. Cons.* 90: 183-91.

Moore, J.E. 2015. Intrinsic growth (rmax) and generation time (T) estimates for the genera *Sousa*, *Orcaella* and *Neophocaena*. *NOAA Tech. Mem.* NOAA-TM-NMFS-SWFSC-550: 14pp.

Parra, G.J., Bilgmann, K., Peters, K.J. and Möller, L.M. 2021. Abundance and potential biological removal of common dolphins subject to fishery impacts in south Australian waters. *Front. Mar. Sci.*: 1237.

Reilly, S.B. and Barlow, J. 1986. Rates of increase in dolphin population size. *Fish. Bull.* 84(3): 527-33.

Roberts, J.O., Webber, D.N., Roe, W.D., Edwards, C.T.T. and Doonan, I.J. 2019. Spatial risk assessment of threats to Hector's and Māui dolphins (*Cephalorhynchus hectori*). *New Zealand Aquatic Environment and Biodiversity Report No.* 214: 175pp.

Slooten, E. and Davies, N. 2011. Hector's dolphin risk assessments: old and new analyses show consistent results. *J. Roy. Soc. New Zealand* 42: 49-60.

Slooten, E. and Dawson, S.M. 2021. Hector's and Māui dolphin bycatch following 2020 management decisions. 25pp. Paper SC/68C/HIM/17 presented to the IWC Scientific Committee, Virtual Meetings, April-May 2021 (unpublished). 25pp. [Paper available from the Office of this Journal].

Slooten, E. and Lad, F. 1991. Population biology and conservation of Hector's dolphin. *Canadian Journal of Zoology - Revue Canadienne de Zoologie* 69: 1701-07.

Taylor, B., Lonergan, M. and Reeves, R. 2018. Panel Comments and Recommendations. Appendix 1 – Hector's and Māui Dolphin Threat Management Plan Review, Risk Assessment Workshop, 9-13 July 2018. 15pp. [Available at: *https://www.doc.govt.nz/globalassets/documents/conservation/native-animals/marine-mammals/maui-tmp/hectors-risk-assessment-workshop-panel-recommendations-appendix-1.pdf*].

Taylor, B.L., Wade, P.R., DeMaster, D.P. and Barlow, J. 2000. Incorporating uncertainty into management models for marine mammals. *Cons. Biol.* 14(5): 1243-52.

Wade, P.R. 1998. Calculating limits to the allowable human-caused mortality of cetaceans and pinnipeds. *Mar. Mamm. Sci.* 14(1): 1-37.

## Adjunct 1

### Summary of a range of life history parameters and associated $r_{max}$ values for Hector's/Māui dolphin populations, based on deterministic application of Lotka's equation or published literature

| Reference | Age at first birth | Calf survival | 1+ survival | Max age | Fecundity | $r_{max}$ |
|---|---|---|---|---|---|---|
| Gormley (2009) | 8.55 | 0.65 | 0.917 | 26 | 0.205 | 0.99 |
| Roberts *et al.* (2019) | 8 | 0.795 | 0.958 | 30 | 0.225 | 1.051 |
| Edwards *et al.* (2018) | 6.9 | 0.9025 | 0.95 | 26 | 0.2375 | 1.055 |
| | 8 | 0.9025 | 0.95 | 23 | 0.2375 | 1.051/1.054* |
| | 8 | 0.8649 | 0.93 | 23 | 0.2325 | 1.027/1.029* |
| | 8 | 0.83 | 0.923 | 26 | 0.225 | 1.033 |
| | 8 | 0.8649 | 0.93 | 23 | 0.186 | 1.01/1.016* |
| | 8 | 0.93 | 0.93 | 23 | 0.186 | 1.015/1.02* |

/*- assumes longevity is infinite.

### References

Edwards, C.T.T., Roberts, J.O. and Doonan, I.J. 2018. Estimation of the maximum rate of intrinsic growth for Hector's dolphin. Final Research Report available from Fisheries New Zealand. 22pp.

Gormley, A.M. 2009. Population modelling of Hector's dolphins. PhD thesis, University of Otago, Dunedin, New Zealand. [Available at: *www.otago.ac.nz/*].

Roberts, J.O., Webber, D.N., Roe, W.D., Edwards, C.T.T. and Doonan, I.J. 2019. Spatial risk assessment of threats to Hector's and Māui dolphins (*Cephalorhynchus hectori*). *New Zealand Aquatic Environment and Biodiversity Report No.* 214: 175pp.

# Appendix 3

# Comment on Item 2.1 $r_{max}$

J.G. Cooke

The value of $r_{max}$ = 0.05 for Hector's (South Island) dolphins used in Roberts *et al.* (2019) is based on the Dillingham *et al.* (2016) review of $r_{max}$ for long-lived species. The value of 0.045 for Māui dolphins is the Hector's value with a correction for small population size. The small-population-size correction covers the effect of demographic stochasticity and possible inbreeding depression, but does not account for ecological or other factors that cause the population to be small.

The Dillingham *et al.* review is about $r_{max}$ for species, which is not synonymous with $r_{max}$ for populations. Deterministically speaking, the limits of the natural range of a population are defined by the contour $r_{max}$ = 0 (Caughley *et al.*, 1988). Populations inside the contour will persist, while any populations that may sometimes be founded outside the contour will eventually disappear. Populations near the borderline of persistence will have lower $r_{max}$ than robust populations in the most favourable parts of the range.

The situation is slightly more complex in the presence of realistic levels of environmental variability. Populations in the core range will have high $r_{max}$, constrained by the biology of the species, and high persistence, while populations on the periphery of the range will have lower average $r_{max}$ and lower persistence.

A question is whether the North Island population of *C. hectori* is a core population with $r_{max}$ and density similar to the maximum for the species, or a marginal population with lower $r_{max}$ and density. in other words, is the species naturally rare in North Island waters or is it a remnant of a much larger population that once occurred in densities comparable to South Island populations?

A combination of low $K$ (carrying capacity) and high $r_{max}$ for the North Island population is ecologically implausible, because it would require that the presence of a relatively small number of individuals would have sufficient impact on the ecosystem to depress the growth rate from the high $r_{max}$ value to zero.

Under a combination of high $r_{max}$ and historically high $K$ for the North Island population, the vulnerability to fishing would need to be considerably higher than that estimated by Roberts *et al.* in order to have depleted the population to its current low level (Cooke *et al.*, 2019; Cooke *et al.*, 2018).

Either way, an $r_{max}$ for Māuis similar to that for Hector's is hard to reconcile with the available information.

Indeed, it is confusing to use the term $r_{max}$ for both species and populations. It would be better to distinguish, as suggested, for example, by Cooke *et al.* (2019), between $r_0$, the mean growth rate of a population at low population size, which will differ between core and marginal populations, and $r_{max}$, the maximum growth rate of a species under good conditions.

**References**

Caughley, G., Grice, D., Barker, R. and Brown, B. 1988. The edge of the range. *J. Animal Ecology* 57(3): 771-85.

Cooke, J., Constantine, R., Hamner, R.M., Steel, D. and Baker, C.S. 2019. Population dynamic modeling of the Māui dolphin based on genotype capture-recapture with projections involving bycatch and disease risk. New Zealand Aquatic Environment and Biodiversity Report No. 216. 38 pp. [Available at: *www.mpi.govt.nz/news-and-resources/publications*].

Cooke, J.G., Steel, D., Hamner, R.M., Constantine, R. and Baker, C.S. 2018. Population estimates and projections of Maui dolphin (*Cephalorhynchus hectori maui*) based on genotype capture-recapture, with implications for management of mortality risk. Paper SC/67b/ASI05 presented to the IWC Scientific Committee, April-May 2018, Bled, Slovenia (unpublished). 16pp. [Paper available from the Office of this Journal].

Dillingham, P.W., Moore, J.E., Fletcher, D., Cortés, E., Curtis, K.A., James, K.C. and Lewison, R.L. 2016. Improved estimation of intrinsic growth rmax for long-lived species: integrating matrix models and allometry. *Ecol. Appl.* 26(1): 322-33

Roberts, J.O., Webber, D.N., Roe, W.D., Edwards, C.T.T. and Doonan, I.J. 2019. Spatial risk assessment of threats to Hector's and Māui dolphins (*Cephalorhynchus hectori*). *New Zealand Aquatic Environment and Biodiversity Report No.* 214: 175pp.

# Appendix 4

# Review of Spatial Risk Assessment for Hector's and Māui Dolphins in New Zealand: Spatial Distribution

Reviewer: Megan C. Ferguson, PhD

*Note: please see the Additional Files for a PDF summary of a PowerPoint presentation to accompany this Appendix.*

## BACKGROUND

In 2019, the IWC Scientific Committee (SC) recommended a detailed independent review of a 'Spatial risk assessment of threats to Hector's and Māui dolphins' (Roberts *et al.*, 2019) with respect to this document's use for informing management measures to address anthropogenic threats to Hector's and Māui dolphins off New Zealand. In 2021, five topics were identified for the risk assessment review, to be considered by independent experts.

This review of Roberts *et al.* (2019) focuses on the density surface models (DSMs) for Hector's and Māui dolphins and the spatial methods used to infer risk to the dolphins by commercial fishing. The IWC SC does not have a strict checklist for evaluating spatially-explicit models. However, my overarching goals for this review are to: (1) identify whether best practices were used in Roberts *et al.* (2019) dolphin and bycatch spatial analyses and to provide references and resources for cases in which alternate preferred methods are available; and (2) to evaluate whether the authors' conclusions are supported by the results of their spatial models and by other available information about the ecology and fisheries that is relevant to the issue of Hector's and Māui dolphin management in New Zealand. Two themes that I repeatedly address in this review are: (1) consideration of spatial and temporal scales (resolution and extent) in the data, model construction, and inference; and (2) identification and evaluation of uncertainty, including both precision and bias. This review is organized into eight sections, corresponding to the eight factors that the Steering Group identified as being of particular interest.

(a) Choice of habitat variables, including variables not considered in Roberts *et al.* (2019) (including temporally dynamic variables, seasonal model relationships, etc.).
(b) Choice of dolphin occurrence data for fitting the model.
(c) Model selection criteria and fitting.
(d) Combination of models for merging coastal and harbour predictions (Māui model only).
(e) Model validation and interpretation of results, including a review of model estimates at the scale of small or hypothesised local dolphin populations.
(f) Reviewing model predictions of spatially resolved bycatch compared to known records including beach cast carcasses and fisher-reported catches from vessels without observers.
(g) Comparing estimates of commercial fisheries deaths from the spatially explicit model with comparable estimates from simpler models, including uncertainty.
(h) Review of estimated patterns of changing fisheries risk over time.

Note that in my review, I use the term 'subpopulation' to mean a 'small or hypothesized local dolphin population', as referenced in item (e) above.

Overall, Roberts *et al.* (2019) spatial risk assessment of threats to Hector's and Māui dolphins represents an enormous amount of thought, effort, and skill. Throughout this review, I have done my best to distinguish among four types of comments: (1) methodological details that I encourage the authors to consider in future density surface model (DSM) analyses; (2) questions about the analyses or interpretation of results that would assist in a thorough evaluation of the spatial risk assessment (Adjunct 1); (3) *recommendations* that would improve the statistical methodology used in the risk assessment but might not considerably change the results that are relevant for making management decisions (Adjunct 2); and (4) *strong recommendations* that may critically change the results that are relevant for making management decisions (Adjunct 3). A complete list of literature cited in this review is provided in the Literature Cited section at the end of the review.

## A. CHOICE OF HABITAT VARIABLES

### A1. Prey

It is reasonable to expect that the distributions of Hector's and Māui dolphins in New Zealand are related to the distribution and availability of their prey because they are relatively small mammals living in a cold water environment. Available information suggests that the diet of Hector's dolphins varies among subpopulations and seasonally (Miller *et al.*, 2013; Miller, 2014).

Roberts *et al.* (2019) considered the predicted occurrence (probability of presence) of six known demersal prey species as potential predictor variables in both the coastal and the North Island harbor dolphin DSMs. The methods used to derive species-specific estimates of occurrence are presented in Appendix 5 (beginning on p.98 of the report). The prey data came from inshore trawl surveys. Roberts *et al.* (2019) do not explicitly state the season(s) in which the trawl surveys occurred. I consulted Stevenson and MacGibbon (2018), who reported that the surveys in 2017 occurred from 24 March to 17 April. Therefore, I assumed that the remainder of the prey data that Roberts *et al.* (2019) used to derive their prey occurrence models were also collected during the same season, which corresponds to 'summer' (1 November-31 April) in the risk assessment. First, I address the methods used to construct the prey occurrence models. Then I address the potential utility of the prey occurrence predictions in predicting dolphin density throughout the waters of New Zealand.

A spatially-explicit generalized additive model (GAM) is a reasonable tool to use to predict species-specific prey occurrence from the trawl data. I have four concerns about the prey model formulation. First, it is unclear why the authors chose to model species presence/absence instead of the abundance or biomass of each species. Abundance and biomass are likely to be more ecologically meaningful as an index of prey availability because it is reasonable to expect that the dolphins target higher density prey patches. Second, the GAM for each prey species incorporated a bivariate smoothing spline for the location (easting and northing) of the trawl station. Based on the methods presented in Appendix 5 (Table A5-3, p.101), it appears that the authors considered only a thin plate regression spline (TPRS) with shrinkage for this bivariate term. These are isotropic smoothing splines, meaning that they assume that the same level of smoothing is appropriate in all directions. However, if there is a stronger gradient in prey distribution (occurrence, in this case) across isobaths (the onshore/offshore axis) than alongshore, then an isotropic smoother will not be able to capture this spatial characteristic. For completeness, it would be best to also evaluate the performance of an anisotropic smoother, such as a tensor product smooth. Third, the coastlines of New Zealand's North and South islands are complex, with many bays, peninsulas, points, and other nonlinear features. Conventional smoothing splines do not always perform well in complex domains such as this because they may smooth over land barriers and inaccurately predict the same probability of occurrence on both sides of a peninsula, for example, when in fact the coastal habitats on either side of the peninsula are very different. Soap film smoothers (Wood *et al.*, 2008) do not smooth across barriers and they are available in the mgcv package. It would be insightful to compare the performance of the conventional smoothing spline used in the analysis with a soap film smoother or similar barrier-respecting smoother. Lastly, the Q-Q plots for red cod and sprat shown in fig. A5-2 (p.101) suggest a lack of fit in the right tail. It would be good to evaluate an alternative link function, like a complementary log-log link.

Roberts *et al.* (2019) considered only demersal prey species and, presumably, only for summer. Miller (2014) found that demersal prey species contributed the most to Hector's dolphin diets on the east coast of the South Island during spring and summer, but stable isotope results suggested that epipelagic prey had a greater long-term contribution. Thus, in addition to the temporal bias in the prey predictor variables that Roberts *et al.* (2019) used in the dolphin DSMs, there is also an ecological bias related to the depth distribution and other characteristics of the prey. For future modeling efforts, I suggest incorporating epipelagic prey species into investigations of the associations between dolphin density and prey biomass, abundance, and species richness.

### A2. Temporal Criteria for Predictor Variables

There was inconsistency in the application of criteria for determining whether data and derived variables used as dynamic predictor variables in the dolphin DSMs were valid. On the one hand, the authors justified the use of a habitat-based approach over a purely spatial DSM due to the temporally dynamic nature of environmental variables such as turbidity, which were believed likely to be associated with dolphin habitat. The coastal dolphin DSM was fitted to satellite-derived variables (turbidity, chlorophyll-a, sea surface temperature, and downwelling light attenuation) from the specific year and month during which the respective dolphin sighting data were collected (Roberts *et al.*, 2019 p.60). However, the prey variables used in all dolphin DSMs represented only the summer season. Additionally, the coastal dolphin DSMs were constructed using only line-transect effort and sightings located on transect segments for which the full suite of predictor variables were available. This resulted in considerable omission of data for the south coast of the South Island simply because predictions of prey species presence were unavailable, regardless of season: '*For SCSI surveys, this resulted in the majority of positive survey line segments being omitted (70.5 to 73.0%) due to unavailability of prey species presence*' (p.29). This leads me to three questions for the authors:

**Q1:** What information (i.e., predictive value) do temporally mis-matched prey occurrence predictions bring to the dolphin DSMs?

**Q2:** Why omit dolphin sightings and effort from line-transect survey segments for which prey occurrence data were missing, if the prey occurrence data were from a mismatched season and the species is known to exhibit seasonal geographic shifts in distribution?

**Q3:** Did you consider incorporating prey occurrence only into the models for the season (summer) during which the prey data were collected?

### A3. Other Potentially Useful Variables
For the temporally dynamic predictor variables that are available at the monthly resolution year-round, Roberts *et al.* (2019) relied solely on ocean colour data.

**Q4:** For the waters around New Zealand, are remotely sensed data available for wind direction and strength, upwelling intensity, ocean surface fronts, eddy kinetic energy, and anything else that might indicate mesoscale oceanographic activity that may affect prey availability and, hence, dolphin density?

### A4. Mismatch in the Temporal Scale of Habitat Variables Used in Model-building vs. Predicting
Roberts *et al.* (2019) used contemporaneous values of temporally dynamic predictor variables during the coastal dolphin DSM construction phase, but then used climatologically averaged values of those predictor variables to derive spatially-explicit estimates of dolphin density:

> *'For the sea surface variables (temperature, turbidity, chlorophyll a concentration and KPAR), location-specific values were extracted for the year and month corresponding to the date of each aerial survey transect (i.e. in the fitting stage Hector's dolphin abundance estimates were related to habitat variables at the time of the survey). Subsequently in the prediction stage, dolphin densities were predicted using seasonally-averaged conditions over the period from 2009-18.' (Roberts et al., 2019 p.15)*

Mannocci *et al.* (2017) note that the choice between contemporaneous vs. climatological covariates in density surface models for highly mobile marine animals should be based on whether interannual variability or seasonal variability dominates the ecosystem:

> *'We posit that animal associations with fine-scale and ephemeral features are best modelled with instantaneous covariates. Associations with large scale and persistent oceanographic features are best modelled with climatological covariates. Associations with mesoscale features are best modelled with instantaneous or contemporaneous covariates if ephemeral processes are present or interannual variability occurs, and climatological covariates if seasonal processes dominate and interannual variability is weak.' (Mannocci et al., 2017, pp.1098-1099)*

There are a number of concerns associated with this mismatch in temporal resolution during the model-building and predicting stages. First, climatological averages dampen extreme or outlying events. If the ecosystem is dominated by interannual variability, and if unusually high values of a given environmental predictor variable (e.g., upwelling intensity) during a particular point in time generates an extremely productive dolphin foraging ground, which in turn attracts an unusually high density of dolphins, this high-density event will not appear in DSM predictions based on climatologically averaged values of the predictor variable. Second, marine habitat is multi-dimensional in covariate space. A point in this multi-dimensional habitat space that is represented by the climatological averages of the associated predictor variables might not lie within the cloud of points associated with the contemporaneous values of the predictor variables. In essence, this issue is similar to extrapolating outside of the covariate space used to fit the model (Bouchet *et al.*, 2020). I **recommend** that the authors conduct a sensitivity analysis to investigate the effects of making predictions from the DSM based on climatological vs. contemporaneous covariates.

## B. CHOICE OF DOLPHIN OCCURRENCE DATA FOR MODEL CONSTRUCTION
I read Mackenzie and Clement (2016). The dolphin aerial line-transect surveys upon which Roberts *et al.* (2019) coastal dolphin DSMs were based were conducted using rigorous field methods that have already been approved by the IWC SC. The line-transect survey data and the associated detection probability analyses documented in Mackenzie and Clement (2016) are of high quality and useful for providing information about dolphin absolute density and absolute abundance in the surveyed areas during the survey periods and years.

I am not an expert on Hector's and Māui dolphins, so I am not qualified to assess whether supplementary data exist (e.g., satellite telemetry, acoustics) that could have been incorporated into the DSM to reduce the spatial or temporal gaps in available information.

Specific comments about the North Island harbor DSM are provided in section D ('*Methods Used to Derive a Composite Māui Dolphin Density Surface Model by Merging Predictions from Coastal and Harbor Models*') below.

## C. DOLPHIN SPATIAL MODEL FORMULATION, FITTING, AND SELECTION CRITERIA

### C1. Spatiotemporal Extent and Resolution
*C1.1 Spatial Extent*

The offshore extent of the coastal dolphin DSM was the 250m isobath. It was difficult to evaluate whether the DSM encompassed the complete range of all subpopulations of Hector's and Māui dolphins. When I conducted a literature search, I found that some authors reported the offshore extent of the distribution in distance alone, and there was no simple means to convert distance to bathymetric depth. Additionally, could not find a scale bar on any of the maps presented in Roberts *et al.* (2019). However, assuming that the dolphin sightings plotted in Appendix 7 (p.112) represent all available information on all subpopulations, it appears that the spatial extent of the dolphin DSM was sufficient. Note, however, that the inshore trawl survey conducted in 2017, which collected some of the prey data used to derive the species-specific predictions of demersal prey occurrence, evidently did not extend as far offshore as the 250m isobath in all areas:

> *'The survey was a two-phase stratified random survey after Francis (1984). The survey area covered depths of 20-200m off the west coast of the South Island from Cape Farewell to Karamea; 25-400m from Karamea to Cape Foulwind; 20-400m from Cape Foulwind to the Haast River mouth; and 10-70m within Tasman and Golden Bays inside a line drawn between Farewell Spit and Stephens Island (Figure 1a-b). The maximum depth on the west coast north of Karamea was limited to 200m because of historically low catch rates in the 200-400m range.' (Stevenson and MacGibbon, 2018, p.3)*

*C1.2 Resolution*

The spatial resolution of the dolphin DSMs was 1km x 1km. I am concerned that this spatial resolution is too small given the sampling resolution of the available data and the relevant ecological phenomena that may be used to reliably predict dolphin density.

A few observations help to put the 1km$^2$ spatial resolution into perspective. The latitudinal distance of the study area exceeds 1,600km. The distance between transects off the west coast of the South Island during the 2015 Hector's dolphin aerial survey ranged from 3.7 to 22.2km, although the most common lag distance was 11.1km (Mackenzie and Clement, 2016). The public dolphin sightings data used to construct the North Island harbor dolphin DSM were opportunistic. The aerial survey of recreational boats upon which the effort proxy for the North Island harbor dolphin DSM was based was a continuous strip transect. The inshore trawl stations at which prey data were collected in 2017 were randomly generated, but required to be a minimum of 5.6km apart (Stevenson and MacGibbon, 2018); I assumed that a similar resolution was used for all surveys that collected the prey data upon which the prey occurrence models were based. Mackenzie and Clement (2016) summarized the results from MacKenzie *et al.* (2012) as follows:

> *'MacKenzie et al. (2012) noted from simulation studies that resulting maps from a DSM analysis were sensitive to the exact location of detections when using a 5×5km prediction grid and recommended that for the purpose of robust inferences about distribution, coarser spatial scales (e.g., cells of hundreds of square kilometres) should be used.' The spatial resolution of the satellite data used to derive the ocean color habitat predictor variables was 500m (Pinkerton et al., 2018).*

The use of a 1km$^2$ spatial resolution with a monthly or seasonal temporal resolution for model-building and a climatological average for predicting dolphin densities is inconsistent with the spatiotemporal scales of ecological phenomena in marine ecosystems. Oceanographic phenomena have characteristic spatial and temporal scales that affect the distribution, density, and movement patterns of cetaceans and other marine life across multiple hierarchical scales (Haury *et al.*, 1978; Mannocci *et al.*, 2017). At fine spatiotemporal scales, cetaceans track ephemeral prey patches over tens of meters to a few kilometres and a few hours (Mannocci *et al.*, 2017). At intermediate scales (mesoscale) spanning tens to hundreds of kilometres, persisting for hours to months, cetaceans target ephemeral and seasonal oceanographic features such as fronts and eddies. At broad scales, cetaceans select persistent water masses and current systems extending over thousands of kilometres that delimit their geographic range. The mesoscale resolution is likely the best analytical resolution to use to provide reliable information about Hector's and Māui dolphin density to inform the present risk assessment. At finer spatial scales, dolphin densities likely cannot be reliably explained or predicted given the information on the environment throughout New Zealand's waters that is available for the risk assessment. Furthermore, my suspicion is that the aerial line-transect survey segments used to construct the coastal dolphin DSM had too many counts of zero dolphins for the assumed probability distribution (more on this in section C2.3, 'Probability Distribution', below). I ***recommend*** using an analytical spatial resolution of 5km x 5km or 10km x 10km for the dolphin DSMs.

### C2. Dolphin Coastal DSM Formulation
*C2.2 Response Variable*

The response variable for the coastal dolphin DSM was not completely specified. However, I assumed that it was identical to the estimator that Mackenzie and Clement (2016) present on p.20:

$$\widehat{N}_t = \sum_{j=1}^{n_j} \frac{s_{ij}}{E\left(p.\left(s_{ij}\right)\right)\widehat{P_{ak}}}$$

Where $\widehat{N}_t$ represents estimated dolphin abundance in a transect segment 1km long and 0.6km wide, including correction factors for perception and availability bias. (The present analysis uses a single abundance estimate for each subpopulation, which Mackenzie and Clement (2016) computed by averaging the corresponding summer and winter abundance estimates for each region.) Using $\widehat{N}_t$ as the response variable, all of the information about survey 'effort' (distance surveyed, effective strip width, transect detection probability, availability bias) is confounded with the sighting *data* (number of dolphins *observed*) on the left-hand side of the DSM model equation and that complicates the estimation of uncertainty in the dolphin density predictions from the DSM. Furthermore, Mackenzie and Clement (2016) estimates of perception bias and availability bias (and associated uncertainty) varied by region, so there may be covariance between these detectability parameters and the environmental predictor variables. Best practices for building a DSM are to incorporate effort as an offset term on the right-hand side of the DSM model equation. Then, in the case of an mgcv-based DSM, Bravington *et al.* (2021)'s variance propagation techniques (i.e., the dsm_varprop function in the dsm package) could be used to incorporate uncertainty from the spatially heterogeneous detection process into the overall estimate of uncertainty in the DSM predictions.

### C2.3 Probability Distribution
All dolphin DSMs assumed a negative binomial error structure. As noted below, the Q-Q plots of residuals from the dolphin coastal DSM provide some indication of lack of fit. The Tweedie distribution is another probability distribution that tends to work well with overdispersed count data and that is easily implemented in mgcv. I have also found that plots of the theoretical and estimated relationship between the mean and variance, as in Ver Hoef and Boveng (2007), can be useful in choosing among probability distributions for DSMs. I **recommend** comparing the fit of negative binomial and Tweedie dolphin coastal DSMs.

### C2.4 Link Function
The link function used to construct the dolphin DSMs was not specified. I assumed that the authors used a logarithmic link function, as would be appropriate for both a negative binomial and Tweedie distribution.

### C2.5 Linear Predictor
Based on the information presented in Table 7 (p.32), seasonality entered the dolphin coastal DSM via factor-smooth interactions with the temporally dynamic ocean color predictor variables. Each habitat variable was allowed to enter the linear predictor for the dolphin coastal DSM as a univariate spline (static and prey variables) or as a factor-smooth interaction (ocean color predictor variables).

The authors note Roberts *et al.* (2019) that, for each smooth term, the argument *k* that specifies the smooth basis dimension was '*constrained to prevent complex relationships that lack biological credibility (fixed to equal four for the initial model term, then three for all subsequent terms)*'. Although a minor point, it is worth noting that setting *k*=4 does not necessarily allow the evaluation of all functions with effective degrees of freedom (EDF) equal to 4. The *choose.k* helpfile in mgcv explains the situation as follows:

> '*One scenario that can cause confusion is this: a model is fitted with k=10 for a smooth term, and the EDF for the term is estimated as 7.6, some way below the maximum of 9. The model is then refitted with k=20 and the EDF increases to 8.7 - what is happening - how come the EDF was not 8.7 the first time around? The explanation is that the function space with k=20 contains a larger subspace of functions with EDF 8.7 than did the function space with k=10: one of the functions in this larger subspace fits the data a little better than did any function in the smaller subspace. These subtleties seldom have much impact on the statistical conclusions to be drawn from a model fit, however.*'

For future reference, it might be better to set *k* slightly higher (e.g., larger than the highest EDF that you want to allow) in order to allow the model to explore a larger range of functions having an acceptable EDF.

One larger concern that I had with the linear predictor was the strict enforcement of seasonality in the static predictors distance offshore, depth, and slope. Considerable information exists that identifies seasonal geographic shifts (onshore/offshore; possibly alongshore) in dolphin distribution. I **strongly recommend** that the authors consider s(distance offshore | season), s(depth | season), and s(slope | season) as potential predictor variables in the DSM. These types of factor-smooth interactions would allow the dolphins' distribution with respect to depth (for example) during summer to be different from their distribution with respect to depth during winter. See Pedersen *et al.* (2019) for this and other hierarchical GAM structures that can be created using the mgcv package. My primary concern here is that it is possible that a term such as

s(depth | season) might perform at least as well as one of the prey variables. If that were the case, the seasonal interaction with depth would be preferred because all of the smooth functions would come from the 'correct' season, eliminating the ambiguity raised in Q1 above.

Another fundamental concern with the linear predictor is that the use of a single intercept term for all subpopulations implies that the underlying densities of all subpopulations (during both seasons) are equal. However, the subpopulations of Hector's dolphin around the South Island are known to be found in varying densities (Mackenzie and Clement, 2016), and the density of Māui dolphins is also known to be very low. If a single DSM is used to model dolphin density throughout New Zealand, and if dolphin density is expected to vary depending on subpopulation, then incorporating a factor for geographic stratum or subpopulation range would allow the intercept (essentially related to background density) in the DSM to vary by subpopulation. This stratum-specific intercept could address some of the discrepancies in abundance estimates derived from the dolphin coastal DSM compared to the line-transect aerial survey analyses that were presented in Table 8 (Roberts *et al.*, 2019 p.35). Furthermore, this type of geographically stratified model could still be used to derive spatially-explicit predictions of dolphin density for the North Island (if warranted) using the existing approach of assuming that the DSM predictions for the North Island were relative densities and then normalizing them based on the expected area-integrated abundance estimate from Baker *et al.* (2016). I ***strongly recommend*** that the dolphin coastal DSM be reformulated to include a geographic stratum factor variable and the resulting area-integrated estimates of abundance for each stratum be compared directly with Mackenzie and Clement (2016) regional abundance estimates to evaluate how well the DSM fits the spatially-explicit abundance data upon which it was based. If the DSM does not result in acceptable abundance estimates for the South Island subpopulations of Hector's dolphins, confidence in its ability to produce reliable predicted densities for Māui dolphins and other subpopulations on the North Island will be considerably reduced.

The present analysis uses a single abundance estimate for each subpopulation, which Mackenzie and Clement (2016) computed by averaging the corresponding summer and winter abundance estimates for each region. If season-specific abundance estimates were used, and if subpopulation density varies seasonally, a factor that accounts for geographic stratum and season could be used.

### C2.6 Uncertainty Estimation: Precision and Bias
Two overarching concerns with Roberts *et al.* (2019) spatial risk assessment are: (1) the lack of unbiased estimates of uncertainty in the dolphin coastal DSM predictions; and (2) the absence of any uncertainty associated with the dolphin DSMs in the Bayesian framework used make inference about key derived quantities, including total number of dolphin deaths and the risk ratio for each commercial fishery.

Roberts *et al.* (2019) estimated the coefficient of variation (CV) in the density predictions from the dolphin coastal DSM by dividing the standard error of the prediction by its predicted value (p.17). I assumed that these calculations were made using the mgcv function predict.gam. If this is the case, then the standard errors were based on the Bayesian posterior covariance matrix of the gam parameters (see the mgcv helpfiles for predict.gam and gamObject). Region-specific estimates of detection probabilities relating to perception and availability bias in the dolphin aerial line-transect surveys were
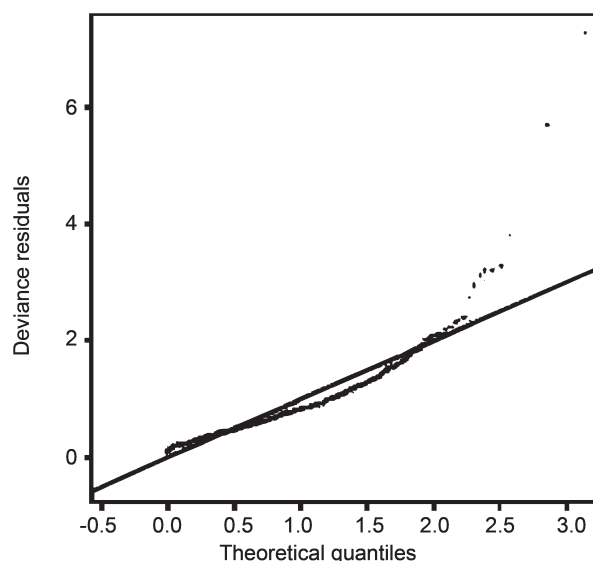


Fig. 1. Figure 9 from p.32 of Roberts *et al.* (2019).

incorporated into the response variable of the dolphin DSM, but the uncertainty in these estimated parameters was not propagated through to the estimated uncertainty in the spatially-explicit dolphin density estimates. Hence, the estimated CVs of the dolphin density predictions were negatively biased (Bravington *et al.*, 2021). The dsm_varprop function in the dsm package allows for this type of variance propagation. See also Ferguson *et al.* (2022) for an example of the varprop machinery applied to a bowhead whale case study. I ***recommend*** that the estimated CVs in the dolphin density predictions be revised to incorporate uncertainty in the estimates of detection probabilities for the dolphin aerial line-transect surveys.

I understand that estimated uncertainty due to marine mammal observer perception and availability bias during the aerial line-transect surveys was ultimately incorporated into the SEFRA model via the prior distribution on total abundance for each dolphin population (Roberts *et al.*, 2019, p.24). However, this method obscures the spatial information available in the estimated parameters and associated uncertainty.

### C3. Model Fitting
The mgcv package has been thoroughly tested and documented, and it is broadly accepted for building DSMs that can be used to inform management decisions. I have no concerns about the use of the tools in mgcv for this analysis.

### C4. Model Selection Criteria
Roberts *et al.* (2019) used a number of standard, complementary methods to select a single best dolphin coastal DSM, including AIC, percent explained deviance, visual examination of Q-Q plots, reviewing maps of predicted density, and expert elicitation.

The authors note (Roberts *et al.*, 2019, p.31) that the '*Ahuru model' was assumed to be the base case model for predicting the spatial abundance of Hector's and Māui dolphins (based on visual inspection of model predictions relative to commercial fishery observer sightings in lower abundance areas)*.' For transparency, it would be helpful if future analyses presented an appendix showing the maps of predictions from all candidate models.

The authors claim that the Q-Q plot of the 'Ahuru model' indicated that the negative binomial distribution was 'appropriate' for the data. I have copied the relevant plot below. As mentioned above, I am slightly concerned about the appearance of a left skew in the mid-range of the theoretical quantiles. Furthermore, although the resolution of the figure makes it difficult to interpret, I think I detect a relatively strong right skew at the right side of the figure. Based on these observations, I ***recommend*** that the authors compare DSMs fit using a Tweedie distribution to the negative binomial models.
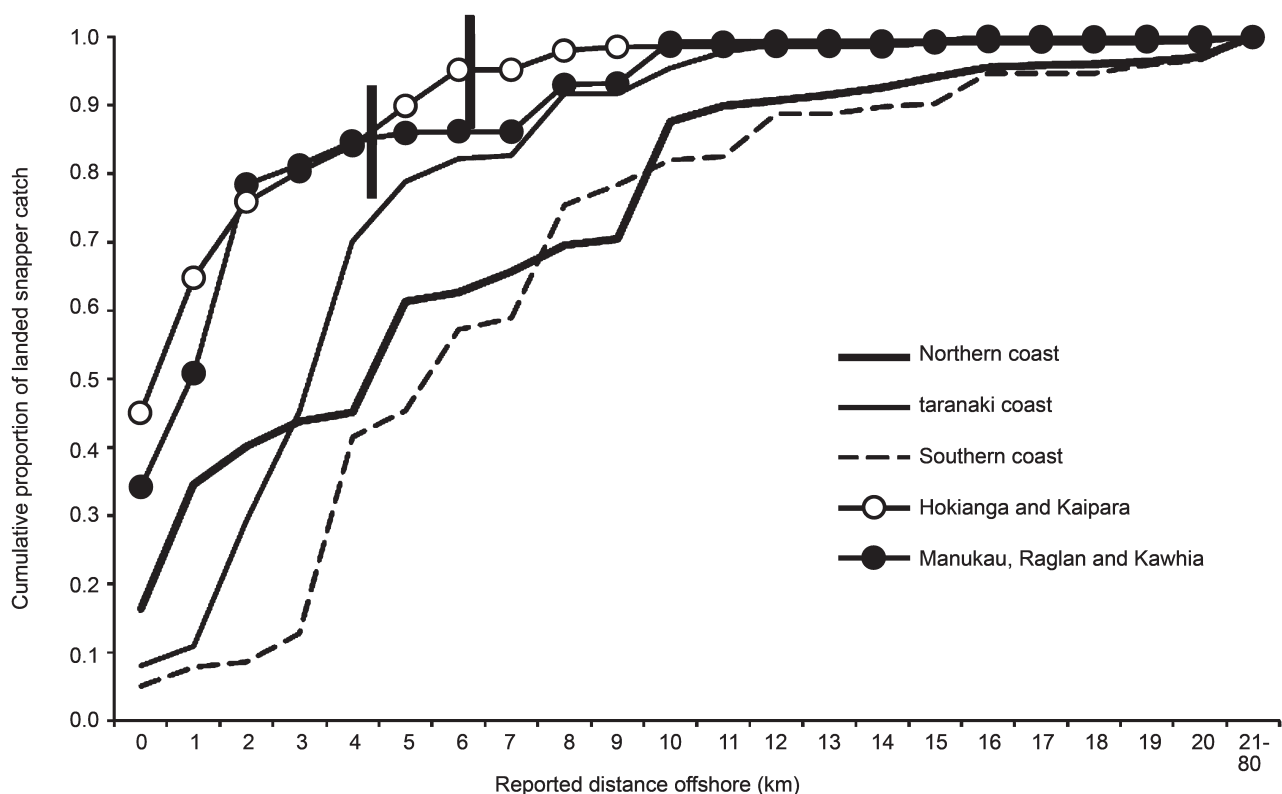


Fig. 2. Figure 12 from Hartill *et al.* (2011, p.25.)

## D. METHODS USED TO DERIVE A COMPOSITE MĀUI DOLPHIN DENSITY SURFACE MODEL BY MERGING PREDICTIONS FROM COASTAL AND HARBOR MODELS

### D1. Dolphin DSM for North Island Harbors
*D1.1 Extent and Resolution*
The public dolphin sightings data used to construct the North Island harbor DSM corresponded to depths shallower than 250m and the data were limited to the summer season (November to April, as defined in the risk assessment). The authors noted that '*the selection of only summer months was necessitated by few sighting events in winter months*'.

**Q5:** Does lack of public dolphin sightings in winter result from fewer people on the water, or is it possible that the data do not adequately represent the true winter distribution? If the latter, then where do the dolphins go during winter? If there is a seasonal migration out of the harbors, this could be a big source of bias in the Māui dolphin winter model, with cascading effects throughout the risk analysis.

A proxy for public observation effort was derived from year-round aerial surveys of coastal recreational fishing craft that were conducted between 2004 and 2007, as reported in Hartill *et al.* (2011). The survey comprised a continuous strip transect:

> '*Most of the flight path was within one kilometre of the coastline, but with some notable exception (Figure 11). Wide loops were flown off harbour entrances, towns and around offshore islands, where fishing is known to occur. The southern flight also routinely passed through a series of waypoints off the South Taranaki Bight, which were based on features which the Wanganui coastguard thought were commonly fished.' (Hartill et al., 2011, p.25)*

Hartill *et al.* (2011) estimated that recreational boats could be detected from a visual range of 3km; therefore, they inferred that all vessels in the harbors were potentially within visual range of the aerial survey observers. However, on the open coast, 'approximately 10% of the reported harvest in the North may have been taken more than three kilometres seaward of the area covered by observers, compared to 1% for the Taranaki area and 6% for the southern coast' (Hartill *et al.*, 2011, p.25). Hartill *et al.* (2011) estimate of 10% unobserved catch on the northern coast is not consistent with Figure 12 in their report (copied above).

When I look at Figure 12 from Hartill *et al.* (2011), I estimate that the cumulative proportion of landed snapper catch on the northern coast from a reported distance offshore of 4km (corresponding to the 1km recreational survey distance from shore plus 3km sighting distance) is approximately 0.45.

**Q6:** Can you clarify how Hartill *et al.* (2011) derived their estimates of 'unobserved' recreational fishing harvest? This is important because it informs whether the data from the recreational vessel aerial survey provide a relatively unbiased estimate of public observation effort.

See my comments in section C1.2 (Resolution) regarding the 1km x 1km analytical resolution. I hold those same concerns for the dolphin harbor DSM. Additionally, use of a larger cell size might alleviate some of the issues mentioned below due to lack of habitat data and subsequent 'flooding' in coastal cells.

### D1.2 Dolphin Harbor DSM Formulation
*D1.2.1 Response Variable*
The methods that Roberts *et al.* (2019) used to incorporate the effort proxy into the dolphin harbor DSM are unclear. The authors report the following:

> '*In order to account for additional sightings effort off-shore of what was visible to the aerial survey, a grid of pseudo-absences was added to the surveyed boat locations, with a 2km spacing, over water shallower than 250m depth (and continuing over land). The use of pseudo-absences was required to prevent spurious model predictions seaward of the recreational boat survey coverage and to minimise the extent to which model predictions of dolphin abundance in harbours were influenced by offshore sightings/ effort and vice versa.' (Roberts et al., 2019, p.18)*

> '*The recreational boat locations (all assumed to be negative for sightings of Hector's/Māui dolphins) were then pooled with the boat-based public sightings of Hector's/Māui dolphins.' (Roberts et al., 2019, p.18))*

> '*As with the coastal habitat modelling, GAMs were produced with sighting counts (including pseudo-absences) specified as the response variable. The candidate predictor variables were the habitat variables for the respective 1km grid cell in which each sighting event was located.' (Roberts et al., 2019, p.19)*

Given the available data, I would have expected that Roberts *et al.* (2019) would create a single analytical grid for the harbor and summarized the number of public dolphin sightings and the number of recreational boat sightings per cell. Then, maintaining consistency with how the coastal DSM was formulated, I expected that they would have estimated relative density per cell as the number of dolphin sightings divided by the number of recreational boat sightings. However, based on the information cited above, it sounds like Roberts *et al.* (2019) built a pointwise model.

**Q7:** Please provide a mathematical formulation for the dolphin harbor DSM that explicitly states the analytical unit that was used to build the model. For example, was the response variable point-based, associated with the locations of the public dolphin sightings, the recreational vessel sightings, and the pseudoabsences? Alternatively, did the response variable comprise the number of dolphin sightings, vessel observations, or pseudoabsences in each cell of a grid that covered the harbors?

**Q8:** If the dolphin harbor DSM was a pointwise model, please clarify how treating the sightings of recreational vessels as zeros in the response variable and incorporating pseudoabsences produces a relatively unbiased estimate of public observation effort.

**Q9:** Please provide a map showing all of the following layers: (1) recreational survey flightlines from Hartill *et al.* (2011); (2) 0:250m depth zone; (3) pseudoabsences used in the dolphin harbor DSM; (4) vessel sighting locations from Hartill *et al.* (2011) that were used in the dolphin harbor DSM; and (5) public dolphin sighting locations used in the dolphin harbor DSM.

### *D1.2.2 Probability Distribution*
The dolphin harbor DSMs also assumed a negative binomial distribution. However, the Q-Q plot of residuals for the selected model (fig. 14, p.38 of Roberts *et al.*, 2019) appears to show that the theoretical quantiles <10 are left-skewed and those >10 are right-skewed. I *recommend* that you compare dolphin harbor DSMs created using the Tweedie distribution to your negative binomial DSMs.

### *D1.2.3 Link Function*
The link function was not reported, so I assumed that a logarithmic link was used. This is standard for a negative binomial distribution.

### *D1.2.4 Linear Predictor*
The linear predictor for all candidate dolphin harbor DSMs included a bivariate spline for latitude and longitude. As noted above in section A1 (Prey), different bivariate splines have different behaviors regarding assumptions about the relative smoothness in each variable (i.e., isotropy vs. anisotropy) and their behavior near barriers. The issue of isotropy in a spatial spline is more of a concern if the units are latitude and longitude rather than converted to units of easting and northing via a geographic projection because the distance between one degree of longitude becomes increasingly less than the distance between one degree of latitude towards the poles.

**Q10:** What type of bivariate spline was used for the spatial term in the dolphin harbor DSMs? Did the spline produce believable predictions in the vicinity of barriers (e.g., points, peninsulas, etc.)?

The report (Roberts *et al.*, 2019, p.19) says that the same habitat layers were tested for the harbor DSMs as for the coastal dolphin DSMs. Furthermore, satellite-derived turbidity measurements were missing from some coastal cells, so a method of 'flooding' was used to fill in the missing turbidity data using the arithmetic mean of the turbidity values in adjacent cells. As discussed further below, this method of flooding introduces an unknown amount of measurement error into the model.

**Q11:** Was only turbidity affected by this issue of missing data in the dolphin harbor DSM, or was it a factor for all of the ocean color satellite data?

**Q12:** Could you provide a map showing the cells that needed to be flooded with turbidity data and also a statistic summarizing the percentage of cells in the dolphin harbor DSM that required flooding?

The report (Roberts *et al.*, 2019, p.35) says that 11.6% of the public dolphins sightings (corresponding to 17.7% of the total number of dolphins sighted by the public) were omitted from the dolphin harbor DSM because the events occurred in 1km cells whose midpoints were on land and, therefore, lacked habitat variables.

**Q13:** Why did you choose to delete these nearshore dolphin sightings rather than use the flooding technique described above to derive an estimate of the missing habitat variables for the portion of the cell that was in the water?

**Q14:** Did you examine the interannual variability in public dolphin sightings in the harbor? If so, was the level of interannual variability in sightings sufficient to justify the use of temporally dynamic predictor variables?

My concern here is that the dolphin data and the estimates of prey occurrence corresponded only to the summer season. If the interannual variability in public dolphin sightings were relatively small, my suspicion is that it would be better to use a purely spatial model for the dolphin harbor DSM because that would eliminate the measurement error associated with 'flooding' the missing dynamic variables.

**Q15:** Are there any behavior data for Hector's or Māui dolphins in the harbors? If so, do those data provide insight into whether environmental variables vs. noise and other types of anthropogenic disturbance are more likely to drive the spatiotemporal variability in dolphin distribution in harbors? If anthropogenic activities may affect dolphin distribution in the harbors, this should be evaluated along with the environmental predictor variables in the harbor DSMs.

Each of the predictor variables other than the bivariate term for location entered the linear predictor as univariate splines with the basis dimension *k* 'constrained to equal 3' (Roberts *et al.*, 2019, p.19). My comment in section C2.5 (Linear Predictor) regarding the effects of setting *k* in the specification of a smoothing spline in mgcv is relevant here.

### D1.3 Dolphin Harbor DSM Uncertainty Estimation: Precision and Bias

I was unable to thoroughly evaluate the methods used to estimate uncertainty in the dolphin harbor DSM due to my lack of understanding of the model formulation, specifically as it relates to the response variable (section D1.2.1). However, given what I understand about the model, I have a few concerns about the estimation of uncertainty.

First, it seems like a proportion of the area covered by the harbor DSMs contained public dolphin sightings but was not surveyed by the aerial surveys for recreational vessels. If this is the case, then it is a source of uncertainty in the model that cannot be quantified and, hence, cannot be accounted for in the risk assessment.

Second, the flooding method that was used to fill in missing values of turbidity introduced measurement error that is unquantified and likely unquantifiable. This issue could be avoided if the interannual variability in dolphin densities in the harbors were low enough to justify creating a purely spatial model.

Third, I am concerned about the lack of winter dolphin sighting data for the North Island harbors (Q5 above). It would be helpful to learn more about potential supplementary data (e.g., acoustics, tagging) that might be available to inform the question of whether the summer density surface is a good approximation for the winter season.

### D1.4 Dolphin Harbor DSM Fitting

The report (Roberts *et al.*, 2019, p.19) states that after 'the optimal model structure was found, the model was refitted using the habitat layer with flooded harbour cells, to include information from the upper reaches of harbours where cells were otherwise missing'.

**Q16:** Why was the complete habitat layer, including the flooded harbor cells, not used in the initial model selection process for the dolphin harbor DSM? Did you try implementing the model selection process on DSMs created using the flooded dataset?

### D1.5 Dolphin Harbor DSM Model Selection Criteria

The criteria used to select the best model seemed sound.

## D2. Merged Dolphin Coastal and Harbor DSM

The methods used to merge the dolphin predictions from the coastal and harbor DSMs seem reasonable, given the existing constraints associated with lack of information about dolphin density for the North Island.

## E. MODEL VALIDATION AND INTERPRETATION OF RESULTS, INCLUDING A REVIEW OF MODEL ESTIMATES AT THE SCALE OF SMALL OR HYPOTHESIZED LOCAL DOLPHIN POPULATIONS

## E1. Final Dolphin DSM Model Validation

### E1.1 Maps of Predicted Dolphin Density

Appendix 7 (Roberts *et al.*, 2019 p.112) provides maps of density predictions from the final dolphin DSM for North Island and South Island, overlaid with other information: public sightings, dead (strandings at sea) dolphins, NOMAD fish observer sightings, fishery-reported captures. However, it is difficult for someone without familiarity with the marine ecology and anthropogenic activities in the waters of New Zealand to interpret these plots. For example, upon first glance, the top map in fig. A7-1 (p.112) seems to show a relatively high density of public dolphin sightings along the northern coast of the North Island, where the dolphin density model predicted low densities.

**Q17:** Can this apparent discrepancy between the large number of public dolphin sightings along the north coast of the North Island and the low predicted densities from the final dolphin DSM (Roberts *et al.*, 2019 p.112) fig. A7-1) be explained based on expected spatial variability in public observation effort around the North Island?

Additionally, I am intrigued by the differences in predicted dolphin density estimates, for both winter and summer, derived from the final merged DSM that are shown in fig. 18 on p.41 of Roberts *et al.* (2019).

**Q18:** Do the maps in fig. 18 (Roberts *et al.*, 2019, p.41) support or contradict the assumption that the habitat preferences of Hector's and Māui dolphins are constant throughout New Zealand's waters?

### E1.2 Dolphin Abundance Estimates by Region/Subpopulation

The authors compared the dolphin subpopulation abundance estimates ('potential habitat capacity') derived from density predictions from the dolphin coastal DSM (the 'Ahuru model') to the analogous estimates in Mackenzie and Clement (2016). Table 8 (Roberts *et al.*, 2019, p.35) and the associated text on p.34 shows that the subpopulation abundance estimates for ECSI and WCSI from the DSM were within the 95% CI of the respective estimates from Mackenzie and Clement (2016). The DSM seasonal estimates for the SCSI exceeded the 95% CI from the Mackenzie and Clement (2016). The DSM seasonal

estimates for the WCNI were nearly 60 to 80 times larger than the best estimate of Māui dolphin abundance (Baker *et al.*, 2016). Another striking result presented in Table 8 was the predicted potential habitat capacity of the area denoted as being outside of the current dolphin core distribution, which amounted to 7,215 dolphins in summer and 11,238 dolphins in winter.

As suggested in section C2.5 (Linear Predictor), I would expect that a reliable DSM that is acceptable for providing management advice would show better agreement between the area-integrated regional predictions of abundance and the respective abundance estimates from Mackenzie and Clement (2016). I think that adding subpopulation or survey subarea as a factor in the DSM will likely improve DSM-predicted subpopulation abundance estimates because this would allow the model to derive estimates of background density separately for each region via region-specific intercept terms.

Additionally, to achieve better agreement between abundance estimates derived from the DSM and estimates of absolute abundance based on conventional analyses for distance sampling data, the authors need to correct for detransformation bias. Detransformation bias arises in a maximum likelihood framework when a nonlinear link function is used with a linear predictor that includes random effects (e.g., smoothing splines), and the response variable (e.g., the natural logarithm of abundance) is transformed via an inverse link function to produce a derived variable (e.g., abundance; Thorson and Kristensen, 2016). The magnitude of this bias depends on the variance of the predictions in the space of the link function and also on the second derivative of the link function (see equation 1a on p.68 of Thorson and Kristensen, 2016). I have found that the associated detransformation bias can be as large as 10% of the area-integrated abundance estimate from a DSM. I am not aware of any existing functions associated with the mgcv package for adjusting model predictions for detransformation bias; however, Thorson and Kristensen (2016) introduce a function that has been widely used in the TMB framework and it is possible to convert a GAM produced by mgcv into a TMB model.

Additionally, I am concerned about the large proportion of dolphin aerial survey data from the SCSI region that were omitted from the coastal dolphin DSM due to missing environmental predictor variables. I wonder if replacing the prey covariate with a factor-smooth interaction between season and depth would allow more of the dolphin aerial survey data to be retained in the model-building phase and, therefore, produce more reliable results in the SCSI region.

### E1.3 Spatial Cross-Validation

The ability to validate the final dolphin DSM is limited by the same factor that complicated the analysis in the first place: lack of information about the spatiotemporal patterns and associated variability in Hector's and Māui dolphins throughout New Zealand during all months of the year. However, the thorough aerial survey coverage on the WCSI and ECSI provide an opportunity to examine the reliability of the coastal DSM when extrapolating to areas outside of the region on which the model was built. I **strongly recommend** that the authors implement spatial cross-validation methods by withholding dolphin aerial survey data from the WCSI during the model-building phase and evaluating how well that model predicts density and abundance in the WCSI. Repeat this exercise for the ECSI.

### E2. Interpretation of Results

### E2.1 Causation vs. Correlation

Although a relatively minor point, I suggest distinguishing between causation and correlation when interpreting results from similar DSMs in the future. For example, the following excerpt from Roberts *et al.* (2019, p.69) interprets results from the harbor and coastal dolphin DSMs in the context of potential future environmental changes:

> *'In addition to potential changes in prey recruitment patterns and spatial distribution, climate change may potentially affect the turbidity of coastal waters through increased rates of glacial melting in summer months, and by affecting wind-driven re-suspension of sediments from the seafloor in winter months. Factors that affect the intensity and spatial extent of coastal turbidity are likely to affect the spatial distribution of Hector's and Māui dolphins, based on their habitat preference (Figure 10). It is also likely that many of their preferred prey species (e.g. red cod, sprat, ahuru and NZ plaice) have an affinity for turbid water, given the moderate positive spatial correlation between these species and coastal turbidity patterns (Figure 8). As such, their prey may also be affected by potential climate warming effects on coastal turbidity.'*

It is reasonable to ask whether the dolphins and prey are targeting the *turbidity*, or if turbid water is a characteristic of the habitat that is targeted, e.g., fish communities associated with river discharge or river deltas. The caution here is that focusing on turbidity as a habitat predictor could be misleading when trying to extrapolate to a novel multidimensional environmental covariate space (e.g., a future climate).

### E2.2 Focus on Precision/Lack of Consideration of Bias

The authors used their estimates of CV in the predicted dolphin densities as one metric for evaluating model performance:

> *'With respect to model predictions of spatial abundance, the coefficient of variation was relatively low (CV less than 0.3) across all coastal regions of both the North and South Island, indicating that the habitat used to make predictions was consistently within the modelled habitat envelope. The main exception to this was the harbours in the west coast of the North Island, where a CV of greater than 1 was obtained for a number of grid cells. This was caused by very high turbidity values beyond the range used to fit the habitat model (Figure 11). This indicates that the habitat models developed are suitable for predicting abundance in coastal and offshore waters, though not into high turbidity harbour areas.' (Roberts et al., 2019, p.33)*

However, as noted above, the estimated CVs in the DSM predictions are negatively biased because they do not incorporate uncertainty in the estimated detectability parameters needed to derive the estimates of dolphin abundance on line-transect segments (the response variable in the coastal dolphin DSM).

Additionally, precision is only one component in uncertainty. Bias is also important to address when evaluating models that may be used for management. The lack of agreement between the area-integrated abundance estimates for subpopulations derived from the DSM and the respective estimates in Mackenzie and Clement (2016) provide evidence of considerable bias when predicting at the regional scale in the present formulation of the model. Additionally, the apparent lack of information about Māui dolphin distribution and density during winter considerably impedes evaluation the model's ability to make spatially-explicit density predictions for this subspecies during 6 months of the year. This is a potentially large source of bias that is not captured in CVs generated by the predict.gam function in mgcv.

## F. EVALUATION OF SPATIALLY-RESOLVED PREDICTED BYCATCH COMPARED TO KNOWN RECORDS, INCLUDING BEACH-CAST CARCASSES AND FISHER-REPORTED CATCHES FROM VESSELS WITHOUT OBSERVERS

The methods used to assess commercial fishery-related deaths are presented in section 4.5 of Roberts *et al.* (2019), beginning on p.21. To summarize, the spatial model of predicted bycatch involved first overlaying the dolphin density predictions from the final merged DSM with a density surface representing fishing intensity (kilometers of set net fishing; number of trawls for the inshore trawl fishery). This produced an estimate of the overlap between dolphin distribution and fishing activities that was then scaled by Mackenzie and Clement (2016) total estimated abundance of Hector's and Māui dolphins (independently for each taxon) to derive an estimate of the total number of individual dolphins of each taxon likely to encounter each fishery. The estimate of predicted bycatch was not spatially explicit. To infer the number of dolphins captured and the number of dolphin deaths for a fishery from the estimated number of dolphins encountering each fishery, the model needed four key variables: (1) vulnerability ($v_g$, where $g$ indexes the type of fishery), the probability that a dolphin would be captured; (2) the probability that a [capture] event is observable given that an observer is on watch ($p_g^{obs}$); (3) the probability that an individual was alive when captured ($\psi_g$)); and the probability that a captured dolphin would be released alive ($\omega_g$).

### F1. Model Definitions and Assumptions
#### F1.1 Definition of a Capture 'Event'
When I initially read the SEFRA model methods presented in section 4.5 (beginning on p.21, and including Table 4 on p.22 of Roberts *et al.*, 2019), I interpreted the definition of a capture event to refer to a single dolphin. However, when I reached the results in section 5, I was unsure whether a capture event could comprise more than one dolphin:

> 'Fisheries observers have recorded a single Hector's dolphin capture in the inshore trawl fisheries. Of the 13 inshore trawl events since 1977 on which Hector's dolphins were reported as captured by the fishing industry, six captured more than one individual, with a maximum of four individuals captured by a single trawl. A mean of 1.92 individuals per positive capture event was obtained for inshore trawls using fishing industry reported capture records (this includes three individuals reported captured on a single trawl in 2019). Assuming that two individuals were captured per inshore trawl event with a capture, this would lead to a doubling of the estimated annual deaths from the risk model (Table A17-1 compared with Table 15).' (Roberts et al., 2019, p.48)

**Q19:** Please clarify the definition of a capture event, as used in the SEFRA model, with respect to the assumed/potential/ expected number of dolphins involved in a single event.

#### F1.2 Assumption about Fisheries Posing Bycatch Risk
The report states that '[t]his implementation of the SEFRA method includes two commercial fishery groups known to pose a risk to Hector's/Māui dolphins - set netting and inshore trawling.'

**Q20:** Are there other fishery groups that may pose risk of Hector's or Māui dolphin bycatch? Could the lack of reported bycatch in other commercial fisheries be due to the small sample size in necropsy data or to lack of observers onboard commercial fishing boats?

#### F1.3 Vulnerabililty
A key assumption in the spatial bycatch threat model is that vulnerability does not vary with space or time. The spatial model does not allow for variability due to habitat, activity state of dolphins, targeted fish species, mesh size, gear depth, etc. Potential implications of this assumption are discussed further below.

#### F1.4 Probability that an Event is Observable, $p_g^{obs}$
Another key assumption in the SEFRA model relates to $p_g^{obs}$, which is the probability of a capture event being observed, conditioned on an observer being on watch. This parameter is inversely related to the cryptic mortality rate ($k$) as $k=1/p_g^{obs}$ (Roberts *et al.*, 2019 p.23) and Appendix 10, p.130). Sources of cryptic mortality include pre-catch losses, ghost fishing mortalities, and post-release mortalities (Roberts *et al.*, 2019 p.130). The parameter $p_g^{obs}$ was assumed to be constant in space and time, and there was little information upon which to base the prior. Potential implications of assumptions about the magnitude and constant nature of $p_g^{obs}$ are discussed further below.
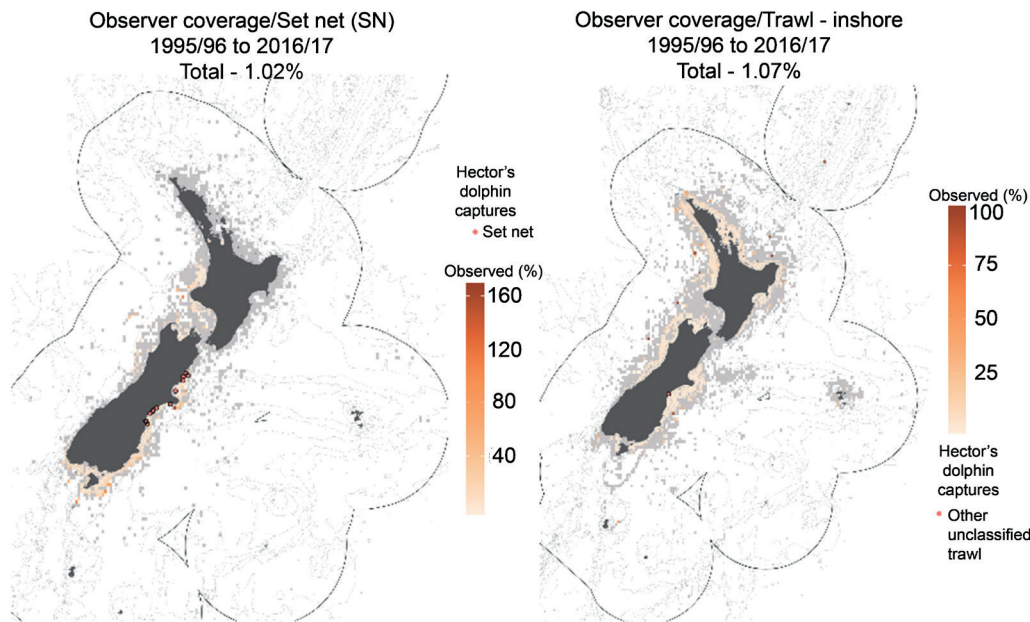
Fig. 3. Figure A15-1 from (Roberts *et al.*, 2019, p.149).

### F1.5 Fishing Effort and Time

Set net fishing effort is measured in terms of kilometers of net, and inshore trawl effort is measured in terms of the number of trawls. The SEFRA model predicts annual deaths based on these defintions of commercial fishing effort.

**Q21:** Does the duration of a fishing activity potentially have an impact on vulnerability or $p_g^{obs}$? Is a net that is in the water for 20 seconds less likely to catch a dolphin than a net that is in the water for 20 hours or 20 days (e.g., ghost gear)? If the time that gear is in the water is related to vulnerability or $p_g^{obs}$, and if there is spatial or temporal variability in this gear 'soak time', then this is an additional source of uncertainty that has not been addressed.

**Q22:** If ghost gear has a greater catchability due to the period of time that it is in the water, do we know how the number of ghost nets has changed over time? This affects the reliability of the assumption that $p_g^{obs}$ has been constant over time.

### F2. Data Used to Build Spatial Model of Commercial Fishery Threat Intensity

### F2.1 Limited to Fisheries Observer Data

Bycatch data used to build the spatial threat intensity model were limed to commercial fisheries protected species captures data:

> *'Commercial fisheries protected species captures data were received 3 July 2018 from Dragonfly Data Science via Fisheries New Zealand. The number of observed dolphin captures by fisheries group and fishing year from 1995/96 to 2016/17 is shown in Table 11. There have been no observed captures of Māui dolphins during this time. All 16 Hector's dolphin captures were from the East Coast of the South Island. The majority of the captures were in set nets, with just one of the captures in an inshore trawl. Only three of the 16 captures were alive.' (Roberts et al., 2019, p.42)*

Additional vessel-reported capture data exist, but these data were not used to build the spatial threat intensity model:

> *'No captures were recorded on observed trawls in a region immediately to the north of Banks Peninsula, where the estimated observed captures were relatively high (Figure A15-3). However, vessel-reported trawl capture events have been recorded there since 2014/15, and all other vessel-reported trawl captures since then were located on the east coast of the South Island and in areas of relatively high model-predicted observed captures (DOC 2019). There were no observed captures in observed WCNI commercial set net fisheries; this is consistent with model predictions, for which estimated numbers of observed captures are close to zero, reflecting very low estimated captures in the present, and low observer coverage in the past when estimated capture rates were higher.' (Roberts et al., 2019, p.48)*

**Q23:** How many vessel-reported captures exist? Where and when (season and year) did they occur? (my understanding is the vessel-reported captures are not shown in maps in Appendix 15).

**Q24:** Why were vessel-reported captures not included in the spatial threat intensity model and the SEFRA model?

*F2.2 Necropsy Data*

Necropsy data were used to validate (visually) the spatial threat intensity model and SEFRA model, and they were used to apportion deaths to different mortality sources by the SEFRA model. Necropsy data were not used to construct the spatial threat intensity model. The sample size in the necropsy database was small: $n$=55 non-calf carcasses; $n$=21 calf carcasses.

Given the lack of fisheries observer coverage (e.g., Appendix 15) and small sample sizes of observed dolphin captures, I **strongly recommend** conducting a sensitivity analysis to investigate whether incorporation of vessel-reported captures and (if feasible) necropsy data in the spatial threat intensity model affects the SEFRA model results, particularly the estimated number of deaths and the risk ratio.

### F3. Model Validation and Interpretation of Results

*F3.1 Available Information and Sensitivity of Model Results to Model Assumptions*

The sample sizes of observed bycatch in the fisheries observer database were small: zero observed Māui dolphin captures and 16 total observed Hector's dolphin captures, 15 of which were in the set net fishery. Fisheries observer coverage was limited and appears to be spatially biased (Roberts *et al.*, 2019, p.149, fig. A15-1, copied below). These small bycatch sample sizes, combined with limited information about the true values of key model parameters ($v_g$, $p_g^{obs}$, $\psi_g$ and $\omega_g$), resulted in limited information for model fitting and validation. For example, the posterior distributions for vulnerability and $p_g^{obs}$ were very similar to their respective prior distributions (Appendix 14, p.146).

The authors highlight the spatial agreement between the observed and predicted captures: 'Notably the agreement between the spatial distribution of predicted and actual observed commercial fishery captures was good (fig. A15-3), indicating that the estimated spatial density of dolphins was a good approximation to the true density and that vulnerability is relatively constant in space' (Roberts *et al.*, 2019, p.63). Due to the data limitations for model building and validation, I am concerned about reaching any strong conclusions about the spatial variability in dolphin density and vulnerability based on these model results. These results suggest that the model appears to fit the data well. However, if there were spatial variability in any of the key model parameters that were assumed constant for each fishery ($v_g$, $p_g^{obs}$, $\psi_g$ and $\omega_g$), given the existing spatial biases in fisheries observer coverage, it is quite possible that the SEFRA model results are biased but we cannot detect this bias using the available information. Therefore, I **strongly recommend** conducting sensitivity analyses (via simulating testing) into the model-derived estimates of dolphin captures, deaths, PST, and risk ratios for scenarios in which the true values of the key model parameters ($v_g$, $p_g^{obs}$, $\psi_g$ and $\omega_g$) varied spatially.

*F3.2 Estimated Uncertainty in Estimated Parameters and Model-Derived Quantities*

As noted previously, the posterior distributions for the estimated parameters and model-derived quantities for the SEFRA model are likely too narrow because the Bayesian inference did not account for the uncertainty associated with the spatially-explicit predictions from the final merged dolphin DSM. Hence, interpretation of the estimated median and CI values from the SEFRA model should proceed with caution.

*F3.3 Scalability of Predicted Risk*

I have one additional minor comment regarding the spatial resolution of the SEFRA model results. The authors state, '[n]ote however that because fishing effort and fisheries observer data are utilised at the level of individual fishing events, cell size has no effect on model outcomes; and maps of captures and risk can be re-generated at any user-defined cell size' (Roberts *et al.*, 2019, p.23). However, this assertion is not entirely correct. The resolution of the dolphin DSM affects the estimated number of dolphin captures per cell. If all SEFRA model specifications were held constant but the resolution of the grid used for the dolphin DSM were changed, the estimated number of dolphin captures would likely change, with cascading effects on other derived quantities (number of dolphin deaths, PST, and risk ratio).

## G. EVALUATION OF SPATIALLY EXPLICIT ESTIMATES AND ASSOCIATED UNCERTAINTY OF DOLPHIN DEATHS FROM COMMERCIAL FISHERIES COMPARED TO ESTIMATES FROM SIMPLER MODELS

Many of the issues addressed above in section F are relevant here: (1) assumptions about which fisheries pose bycatch risk; (2) lack of a temporal metric associated with fishing effort; (3) sensitivity to, and lack of information about, key model parameters ($v_g$, $p_g^{obs}$, $\psi_g$ and $\omega_g$); (4) likely underestimation of uncertainty in derived quantities because the uncertainty from the dolphin DSM was not propagated through the SEFRA model; (5) lack of independent information with which to validate the SEFRA model predictions.

## H. REVIEW OF ESTIMATED PATTERNS OF CHANGING FISHERIES RISK OVER TIME

Due to the limited information on dolphin distribution and density, bycatch, and spatiotemporal variability in key SEFRA model parameters ($v_g$, $p_g^{obs}$, $\psi_g$ and $\omega_g$), the ability to validate the estimated patterns of changing fisheries risk over time is limited. Therefore, I **strongly recommend** conducting sensitivity analyses (via simulating testing) into the model-derived estimates of dolphin captures, deaths, PST, and risk ratios for scenarios in which the true values of the key model parameters ($v_g$, $p_g^{obs}$, $\psi_g$ and $\omega_g$) and the dolphin distribution vary over time.

## SUMMARY

I would like to reiterate the statement that I made in the beginning of my review that Roberts *et al.* (2019)'s spatial risk assessment of threats to Hector's and Māui dolphins represents an enormous amount of thought, effort, and skill.

In addition to the list of questions and recommendations that I provide in Adjuncts 1, 2, and 3, I want to highlight two points. First, due to the relatively limited information that is available to be used to inform each stage in the Hector's and Māui dolphin risk assessment, it is particularly important to strive for an honest accounting of uncertainty, including both precision and bias. In many cases, it is likely easier to estimate precision than bias. However, even acknowledging that bias exists and the likely direction of that bias (positive or negative) is useful.

Second, the problem that concerns me the most about the present risk assessment is that the model-based estimates of dolphin deaths and risk ratios associated with the commercial fisheries are sensitive to the estimated values of key model parameters ($v_g$, $p_g^{obs}$, $\psi_g$ and $\omega_g$) about which limited information exists. For example, a doubling of vulnerability directly translates to a doubling of the number of dolphin deaths and the risk ratio. The risk assessment that Roberts *et al.* (2019) presented for the case of equal detection probability among non-fishery mortality events and a calibration coefficient value of ϕ=0.2 resulted in median estimates of 41 and 5 Hector's dolphin deaths in the set net and inshore trawl fisheries, respectively, and analogous values of 0.1 and 0.0 for Māui dolphins. If, for example, the true value of Hector's dolphin vulnerability in set nets were twice as large as estimated - which is reasonable given the considerable uncertainty in this parameter estimate – this would double the risk ratio. What is the likelihood that the twofold increase in Hector's dolphin deaths (a change from 41 to 82 individuals) would be detected?

### Literature Cited

Baker, C.S., Steel, D., Hamner, R.M., Hickman, G., Boren, L., Arlidge, W. and Constantine, R. 2016. Estimating the abundance and effective population size of Māui dolphins using microsatellite genotypes in 2015–16, with retrospective matching to 2001–16. Department of Conservation, Auckland.

Bouchet, P.J.-F., Miller, D.L., Roberts, J., Mannocci, L., Harris, C.M. and Thomas, L. 2020. dsmextra: Extrapolation assessment tools for density surface models. *Methods in Ecology and Evolution* 11(11): 1464-69. [Available at: https://doi.org/10.1111/2041-210X.13469].

Bravington, M.V., Miller, D.L. and Hedley, S.L. 2021. Variance propagation for density surface models. *Journal of Agricultural, Biological and Environmental Statistics* 26: 306-23. [Available at: https://doi.org/10.1007/s13253-021-00438-2].

Ferguson, M.C., Miller, D.L., Clarke, J.T., Brower, A.A., Willoughby, A.L. and Rotrock, A.D. 2022. Spatial modeling, parameter uncertainty, and precision of density estimates from line-transect surveys: a case study with Western Arctic bowhead whales. 53pp. Paper SC/68D/ASI/01 presented to the IWC Scientific Committee, Virtual Meetings, April-May 2022 (unpublished). 53pp. [Paper available from the Office of this Journal].

Hartill, B., Vaughan, M. and Rush, N. 2011. Recreational harvest estimate for SNA 8 in 2006–07. 48pp. New Zealand Fisheries Assessment Report 2011/51.

Haury, L.R., McGowan, J.A. and Wiebe, P.H. 1978. Patterns and processes in the time-space scales of plankton distributions. pp.277-327. *In*: Steele, J.H. (eds). *Spatial Pattern in Plankton Communities*. Plenum, New York.

Mackenzie, D. and Clement, D. 2016. Abundance and distribution of WCSI Hector's dolphin. *New Zealand Aquatic Environment and Biodiversity Report No. 1168 to the Ministry for Primary Industries*. 79pp. [Plus supplementary material].

MacKenzie, D.I., Clement, D. and Mattlin, R. 2012. Abundance, distribution and productivity of Hector's (and Māui's) dolphins (Final Research Report, PRO2009-01B). (Unpublished report held by the Ministry for Primary Industries.).

Mannocci, L., Boustany, A.M., Roberts, J.J., Palacios, D.M., Dunn, D.C. and Halpin, P.N. 2017. Temporal resolutions in species distribution models of highly mobile marine animals: Recommendations for ecologists and managers. *Diversity and Distributions* 23(10): 1098-109. [Available at: https://doi.org/10.1111/ddi.12609].

Miller, E., Lalas, C., Dawson, S., Ratz, H. and Slooten, E. 2013. Hector's dolphin diet: The species, sizes and relative importance of prey eaten by *Cephalorhynchus hectori*, investigated using stomach content analysis. *Mar. Mamm. Sci.* 29: 606–28.

Miller, E.J. 2014. Ecology of Hector's dolphin (*Cephalorhynchus hectori*): Quantifying diet and investigating habitat selection at Banks Peninsula (Thesis, Doctor of Philosophy), University of Otago. 168pp.

Pedersen, E.J., Miller, D.L., Simpson, G.L. and Ross, N. 2019. Hierarchical generalized additive models in ecology: an introduction with mgcv. *PeerJ* 7: e6876. [Available at: https://doi.org/10.7717/peerj.6876].

Pinkerton, M.H., M. Espino, M., Wood, S. and Zeldis, J. 2018. Measuring the effects of bivalve mariculture on water quality in northern New Zealand using MODIS-Aqua satellite observations. *Aquaculture Environment Interactions* 10: 529–45.

Roberts, J.O., Webber, D.N., Roe, W.D., Edwards, C.T.T. and Doonan, I.J. 2019. Spatial risk assessment of threats to Hector's and Māui dolphins (*Cephalorhynchus hectori*). *New Zealand Aquatic Environment and Biodiversity Report No.* 214: 175pp.

Stevenson, M.L. and MacGibbon, D.J. 2018. Inshore trawl survey of the west coast South Island and Tasman and Golden Bays, March-April 2017 (KAH1703). *New Zealand Fisheries Assessment Report* 2018/18: 92pp.

Thorson, J.T. and Kristensen, K. 2016. Implementing a generic method for bias correction in statistical models using random effects, with spatial and population dynamics examples. *Fisheries Research* 175: 66-74. [Available at: https://doi.org/10.1016/j.fishres.2015.11.016].

Ver Hoef, J.M. and Boveng, P.L. 2007. Quasi-Poisson vs. negative binomial regression: How should we model overdispersed count data? *Ecology* 88(11): 2766-72. [Available at: https://doi.org/10.1890/07-0043.1].

Wood, S.N., Bravington, M.V. and Hedley, S.L. 2008. Soap film smoothing. *J. Roy. Stat. Soc.* B70(5): 931-55.

**Adjunct 1**

**Questions about Roberts *et al.*'s (2019) analyses or interpretation of results that would assist in a thorough evaluation of the spatial risk assessment**

Q1: What information (i.e., predictive value) do temporally mis-matched prey occurrence predictions bring to the dolphin DSMs?

Q2: Why omit dolphin sightings and effort from line-transect survey segments for which prey occurrence data were missing, if the prey occurrence data were from a mismatched season and the species is known to exhibit seasonal geographic shifts in distribution?

Q3: Did you consider incorporating prey occurrence only into the models for the season (summer) during which the prey data were collected?

Q4: For the waters around New Zealand, are remotely sensed data available for wind direction and strength, upwelling intensity, ocean surface fronts, eddy kinetic energy, and anything else that might indicate mesoscale oceanographic activity that may affect prey availability and, hence, dolphin density?

Q5: Does lack of public dolphin sightings in winter result from fewer people on the water, or is it possible that the data do not adequately represent the true winter distribution? If the latter, then where do the dolphins go during winter? If there is a seasonal migration out of the harbors, this could be a big source of bias in the Māui dolphin winter model, with cascading effects throughout the risk analysis.

Q6: Can you clarify how Hartill *et al.* (2011) derived their estimates of 'unobserved' recreational fishing harvest? This is important because it informs whether the data from the recreational vessel aerial survey provide a relatively unbiased estimate of public observation effort.

Q7: Please provide a mathematical formulation for the dolphin harbor DSM that explicitly states the analytical unit that was used to build the model. For example, was the response variable point-based, associated with the locations of the public dolphin sightings, the recreational vessel sightings, and the pseudoabsences? Alternatively, did the response variable comprise the number of dolphin sightings, vessel observations, or pseudoabsences in each cell of a grid that covered the harbors?

Q8: If the dolphin harbor DSM was a pointwise model, please clarify how treating the sightings of recreational vessels as zeros in the response variable and incorporating pseudoabsences produces a relatively unbiased estimate of public observation effort.

Q9: Please provide a map showing all of the following layers: (1) recreational survey flightlines from Hartill *et al.* (2011); (2) 0:250m depth zone; (3) pseudoabsences used in the dolphin harbor DSM; (4) vessel sighting locations from Hartill *et al.* (2011) that were used in the dolphin harbor DSM; and (5) public dolphin sighting locations used in the dolphin harbor DSM.

Q10: What type of bivariate spline was used for the spatial term in the dolphin harbor DSMs? Did the spline produce believable predictions in the vicinity of barriers (e.g., points, peninsulas, etc.)?

Q11: Was only turbidity affected by this issue of missing data in the dolphin harbor DSM, or was it a factor for all of the ocean color satellite data?

Q12: Could you provide a map showing the cells that needed to be flooded with turbidity data and also a statistic summarizing the percentage of cells in the dolphin harbor DSM that required flooding?

Q13: Why did you choose to delete these nearshore dolphin sightings rather than use the flooding technique described above to derive an estimate of the missing habitat variables for the portion of the cell that was in the water?

Q14: Did you examine the interannual variability in public dolphin sightings in the harbor? If so, was the level of interannual variability in sightings sufficient to justify the use of temporally dynamic predictor variables? My concern here is that the dolphin data and the estimates of prey occurrence corresponded only to the summer season. If the interannual variability in public dolphin sightings were relatively small, my suspicion is that it would be better to use a purely spatial model for the dolphin harbor DSM because that would eliminate the measurement error associated with 'flooding' the missing dynamic variables.

Q15: Are there any behavior data for Hector's or Māui dolphins in the harbors? If so, do those data provide insight into whether environmental variables vs. noise and other types of anthropogenic disturbance are more likely to drive the spatiotemporal variability in dolphin distribution in harbors? If anthropogenic activities may affect dolphin distribution in the harbors, this should be evaluated along with the environmental predictor variables in the harbor DSMs.

Q16: Why was the complete habitat layer, including the flooded harbor cells, not used in the initial model selection process for the dolphin harbor DSM? Did you try implementing the model selection process on DSMs created using the flooded dataset?

Q17: Can this apparent discrepancy between the large number of public dolphin sightings along the north coast of the North Island and the low predicted densities from the final dolphin DSM (Roberts *et al.* 2019, fig. A7-1, p.112) be explained based on expected spatial variability in public observation effort around the North Island?

Q18: Do the maps in Figure 18 (Roberts *et al.*, 2019, p.41) support or contradict the assumption that the habitat preferences of Hector's and Māui dolphins are constant throughout New Zealand's waters?

Q19: Please clarify the definition of a capture event, as used in the SEFRA model, with respect to the assumed/potential/expected number of dolphins involved in a single event.

Q20: Are there other fishery groups that may pose risk of Hector's or Māui dolphin bycatch? Could the lack of reported bycatch in other commercial fisheries be due to the small sample size in necropsy data or to lack of observers onboard commercial fishing boats?

Q21: Does the duration of a fishing activity potentially have an impact on vulnerability or $p_g^{obs}$? Is a net that is in the water for 20 seconds less likely to catch a dolphin than a net that is in the water for 20 hours or 20 days (e.g., ghost gear)? If the time that gear is in the water is related to vulnerability or $p_g^{obs}$, and if there is spatial or temporal variability in this gear 'soak time', then this is an additional source of uncertainty that has not been addressed.

Q22: If ghost gear has a greater catchability due to the period of time that it is in the water, do we know how the number of ghost nets has changed over time? This affects the reliability of the assumption that $p_g^{obs}$ has been constant over time.

Q23: How many vessel-reported captures exist? Where and when (season and year) did they occur? (My understanding is the vessel-reported captures are not shown in maps in Appendix 15.)

Q24: Why were vessel-reported captures not included in the spatial threat intensity model and the SEFRA model?

## Adjunct 2

### *Recommendations* that would improve the statistical methodology used in the risk assessment but might not considerably change the results that are relevant for making management decisions

(1) I *recommend* that the authors conduct a sensitivity analysis to investigate the effects of making predictions from the DSM based on climatological vs. contemporaneous covariates.
(2) I *recommend* using an analytical spatial resolution of 5km x 5km or 10km x 10km for the dolphin DSMs.
(3) I *recommend* comparing the fit of negative binomial and Tweedie dolphin coastal DSMs.
(4) I *recommend* that the estimated CVs in the dolphin density predictions be revised to incorporate uncertainty in the estimates of detection probabilities for the dolphin aerial line-transect surveys.
(5) I *recommend* that you compare dolphin harbor DSMs created using the Tweedie distribution to your negative binomial DSMs.

## Adjunct 3

### *Strong recommendations* that may critically change the results that are relevant for making management decisions

(1) I ***strongly recommend*** that the authors consider s(distance offshore | season), s(depth | season), and s(slope | season) as potential predictor variables in the DSM. These types of factor-smooth interactions would allow the dolphins' distribution with respect to depth (for example) during summer to be different from their distribution with respect to depth during winter.

(2) I ***strongly recommend*** that the dolphin coastal DSM be reformulated to include a geographic stratum factor variable and the resulting area-integrated estimates of abundance for each stratum be compared directly with MacKenzie and Clement's (2016) regional abundance estimates to evaluate how well the DSM fits the spatially-explicit abundance data upon which it was based.

(3) I ***strongly recommend*** that the authors implement spatial cross-validation methods by withholding dolphin aerial survey data from the WCSI during the model-building phase and evaluating how well that model predicts density and abundance in the WCSI. Repeat this exercise for the ECSI.

(4) I ***strongly recommend*** conducting a sensitivity analysis to investigate whether incorporation of vessel-reported captures and (if feasible) necropsy data in the spatial threat intensity model affects the SEFRA model results, particularly the estimated number of deaths and the risk ratio.

(5) I ***strongly recommend*** conducting sensitivity analyses (via simulating testing) into the model-derived estimates of dolphin captures, deaths, PST, and risk ratios for scenarios in which the true values of the key model parameters ($v_g$, $p_g^{obs}$, $\psi_g$ and $\omega_g$) varied spatially.

(6) I ***strongly recommend*** conducting sensitivity analyses (via simulating testing) into the model-derived estimates of dolphin captures, deaths, PST, and risk ratios for scenarios in which the true values of the key model parameters ($v_g$, $p_g^{obs}$, $\psi_g$ and $\omega_g$) and the dolphin distribution vary over time.
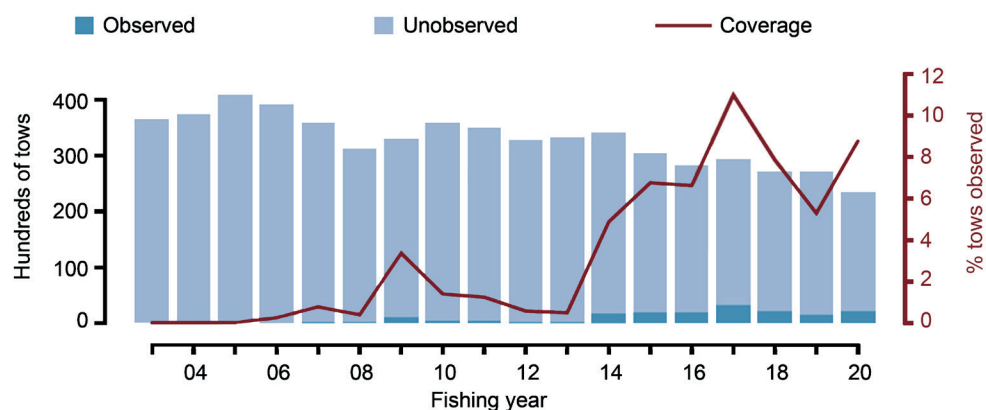


Fig. 1. Fishing effort and observer cover in inshore trawl fishery between 2003 and 2021 (source: Fisheries New Zealand. https://protectedspeciescaptures.nz/PSCv6/released/hectors-dolphin/inshore-trawl/all-vessels/eez/2019- 20/).
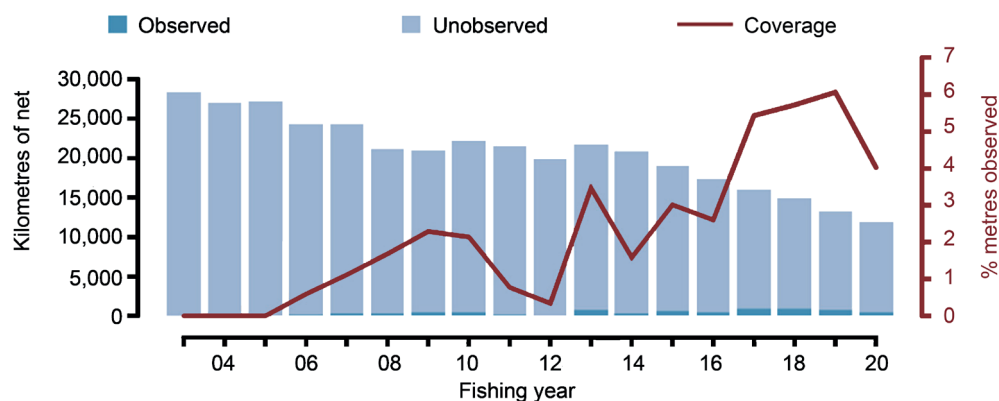


Fig. 2. Fishing effort and observer cover in setnet fishery between 2003 and 2021 (source: Fisheries New Zealand. https://protectedspeciescaptures.nz/PSCv6/released/hectors-dolphin/setnet/all-vessels/eez/2019-20/).

# Appendix 5

# Spatial Risk Assessment of Threats to Hector's and Māui Dolphins

Reviewer: Hélène Peltier

At the request of the International Whaling Commission, this work aims to review the framework used by the New Zealand government to implement management measures for Māui and Hector's dolphins. This spatial risk assessment of threats was untertaken to inform a Threat Management Plan for both sub-species. Based on Spatiallly-Explicit Fisheries Risk Assessment (SEFRA), the method evaluates encounters between dolphins and lethal threats as a function of their overlap in space. The probability of death per encounter is estimated from fisheries observer data or other datasets such as strandings.

This part of the review is mostly focused on bycatch estimates, available data and their likely implications.

## 1. FISHERY OBSERVATION PROGRAMME

As bycatch seems to be one of the major threats affecting Hector's and Māui dolphins, observed bycaught individuals and observer effort are crucial parameters. Unfortunately, if the table 11 describes the observed commercial fishery bycatch, very few information is presented in the report about the observation effort. The overall observer coverage is summed over the fishing years 1995/96 to 2016/17 and equals 1.02% for set nets and 1.07% of inshore trawlers fishing effort.

The sampling strategy of national observer programme is not described. The randomization of sampling strategy should include as many vessels as possible, ideally through relevant stratification. In other words, observer cover of 5% on 80% of vessels is more relevant and significant than 30% of observation cover on 20% of fishing vessels. As information is not provided, the significance of observer programme cannot be discussed.

More information on fishery observer programme is available online and in different reports from Department of Conservation and Ministry for Primary Industries (Fisheries New Zealand).

### Temporal distribution
It seems that before 2006/07, observer effort on inshore trawls was sparse and only locally deployed (for example a bycatch of Hector's dolphin had been reported in 1997/98 in the Pegasus Bay-Canterbury Bight area) (Rowe, 2009). Observer programme is officially in place since 2006/07, but remains inequal as it ranged from 0.4 to 11% of tows observed between 2006 and 2017 (Fig. 1).

The establishment of observer programme on setnets was just as disparate, as a dedicated programme occurred in 1997/98 in the Pegasus Bay-Canterbury Bight during which 8 Hector's dolphins were recorded. Setnet fishery was then observed in 2005/06 in Nelson/Marlborough regions, and in 2006/07 in Kaikoura, Nelson and Southland. Since 2006/07, the observer effort exceeded 3% only 3 times (2012/13, 2014/15 and 2016/2017) (Fig. 2).

This very irregular temporal distribution of observer coverage is an issue as soon as it is not taken into account by the modelling process. This work considers bycatch time series from 1995/96 to 2016/17, which could suggest a relative stability in observer cover. But a deeper exploration of available data seems to demonstrate important gaps and heterogeneity in observer programmes. As bycatch data are used in the modelling process as non-spatialized input data summed over the whole period, the absence of observer onboard and the absence of bycatch are mixed up.

### Spatial distribution
It seems that setnets equal or below 8m in length were never observed by observers (Fisheries New Zealand, 2022). This fishery was constituted of 120 to 130 vessels in 2020/21 that operated in semi-enclosed waters including harbours and estuaries (the total setnet fishery encompassed 170-190 vessels). If this fishery is unlikely to catch oceanic species, the very coastal habitat range of Hector's and Māui dolphins can make them vulnerable to this métier. The coastal and harbour distribution of dolphins is confirmed by public sighting observations and results provided by this report. Except engine power, métier characteristics would be comparable in terms of net size, soak time, and target species (Fisheries New Zealand, 2022).

Fishing areas of <8m setnets seemed to have a great overlap with both Hector's and Māui dolphins in Northern Island and South Eastern Island (Figs. 3 and 4). The absence of bycatch in setnets in Northern Island could therefore be due to the absence of at-sea monitoring in this area. This observation is all the more worrying as the West Coast of Northern Island
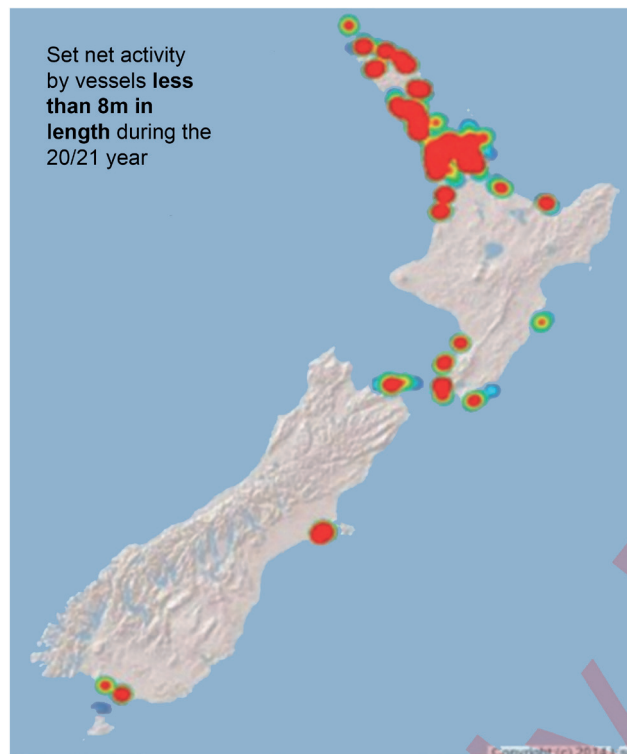
Fig. 3. Set net activity by vessels <8m during fishing season 2020/21 (source: Fisheries New Zealand, 2022).

constitutes the only remaining habitat of rare Māui dolphins. The jeopardizing impact of small- scale fisheries on coastal cetaceans was described in different occasions, despite the difficulties to observe their fishing effort and practices (Mangel *et al.*, 2010; Rojas-Bracho and Reeves, 2013). As fisheries observer data are used as input data, the spatial heterogeneity of observer covers and total absence of monitoring on 70% of vessels can artificially highlight and/or turn off some bycatch risk areas. The observer effort, in addition of being very low, is not randomly distributed across the fishery, both in terms of vessels and also fishing areas. If these conclusions are common with most of countries, they must be considered in analytical choices in order to integrate these biases in final estimates (Authier *et al.*, 2021; Rouby *et al.*, 2022).

**Observer biases**

Two main biases were identified in observer programmes: (i) the deployment effect, or non-random assignment of observers to vessels and ports due to the fact that accepting an observer on board is at the vessel master's discretion, and (ii) the observer effect, i.e. a change in fishing practices when an observer is present (Amandè *et al.*, 2012; Benoît and Allard, 2009; Faunce and Barbeaux, 2011; Stratoudakis *et al.*, 1998). These biases are likely to hinder the detection of bycatch by observers.

To sum up, the lack of information on observer cover hinders our possibility to discuss modelling strategy. Further investigations highlight important spatial and temporal heterogeneity, that may have important consequences such as underestimation of Hector's/Māui dolphin bycatch.

## 2. FISHERIES

By construction, the analysis presented in the report considers a static and constant fishing effort since 1995/96, including fishing practices and gear dimensions. As no information is available in the report about gear types (pelagic/bottom trawls? Gillnets/trammel nets?) and their temporal evolution, it remains delicate to assume con- sequences of such analytical choices, like using 2014/15 to 2016/17 as fishing effort paragon.

On Fisheries New Zealand website some available information on fishing effort showed around 50% reduction of net length between 2002/03 and 2019/20. Is this reduction related to changes in net dimensions or number of vessels? If related to changes in net dimensions, is this length reduction followed by an increase of net height, as observed in other countries? If so the catchability of dolphins could raise and consequences on dolphin bycatch could be worrying. If this decrease is related to a diminution of fishing vessels, it supposes that greater net length operated in New Zealand waters when the fisheries wasn't involved in observer programmes. This may have resulted in an important ignorance and underestimation
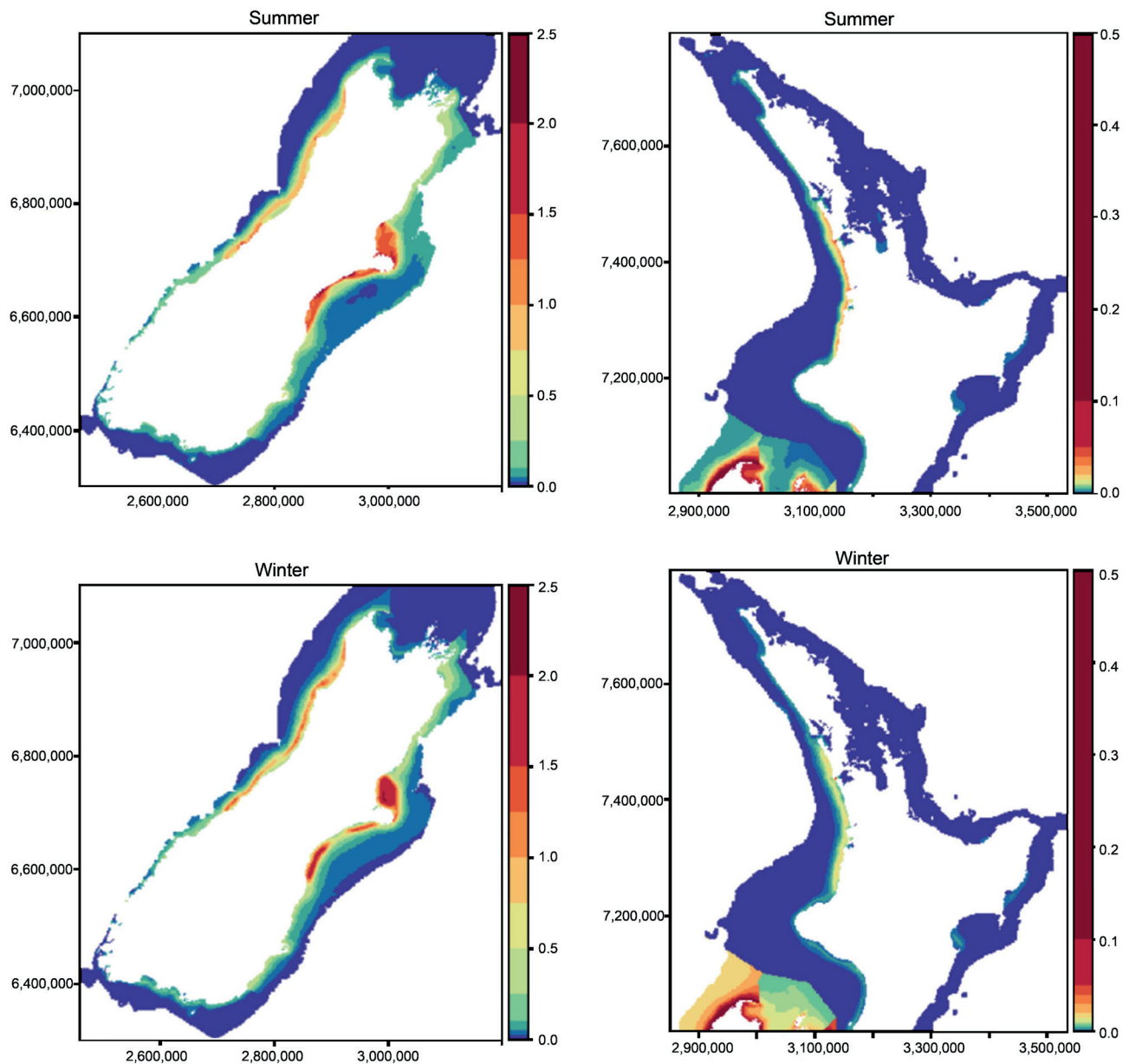
Fig. 4. Final seasonal estimated spatial density of Hector's and Māui dolphins in winter used for spatial risk assessment, North and South Islands.

Table A

Hector's and Māui dolphin annual deaths from various causes.

| | Mean annual deaths (5th to 95th percentile estimates) | | |
|---|---|---|---|
| | Commercial set net | Inshore trawl | Toxoplasmosis* |
| Māui | 0.10 (0-0.25) | 0.02 (0-0.05) | 1.9 (1.1-3.0) |
| Hector's | 44 (21-80) | 14 (1-43) | 334 (132-652) |

*It is important to note that commercial fisheries mortalities (set net and inshore trawl) are based on fisheries observer data and have been estimates with high certainty. Toxoplasmosis deaths have been estimates from necropsy results, which relies on the relative detectability of dolphin carcasses that have died from various causes, resulting in uncertainty that may not be reflected in the ranges above.

of former bycatch levels. The same pattern was detected in trawl fisheries, that would also require more information to describe the underlying process and its potential consequences on cetacean bycatch.

## 3. HECTOR'S AND MĀUI DOLPHIN BYCATCH

The underlying assumption of the model is that bycatch risk is directly correlated with fishing effort-dolphin distribution overlap. Has this direct link been demonstrated in the case of Hector's/Māui dolphin bycatch? The case of common dolphin in NE Atlantic (Bay of Biscay) suggests trophic interactions between dolphins and midwater pair trawlers. In winter, while dolphins share same preys (small pelagic fishes) as some target fishes (such as albacore tuna, common hake or common seabass), high levels of bycatch are recorded (ICES, 2023; Meynier *et al.*, 2008; Morizur *et al.*, 1999; Spitz *et al.*, 2013). The same gears target small pelagic fishes in spring, and no dolphin bycatch occurs. This example, and others across the world (Jog *et al.*, 2022) highlights the complexity of interactions between top predators and fishing gears, and questions the underlying hypothesis used in the model. This hypothesis of 'co-occurrence = causality' leads that any changes in fisheries and/or dolphin distribution will *de facto* generate important changes in risk ratio, maybe without any biological truth.

The distribution of Hector's dolphins appeared to have changed locally since early 2000's (Carome *et al.*, 2022), but main hotspots seemed relatively consistent over time (Brough *et al.*, 2019). I guess that distribution of Hector's/Māui dolphins is not available at yearly resolution. But due to modelling choices, the hypothesis of fisheries and dolphin distribution stability can have consequences on bycatch mortality estimates.

My other concern is the calculation of mortality related to different sources. The model estimates total mortality and subtracts commercial fishery deaths. Remaining mortality is proportionally attributed to other causes of death. By construction, all causes of death are linked and scaled to bycatch. The huge weight of bycatch estimates in this modelling approach requires even more precautions and clarification of fishery and observer cover dynamics.

## 4. ADDITIONAL COMMENTS

I have troubles understanding the choice of using PST instead of PBR as threshold. This choice is not in the line of conservative approaches that could be expected in circumstances of endangered and critically endangered sub-species. Previously estimated at 24 individuals for hector's dolphins (Slooten and Dawson, 2021), the PBR is well described and admitted by scientific community. The PST provide 3 times higher threshold for this species. The choice of thresholds other than PBR can be justified by different circumstances. In case of lack of data, a proportion of population as threshold can be considered (ex: 1.7% suggested by ASCOBANS in NE Atlantic for data deficient species). The modified PBR (mPBR), developed and used in context of small cetacean bycatch by OSPAR agreement (OSLO-PARIS Agreement in NE Atlantic), tends to be more conservative than PBR as it includes uncertainties around bycatch estimates (Genu *et al.*, 2021; Taylor *et al.*, 2022). In case of robust available data on abundance, demographic and anthropic threats, the choice is usually made to use the Removals Limit Algorithm as threshold.

Strandings were used in order to distribute remaining mortality according to necropsy results. This strong assumption supposes: (1) that cause of death doesn't infer on buoyancy of dolphins; (2) an equal probability of stranding/discovery according to the location of death at sea (and therefore cause of death in case of very spatially localised threats); and (3) a temporal stability of stranding/discovery probability over the study period in case of temporal changes in threats. In my opinion the proportion of individuals according to their cause of death can be used as temporal indicator of magnitude of threats, but probably not to estimate absolute deaths.

## REFERENCES

Amandè, M.J., Chassot, E., Chavance, P., Murua, H., Molina, A.D. and de Bez, N. 2012. Precision in bycatch estimates: the case of tuna purse-seine fisheries in the Indian Ocean. 69,. *ICES J. Mar. Sci. J. Cons.* 69: 1501–10. [Available at: https://doi.org/10.1093/icesjms/fss106].

Authier, M., Rouby, E. and Macleod, K. 2021. Estimating cetacean bycatch from non-representative samples (I): A simulation study with regularized multilevel regression and post-stratification. *Front. Mar. Sci.* 8. [Available at: https://doi.org/10.3389/fmars.2021.719956].

Benoît, H.P. and Allard, J. 2009. Can the data from at-sea observer surveys be used to make general inferences about catch composition and discards? *Can. J. Fish. Aquat. Sci.* 66: 2025-39. [Available at: https://doi.org/10.1139/F09-116].

Brough, T., Rayment, W., Slooten, E. and Dawson, S. 2019. Fine scale distribution for a population of New Zealand's only endemic dolphin (*Cephalorhynchus hectori*) shows long-term stability of coastal hotspots. *Mar. Mammal Sci.* 35: 140–63. [Available at: https://doi.org/10.1111/mms.12528].

Carome, W., Slooten, E., Rayment, W., Webster, T., Wickman, L., Brough, T. and Dawson, S.M. 2022. A long- term shift in the summer distribution of Hector's dolphins is correlated with an increase in cruise ship tourism. *Aquat. Conserv. Mar. Freshw. Ecosyst.* 32: 1660–74. [Available at: https://doi.org/10.1002/aqc.3881].

Faunce, C.H. and Barbeaux, S.J. 2011. The frequency and quantity of Alaskan groundfish catcher-vessel landings made with and without an observer. *ICES J. Mar. Sci. J. Cons.* 68: 1757–63. [Available at: https://doi.org/10.1093/icesjms/fsr090].

Fisheries New Zealand. 2022. Regulatory Impact Statement: Wider Rollout of On-board Cameras. Ministry for Primary Industries.

Genu, M., Gilles, A., Hammond, P.S., Macleod, K., Paillé, J., Paradinas, I., Smout, S., Winship, A.J. and Authier, M. 2021. Evaluating strategies for managing anthropogenic mortality on marine mammals: An R implementation with the package RLA. *Front. Mar. Sci.* 8.

ICES. 2023. Workshop on mitigation measures to reduce bycatch of short-beaked common dolphins in the Bay of Biscay (WKEMBYC2, outputs from 2022 meeting). *ICES Sci. Rep.* 5(3): 66pp. [Available from: https://doi.org/10.17895/ices.pub.21940337].

Jog, K., Sutaria, D., Diedrich, A., Grech, A. and Marsh, H. 2022. Marine mammal interactions with fisheries: Review of research and management trends across commercial and small-scale fisheries. *Front. Mar. Sci.* 9.

Mangel, J.C., Alfaro-Shigueto, J., Van Waerebeek, K., Cáceres, C., Bearhop, S., Witt, M.J. and Godley, B.J. 2010. Small cetacean captures in Peruvian artisanal fisheries: high despite protective legislation. *Biol. Cons.* 143: 136-43.

Meynier, L., Pusineri, C., Spitz, J., Santos, M.B., Pierce, G.J. and Ridoux, V. 2008. Intraspecific dietary variation in the short-beaked common dolphin *Delphinus delphis* in the Bay of Biscay: importance of fat fish. *Mar. Ecol. Prog. Ser.* 354: 277–87. [Available at: https://doi.org/10.3354/meps0724].

Morizur, Y., Berrow, S.D., Tregenza, N.J.C., Couperus, A.S. and Pouvreau, S. 1999. Incidental catches of marine-mammals in pelagic trawl fisheries of the northeast Atlantic. *Fish. Res.* 41: 297-307.

Rojas-Bracho, L. and Reeves, R.R. 2013. Vaquitas and gillnets: Mexico's ultimate cetacean conservation challenge. *Endanger. Species Res.* 21: 77-87. [Available at: https://doi.org/10.3354/esr00501].

Rouby, E., Dubroca, L., Cloatre, T., Demaneche, S., Genu, M., Macleod, K., Peltier, H., Ridoux, V. and Authier, M. 2022. Estimating Bycatch From Non-representative Samples (II): A Case Study of Pair Trawlers and Common Dolphins in the Bay of Biscay. *Front. Mar. Sci.* 8: 795942. [Available at: https://doi.org/10.3389/fmars.2021.795942].

Rowe, S.J. 2009. Conservation Services Programme observer report, 1 July 2004 to 30 June 2007. Marine Conservation Services Series 1, Department of Conservation. [Available for download at: https://www.doc.govt.nz/globalassets/documents/science-and-technical/sap250entire.pdf]

Slooten, E. and Dawson, S.M. 2021. Delays in protecting a small endangered cetacean: lessons learned for science and management. *Front. Mar. Sci.* 8: 606547.

Spitz, J., Chouvelon, T., Cardinaud, M., Kostecki, C. and Lorance, P. 2013. Prey preferences of adult sea bass *Dicentrarchus labrax* in the northeastern Atlantic: implications for bycatch of common dolphin *Delphinus delphis*. *ICES J. Mar. Sci. J. Cons.* 70: 452–61. [Available at: https://doi.org/10.1093/icesjms/fss200].

Stratoudakis, Y., Fryer, R.J. and Cook, R.M. 1998. Discarding practices for commercial gadoids in the North Sea. *Can. J. Fish. Aquat. Sci.* 55: 1632–44. [Available at: https://doi.org/10.1139/f98-029].

Taylor, N., Authier, M., Banga, R., Genu, M. and Gilles, A. 2022. *Marine Mammal By-catch, OSPAR, 2023: The 2023 Quality Status Report for the Northeast Atlantic*. OSPAR Commission, London.

# Appendix 6

# Spatial Risk Assessment of Threats to Hector's and Māui Dolphins in New Zealand

Reviewer: Michael E. Grigg, PhD

## 1. REVIEW OF AVAILABLE DOCUMENTS

In preparation for this review, the following documents contained information relevant to the review process for this report, all referenced below:

Roberts *et al.* (2019), Massie *et al.* (2010), Zulpo *et al.* (2018), Dubey (1995), Coupe (2021), Roe *et al.* (2013), Gibson and et al. (2011), Barbieri and et al. (2016), Miller and et al. (2004), Miller *et al.* (2008), Barbosa *et al.* (2015), VanWormer *et al.* (2014), Boothroyd and Grigg (2002), Kennard *et al.* (2021), Miller *et al.* (2023) and Anon. (2019).

## 2. REVIEW OF SPATIAL RISK ASSESSMENT OF THREATS TO HECTOR'S AND MĀUI DOLPHINS: NON-FISHERY ANTHROPOGENIC THREATS

The Scientific Committee of the IWC established a steering group to assemble independent experts to review aspects of the spatial risk assessment management framework developed by the New Zealand government. The following is a short-written review that assesses non-fisheries related anthropogenic threats to the management and recovery of coastal-dwelling Hector's and Māui dolphins of New Zealand. The review is divided into the following sections.

(1)  Overview of non-fishery anthropogenic threats with emphasis on infectious threats.
(2)  Use of beach-cast necropsies as a means of estimating non-fishery deaths.
(3)  The relative detectability of carcasses from animals dying of different causes, and resulting sources of bias (seasonal, spatial, factors affecting carcass buoyancy).
(4)  Implications of other evident patterns or biases for estimation of risk (sex or age bias, seasonal patterns).
(5)  Identifying data or research priorities to improve understanding of toxoplasmosis, recreational netting, climate change, other anthropogenic stressors.

### 2.1 Overview of non-fishery anthropogenic threats with emphasis on infectious threats

A range of human-induced threats have been identified affecting the health and resilience of New Zealand's native Hector's and Māui dolphins. Hector's dolphins number around 15,000 and are coastal-dwelling predominantly in waters around the South Island, whereas Māui dolphins are a critically threatened species (by one estimate, only ~63 animals left), also coastal-dwelling, and found on the West Coast of the North Island. Human-induced (anthropogenic) threats include fishing, infectious diseases (primarily toxoplasmosis), and mining activities. Historically, accidental drowning in commercial fishing nets was regarded as the primary anthropogenic threat to the sustainability of New Zealand's native dolphins. Risk assessment estimates for fishing are currently 1 Māui dolphin death every 9 years, and 58 Hector's dolphin deaths per year according to Roberts *et al.* (2019). What is not clear is how this estimate impacts the management and conservation of dolphin population recovery, and reproductive potential.

In the past decade, several publications highlighted the significance of infectious diseases as an important cause of mortality in dolphins. Specifically, toxoplasmosis, a protozoal disease caused by the cat parasite *Toxoplasma gondii* has emerged as the most significant threat. Cats defecate extraordinarily high numbers of environmentally stable, highly infectious oocysts that can be transported into the coastal environment by run-off from land sources; this parasite can, and does, infect any warm-blooded vertebrate, including dolphins, that ingest contaminated food or water. Coastal areas adjacent to large rivers, or near cities with high cat densities are thought to be specific hot spots for exposure to toxoplasmosis. Indeed, fatalities have been recorded at the mouth of the Waikato River on the North Island, as well as several rivers on the South Island. Not only does the parasite cause dolphin death, it can also cause reproductive loss and behavioural changes that impact neonate survival and dolphin susceptibility to predation.

According to the report provided by Roberts *et al.* (2019) toxoplasmosis deaths (estimated from necropsy results) appear to cause a ~6-17-fold higher number of mortalities than that estimated from commercial fisheries. Risk assessment estimates for toxoplasmosis are currently 1.9 Māui dolphin deaths per year, and 334 Hector's dolphin deaths per year – see Table A.

As indicated in the Table legend, it is important to clarify that toxoplasmosis deaths estimated from necropsy results are subject to greater uncertainty and possible bias, principally the result of capturing only a limited number of carcasses that strand. With that being said, the ability to perform a full pathological workup is invaluable because it allows investigators to examine a wide-range of variables that may impact the health and resilience of these iconic species. A targeted approach to increase the number of necropsy evaluations should reduce this uncertainty and this approach has been invaluable in the study (and conservation) of other threatened and endangered marine mammal populations, such as Hawaiian monk seals, and southern sea otters.

### 2.2 Use of beach-cast necropsies as a means of estimating non-fishery deaths

A key objective of the report by Roberts *et al.* (2019) was to provide an in depth synopsis of the various threats impacting Hector's and Māui dolphins since the last TMP (Threat Management Plan) risk assessment of 2007. The Roberts assessment includes several new inputs, including spatial abundance estimates for near-shore dolphins, a spatially resolved commercial fisheries dataset, and it incorporates a new *k* variable in its spatial risk model to represent non-fishery deaths due to infectious diseases, shark predation, and other anthropogenic variables.

This was done specifically to delineate between commercial fishery and non-fishery causes of death. The *k* variable addition largely reflects the awareness that infectious diseases have emerged as a significant cause of mortality, which was not systematically investigated or addressed prior to 2007.

Roe *et al.* (2013) utilized necropsy data from 57 bycatch or beach-cast dolphins recovered by the New Zealand Department of Conservation marine mammal stranding investigation response between 2007-11. This study poignantly highlighted the role infectious diseases are playing in the health and conservation of dolphins. It specifically established that the protozoan parasite *Toxoplasma gondii* was the most significant infectious disease risk factor contributing to dolphin mortality. Of the 57 recovered dolphins, however, 29 were excluded from necropsy due to decomposition, which is significant and could conceivably contribute to reporting bias because some infectious disease agents promote the decomposition process. Of the 28 dolphin carcasses of sufficient quality for full necropsy evaluation, 7 *Toxoplasma* fatal infections were identified: 5 in Hector's dolphins, 2 in Māui dolphins. Significantly, only a single genotype of *Toxoplasma gondii* was recovered despite wide geographic sampling. Whether parasite genotype is a significant variable in disease outcome has not been assessed, and this variable has proved critical in the success and recovery of other threatened marine mammal species, including the southern sea otter (Miller *et al.*, 2004; Kennard *et al.*, 2021). What information is known about *Toxoplasma* genetic diversity in New Zealand? The Type II strain identified that caused lethal infections in all dolphins analyzed from the Roe *et al.* (2013) study is a common strain, not atypical as claimed in their paper, that causes largely asymptomatic infections in humans in Europe and North America (Boothroyd and Grigg, 2002). In the additional samples included in the Roberts *et al.* (2019) report, was parasite genotype identified? This is an oversight if the analysis was done, but not included as a variable in the threat assessment analysis performed. Moreover, is the strain that was identified infecting and causing death in dolphins the same one that is largely circulating in cats and their prey in New Zealand? A PhD thesis published in 2021 by Coupe at Massey University suggests that this is not the case, as no Type II isolate was identified infecting cats. Clearly wider sampling is required, both in dolphins, and in surrounding near-shore animals to evaluate whether parasite genotype is a significant factor contributing to death in dolphins, as has been performed and concluded to be the case in other published studies (Miller *et al.*, 2008, VanWormer *et al.*, 2014).

Perhaps more surprising was the presence of parasite DNA in 17 of 28 dolphins (61%), including 1 neonate, and 1 pregnant female with a transplacental infection. This is high and would predict high environmental carriage of oocysts. I am not an expert on Hector's or Māui dolphins, so it is not clear to me whether validated tests exist to screen dolphin sera for circulating antibodies that indicate previous infection. Serology tests that screen for the presence of *Toxoplasma* antibodies should be applied to identify previous exposure, and hence, chronic infection. These are routinely performed on other marine mammals and could be used to identify the degree of chronic carriage in live sampled, as well as, stranded, beach-cast, or bycatch animals. This is significant for many reasons: (1) not all parasite infections result in dolphin mortality, but may impact their behaviour and/or risk of predation, which is not currently addressed in any of the disease risk models; (2) are the genotype(s) associated with asymptomatic chronic carriage different from the Type II strain associated with acute mortality in the Roe *et al.* (2013) study, which would suggest that parasite genotype is a significant risk factor; (3) the high prevalence rate indicates high exposure rates and suggests a significant environmental load; (4) transplacental transmission identifies the ability of this parasite to contribute to reproductive loss; and (5) omitting neonates from screening for *Toxoplasma* infection may greatly impact the population recovery and growth estimates, because this parasite certainly impacts both neonatal survival and the frequency estimates within the threat assessment threshold. Currently the majority of neonates are listed as 'died from maternal separation' and are not factored into the spatial risk assessment model.

Roberts *et al.* (2019) expanded the dataset from the Roe *et al.* (2013) study, to extend the necropsy data through 2012-18. During the 11-year period analysed (2007-18), a total of 76 dolphins underwent full necropsy examinations (Table A4-1). Twenty-one (21) were listed as calves, and these dolphins were not included in the spatial risk assessment (Table A4-2). Molecular screening of aborted calves from 3 separate studies carried out on marine mammal populations all previously identified a significant prevalence of protozoan infection in beach-cast calves (Gibson *et al.*, 2011; Barbosa *et al.*, 2015; Barbieri *et al.*, 2016), this is an oversight.

Of the 55 sub-adult and adult dolphins included, 12 were listed as commercial fishing- related deaths, whereas 19 were listed as infectious diseases-related deaths, with *Toxoplasma* infection accounting for the majority, or 9/19 (47%), of the mortalities (Table A4-3). The principal challenge that exists when considering the relative and contributing roles of fishery versus non-fishery-related deaths is: (1) how accurate are the estimates from the necropsy findings that are being applied to the population at large; and (2) how best to incorporate the necropsy results into the spatial risk models.

*Toxoplasma* is increasingly considered a significant risk to New Zealand dolphin population resilience, as it has been for other iconic marine mammal species, including Hawaiian monk seals, southern sea otters, beluga whales, and Caspian seals. This certainly highlights the necessity to test all dolphin carcasses, regardless of age or decomposition status, for the presence of toxoplasmosis, whether or not it is considered the primary cause of death. This will facilitate a more accurate baseline for the prevalence of infection. *Toxoplasma* tissue cysts are surprisingly resilient to the processes of decomposition, and parasite DNA can be successfully extracted from highly autolyzed tissue. Further, the extracted DNA will be available to assay for the presence other infectious agents (in the future) that may be important to investigate for trends occurring in real-time, such as epizootics, and whether disease agents identified and genotyped during a particular disease outbreak (that result in mortality) were present and circulating in the population prior to the epizootic. Also, current necropsy estimates likely under-represent the true number of dolphins succumbing to infectious disease-related effects, including reproductive fitness, maternal separation, pneumonia, myocarditis, and risk of predation. This is because neonates are not included in the spatial risk assessment estimates, as they are largely thought to die from malnutrition due to maternal separation. Moreover, this parasite is also a primary cause of pneumonia in many dogs and cats. In marine mammals, it is associated with causing myositis and myocarditis, other important contributing factors to dolphin health and sustainability, as tabled in Appendix 4, but not included as variables in the spatial risk assessment estimates.

Finally, because the parasite can impact neonate feeding behaviour, the reproductive fitness of females (7 out of the 9 deaths attributed to *Toxoplasma* were in reproductive-aged females), and other disease processes highlighted in Appendix 4, the Roberts *et al.* (2019) study likely underrepresents the true impact of *Toxoplasma* infection on the health and resilience of the New Zealand dolphin populations and their recovery.

In conclusion, available necropsy data has highlighted the role of infectious diseases in the health and resilience of the New Zealand dolphins. It also establishes the need to prioritize increased surveillance and the number of carcasses available for necropsy evaluation, as current datasets are subject to greater uncertainty, principally because too few carcasses have been assessed. Furthermore, all dolphin samples need to be assessed to limit potential bias, and a critical oversight was to exclude neonates from the risk assessment models. Further, the analysis should not be limited to assessing only primary cause death, as contributing factors that impact feeding behaviour, reproductive fitness, risk of predation, parasite genotype or other disease processes need to be considered when assessing the overall health and resilience of the species, and its ability to undergo population recovery.

### 2.3. Implications of other evident patterns or biases for estimation of risk (sex or age bias, seasonal patterns)

According to the Roberts *et al.* (2019) assessment, a strong seasonality of *Toxoplasma* exposure and death was observed – with all nine deaths occurring between October-December, which is also during the calving season, and when oocyst loading and transmission within the environment was predicted to be at its highest, due to seasonal storm events. Whether transplacental transmission has negatively impacted reproductive fitness was not assessed. A Spatial Threat Intensity for Toxoplasmosis was estimated using seasonally adjusted hydrological models and human-based population density values as a proxy to estimate cat density, an approach successfully adopted by VanWormer *et al.* (2014) to resolve high-risk areas for *Toxoplasma* exposure among threatened sea otters. This is a reasonable approach to identify those regions where a sustained, year-long effort to receive all beach-cast dolphins for necropsy evaluation will be informative to assess whether real differences exist due to seasonal or spatial biases in exposure potential. A major recommendation is to increase the sample size of necropsied dolphins to improve model parameters for estimating proportional causes of death and/ or exposure. Also factors based on: (1) loss of wetland; (2) impervious surfaces; (3) size of local cat populations; and (4) the detection of *Toxoplasma* oocysts in bio-sentinel populations (such as mussels, which are filter-feeders previously demonstrated to concentrate oocysts, Miller *et al.*, 2008) as a proxy to demonstrate the relative load of *Toxoplasma* oocysts in the environment should be incorporated into the spatial risk assessment model. Indeed, in the published PhD thesis from Coupe (2021) ~2% of mussel hemolymph samples collected from coastal field sites in key Māui dolphin habitats tested positive for *Toxoplasma*, indicating that the parasite is present in sufficient quantity in natural populations of mussels in coastal waters that represent key habitats for the New Zealand dolphins. What remains to be clarified is the source of infection and route of entry of the parasite into the dolphin population. Work performed in California, USA previously identified the uptake and transmission of *Toxoplasma* oocysts by migratory, filter-feeding fish (Massie *et al.*, 2010) so assessment of dolphin prey consumption may influence screening approaches to identify relevant sources for exposure.

In the Roberts *et al.* (2019) assessment, the question was raised whether dolphin sex was a significant variable predicting the outcome of infection. Certainly 7 out of 9 dead dolphins were female, but the number is too small to draw any definitive conclusions. Thus far, no strong sex bias has been identified in other marine mammal populations studied

that were exposed to *Toxoplasma* infection. In the study by Kennard *et al.* (2021), among sea otters infected with Type II strains (the same genotype that resulted in acute mortality in the New Zealand dolphins) more males were infected than females (11 vs. 3, respectively), but there was no statistical difference based on sex among infections that caused acute mortality. Importantly, the majority of Type II infected otters did not succumb to significant disease or result in acute mortality. This variable has not been assessed in the dolphins, because only those dolphins that died acutely from *Toxoplasma* were genotyped. Whether *Toxoplasma* infection is impacting the reproductive fitness of the species is perhaps a more germane point to focus on. We do not yet know the degree to which transplacental transmission resulting in congenital infection occurs, which may represent a pivotal variable that impacts population recovery, and this has not been incorporated into the spatial threat assessment. The assumption in cetaceans is that the majority of congenital infections occur horizontally, when a primary infection occurs in a non-immune pregnant female. However, in cattle, the dog parasite *Neospora caninum*, is transmitted predominantly vertically among immune dams that become immunosuppressed during pregnancy, resulting in a recrudescent infection that crosses the placenta to cause congenital infection and abortion, impacting reproductive fitness. In dogs, transplacental infection of *Toxoplasma gondii* often occurs with no clinical signs, but the parasite is detected and isolated in neonates. In cats, kittens born to infected queens during gestation are often infected with *Toxoplasma* transplacentally or via suckling. What has not been addressed in the risk assessment model is the extent to which transplacental infection is occurring in New Zealand dolphins, and whether this is impacting reproductive fitness among reproductive-aged females.

### 3. IDENTIFYING DATA OR RESEARCH PRIORITIES TO IMPROVE UNDERSTANDING OF TOXOPLASMOSIS, RECREATIONAL NETTING, CLIMATE CHANGE, OTHER ANTHROPOGENIC STRESSORS

In the Roberts *et al.* (2019) assessment, multiple tables and figures were generated (see table 14, figure 19, figures A12-2) that list a multitude of anthropogenic stressors and other variables potentially affecting the health and resilience of New Zealand dolphins. The variables were listed as frequency estimates, with 1.0 indicating greatest risk, and 0.0 lowest. But no ranking was based on the absolute number of cases assessed for each variable, which is problematic, as it can inflate the perceived threat. In absolute terms, the threat from, for example, an oil spill, or aquaculture, is relatively negligible, compared to other variables such as infectious diseases or fishery bycatch. My recommendation is to disclose absolute numbers for each variable, to better inform relative risk and prioritise research on the most relevant variables that impact dolphin recovery efforts.

Māui dolphin abundance estimates were strikingly different, depending on the estimate method utilized (Table 8 versus the population size estimate published by Baker *et al.*, 2016). This needs to be carefully addressed. The WCNI estimates (West Coast North Island) calculated for a seasonally adjusted population size for Māui dolphins was 3,690 in the summer, and 5,223 in the winter. These values are far in excess of the genetic marker-recapture based population size estimate of just 63 Māui dolphins from the Baker *et al.* (2016) study. For the purposes of evaluating conservation strategies that impact the recovery of the Māui dolphin population to a sustainable level, better agreement between these two datasets needs to be ascertained.

Given that *Toxoplasma gondii* is known to significantly impact the health and population recovery of other marine mammal populations (Miller *et al.*, 2004; Gibson *et al.*, 2011; Barbieri *et al.*, 2016) one approach may be to quantify population level risks associated with *Toxoplasma* by comparing the Potential Biological Removal Factor (PBR) (Wade, 1998) with the Potential Removal due to Protozoan Infection (PRP). PBR is typically used to assess or monitor population level impacts of anthropogenic associated mortalities in marine mammals - and is defined as the maximum number of animals that can be removed from a marine mammal stock, not including natural mortalities, while allowing that stock to reach or maintain an optimum sustainable population level according to the Marine Mammal Protection Act (MMPA) as published by Wade (1998). It is calculated as PBR=$0.5\ RMAXN_{MIN}F_R$, where $r_{MAX}$ is the theoretical (estimated) net productivity rate of the stock at a small population size, or in other words, the per capita rate of increase in a marine mammal [a default value of $r_{MAX}$= ~1.04 was used based on cetacean-commercial fishery interactions within the US]; and $N_{MIN}$ is the minimum population estimate, and $F_R$ is the recovery factor for the marine mammal stock in question.

One approach could be to assess the threat of protozoan parasites to marine mammal stock recovery. Assuming it is possible to define the 'Potential Removal due to Protozoans' (PRP) as $PRP = P_{TG}\ P_D\ R_{STR}\ N_{MIN}$ where $P_{TG}$ is the prevalence of the protozoan parasite, $P_D$ is the percent death in the marine mammal stock of those infected with the protozoan parasite, $R_{STR}$ is the annual rate of stranding, and $N_{MIN}$ is the minimum population estimate. It may prove possible to calculate both $P_{TG}$ and $P_D$ using annual (2007-18) information (prevalence and cause of death) from dead, stranded individuals in both the Hector and Māui dolphin populations. This will establish a baseline for ongoing comparative purposes to monitor resiliency and population growth and recovery. Taking the data from Roberts *et al.* (2019), it should be possible to calculate both a yearly and an overall average PRP and PBR for Hector's and Māui dolphin populations for the period 2007-18. But to do this successfully, and to inform on the relative conservation strategies enacted to promote species recovery, accurate minimum population estimates need to be established. Regardless of the methodology used, and the estimate produced, the approach above is capable of estimating change on a year-to-year basis and should establish whether individuals with a defined intrinsic growth rate are positively recovering.

According to Roberts *et al.* (2019), Māui dolphin population status and trajectory (page based using genetic capture-recapture observations produced median estimates consistent with declining (lamda = 0.980) or increasing (lamda = 1.025) population trends depending on how survival was parameterized (Baker, 2016a). In our unpublished work investigating the recovery of Hawaiian monk seals, we used this approach to calculate potential removal of seals by protozoan parasites (2001-15), and we calculated average intrinsic growth rates for Hawaiian monk seals at 6 northwest Hawaiian Islands (Kure, Midway, Pearl and Hermes, Lisianski, Laysan, French Frigate Shoals) based on estimates provided by Baker *et al.* (2011). In calculating the annual stranding rate for our population, we assumed that the majority of carcasses were recovered, which will not be the case. However, both the percent infected with *T. gondii* ($P_{TG}$) and the percent mortality associated with *T. gondii* ($P_D$) rely on the number stranded, which is directly reflected by the annual stranding rate, or number of carcasses recovered relative to the overall population. Assumedly, both the prevalence and mortality rate associated with a protozoan parasite, specifically *T. gondii*, will remain constant relative to the number of strandings if it is a true representation of both variables in the general population. Thus, while the PRP equation would rely on the number of recovered dead Māui dolphins relative to the reconstructed population size, it should provide an accurate risk of removal if the prevalence and rate of associated mortality remain constant. Regardless of the assumptions, this approach provides a means to quantify and monitor the risk over time associated with a protozoan parasite. It is also readily comparable to PBR to gauge the relative significance of the risk posed by protozoan parasites to a species such as the Māui dolphins.

In conclusion, the Roberts *et al.* (2019) report proposed a number of future research priorities to generate improved coastal toxoplasmosis risk assessment datasets, but they did not advocate increasing the number of necropsy's evaluated, which is my strongest recommendation:

(1) **To produce cat habitat suitability models with field-based surveys of stray and feral cat density**. The published thesis by Coupe (2021) has already identified seroprevalence rates of >60% in companion cats, established that 1.6% of feral/stray cats assayed were actively shedding oocysts and produced estimated oocyst load calculations of between 190-240 oocysts/sq. meter, which is significant, and in line with studies pursued in California, where documented land-to-sea transmission of *Toxoplasma* infection into threatened marine mammal species has been established.

(2) **Inform on coastal *Toxoplasma* oocyst density estimates by screening tissues of filter-feeding species collected from target at-risk coastal locations.** Again, the published thesis by Coupe (2021) showed that 13/104 mussels were PCR positive for *Toxoplasma*, and that within field sites in key Māui dolphin habitats, ~2% of mussels in the river terminal tested positive for *Toxoplasma.*

(3) **Assess whether parasite genotype impacts disease**. To reliably ascertain this, all stranded dolphins, regardless of age and disease presentation, should be tested by PCR and/or serology to determine *Toxoplasma* exposure, and genotype the infecting parasite to identify all strains of *Toxoplasma* capable of infecting New Zealand dolphins. Further, prey animals and cats upstream of river terminals frequented by New Zealand dolphins should be assessed for *Toxoplasma* infection status and genotype, to ascertain population genetic diversity of the parasite species in New Zealand, in general.

(4) **Determine the preferred prey species and whether diet preference increases the biological risk of infection and the development of acute toxoplasmosis.** This is critical information to generate, and to date, I have not seen anything published that informs on this proposed research mandate.

Finally, I found no information on proposed Live Capture Initiatives to assess health and serostatus in order to profile the risk of exposure to infectious disease agents that may impact the health and resilience of the New Zealand dolphin species. To my knowledge, no serology has been performed, nor has it been proposed to screen for the prevalence of infection across all dolphins, regardless of disease state, to generate baseline datasets. Nor has the ability of the parasite to cause congenital infection and alter reproductive fitness been assessed. Whether vaccination is possible, or being considered, ongoing discussions among care-takers of the critically endangered Hawaiian monk seal have expressed an interest in utilizing the S48 ToxoVax vaccine, that is routinely used in New Zealand among livestock populations, as one possible prevention initiative to consider.

## References

Anon. 2019. Fisheries New Zealand discussion paper 2019/03 ISSN No. 2624-0165.

Baker, J.D., Harting, A.L., Wurth, T.A., Johanos, T.C. 2011. Dramatic shifts in Hawaiian monk seal distribution predicted from divergent regional trends. *Mar. Mamm. Sci.* 27(1): 78-93. [Available at: https://doi.org/10.1111/j.1748-7692.2010.00395.x].

Baker, C.S., Steel, D., Hamner, R.M., Hickman, G., Boren, L., Arlidge, W. and Constantine, R. 2016. Estimating the abundance and effective population size of Māui dolphins using microsatellite genotypes in 2015–16, with retrospective matching to 2001–16. Department of Conservation Auckland.

Barbieri, M.M., Kashinsky, L., Rotstein, D.S., Colegrove, K.M., Haman, K.H., Magargal, S.L., Sweeny, A.R., Kaufman, A.C., Grigg, M.E., Littnan, C.L. 2016. Protozoal-related mortalities in endangered Hawaiian monk seals *Neomonachus schauinslandi. Dis Aquat Organ.* 121(2):85-95. [Available at: https://doi.org/10.3354/dao03047]. PMID: 27667806

Barbosa, L., Johnson, C.K., Lambourn, D.M., Gibson, A.K., Haman, K.H., Huggins, J.L., Sweeny, A.R., Sundar, N., Raverty, S.A. and Grigg, M.E. 2015. A novel Sarcocystis neurona genotype XIII is associated with severe encephalitis in an unexpectedly broad range of marine mammals from the northeastern Pacific Ocean. *Int. J. Parasitol.* 45(9-10): 595-603. [Available at: https://doi.org/10.1016/j.ijpara.2015.02.013].

Boothroyd, J.C. and Grigg, M.E. 2002. Population biology of *Toxoplasma gondii* and its relevance to human infection: do different strains cause different disease? *Curr. Opin. Microbiol.* 5(4): 438-42. [Available at: https://doi.org/10.1016/s1369-5274(02)00349-1].

Coupe, A. 2021. Investigating *Toxoplasma gondii* in the marine environment in New Zealand: from cats to kai moana (shellfish). PhD thesis, Massey University.

Dubey, J.P. 1995. Duration of immunity to shedding of Toxoplasma gondii oocysts by cats. *J. Parasitol.* 81(3): 410-5.

Gibson and et al. 2011. Polyparasitism is associated with increased disease severity in *Toxoplasma gondii*-infected marine sentinel species. *PLoS Neg. Trop. Dis.* 5(5): e1142.

Kennard, A., Miller, M.A., Khan, A., Quinones, M., Miller, N., Sundar, N., James, E.R., Greenwald, K., Roos, D.S. and Grigg, M.E. 2021. Virulence shift in Type X *Toxoplasma gondii*: natural cross QTL identifies ROP33 as rodent Vir locus. *BioRxiv - Pathology Early View*. [Available at: https://doi.org/10.1101/2021.03.31.437793].

Massie, G.N., Ware, M.W., Villegas, E.N. and Black, M.W. 2010. Uptake and transmission of Toxoplasma gondii oocysts by migratory, filter-feeding fish. *Vet. Parasitol.* 169(3-4): 296-303. [Available at: https://doi.org/10.1016/j.vetpar.2010.01.002].

Miller, M.A., Grigg, M.E., Kreuder, C., James, E.R., Melli, A.C., Crosbie, P.R., Jessup, D.A., Boothroyd, J.C., Brownstein, D., Conrad, P.A. 2004. An unusual genotype of Toxoplasma gondii is common in California sea otters (*Enhydra lutris nereis*) and is a cause of mortality. *Int J Parasitol.* 34(3):275-84. [Available at: https://doi.org/10.1016/j.ijpara.2003.12.008.] PMID: 15003489.

Miller, M.A., Miller, W.A., Conrad, P.A., James, E.R., Melli, A.C., Leutenegger, C.M., Dabritz, H.A., Packham, A.E., Paradies, D., Harris, M., Ames, J., Jessup, D.A., Worcester, K. and Grigg, M.E. 2008. Type X *Toxoplasma gondii* in a wild mussel and terrestrial carnivores from coastal California: new linkages between terrestrial mammals, runoff and toxoplasmosis of sea otters. *Int. J. Parasitol.* 38(11): 1319-28. [Available at: https://doi.org/10.1016/j.ijpara.2008.02.005].

Miller, M.A., Newberry, C.A., Sinnott, D.M., Batac, F.I., Greenwald, K., Reed, A., Young, C., Harris, M.D., Packham, A.E. and Shapiro, K. 2023. Newly detected, virulent *Toxoplasma gondii* COUG strain causing fatal steatitis and toxoplasmosis in southern sea otters (*Enhydra lutris nereis*). *Front. Mar. Sci.* 10. [Available at: https://doi.org/10.3389/fmars.2023.1116899].

Roberts, J.O., Webber, D.N., Roe, W.D., Edwards, C.T.T. and Doonan, I.J. 2019. Spatial risk assessment of threats to Hector's and Māui dolphins (*Cephalorhynchus hectori*). *New Zealand Aquatic Environment and Biodiversity Report No.* 214: 175pp.

Roe, W.D., Howe, L., Baker, E.J., Burrows, L. and Hunter, S.A. 2013. An atypical genotype of *Toxoplasma gondii* as a cause of mortality in Hector's dolphins (*Cephalorhynchus hectori*). *Vet. Parasitol.* 192(1-3): 67-74.

VanWormer, E., Miller, M.A., Conrad, P.A., Grigg, M.E., Rejmanek, D., Carpenter, T.E. and Mazet, J.A.K. 2014. Using molecular epidemiology to track *Toxoplasma gondii* from terrestrial carnivores to marine hosts: implications for public health and conservation. *PLoS Negl. Trop. Dis.* 29(8): 5. [Available at: https://doi.org/10.1371/journal.pntd.0002852].

Wade, P.R. 1998. Calculating limits to the allowable human-caused mortality of cetaceans and pinnipeds. *Mar. Mamm. Sci.* 14(1): 1-37.

Zulpo, D.L., Sammi, A.S., Dos Santos, J.R., Sasse, J.P., Martins, T.A., Minutti, A.F., Cardim, S.T., de Barros, L.D., Navarro, I.T. and Garcia, J.L. 2018. Toxoplasma gondii: A study of oocyst re-shedding in domestic cats. *Vet. Parasitol.* 249: 17-20. [Available at: https://doi.org/10.1016/j.vetpar.2017.10.021].

# Appendix 7

# Spatial Risk Assessment of Threats to Hector's and Māui Dolphins

Reviewer: Brian Brost

Roberts et al. use a spatially-explicit fishery risk assessment (SEFRA) approach to estimate risk to Hector's and Maui dolphins from fishery and non-fishery sources. The SEFRA approach is unique because it accounts for risks and populations that are heterogeneous in space. The effort invested in its application to Hector's and Maui dolphins is evident. My primary concern relates to the general SEFRA model specification and structure, as well as the lack of available data to inform some of the necessary quantities. I had difficulty following the description of the methods, terminology, and notation. I apologize if any of my comments are consequently off the mark.

The Poisson likelihood presented for the bycatch portion of the SEFRA approach is a function of 4 parameters: dolphin abundance ($N_s$), vulnerability ($v_g$), the probability bycatch is observable ($p_g$), and the probability that a captured dolphin is alive ($\psi_g$). The dolphin bycatch data contain no information about $N_s$ and $p_g$, so it's no surprise their posteriors and priors match (pg. 48). Any claim to "estimate" these parameters (e.g., pgs. 22 and 23) is misleading because they aren't identifiable. Even if the data were informative, there are some structural issues with the SEFRA model specification. Consider $y = \alpha \times \beta$ as an example. No matter how much data ($y$) are available, it's impossible to distinguish between $\alpha$ and $\beta$—any given change in $\alpha$ can be offset by a change in $\beta$. It's likely that $v_g$ and $p_g$, in particular, are similarly confounded in this model.

Estimating $\psi_g$ and $v_g$ are separate objectives and use the bycatch data in different ways. It's better practice (and more transparent) to tackle each objective in turn using appropriate error structures. For example, estimating $\psi$ involves a comparison of dead vs. live captures. Such a model is very easy to implement. In fact, given the prior specification in Table 5, obtaining 10,000 samples from the posterior distribution for $\psi$ in the set net fishery is a single, short line of R code: rbeta$(10000, 3 + 1, 12 + 3)$. This model does not use a Poisson likelihood or involve $p_g$.

Estimation of vulnerability ($v_g$), on the other hand, involves modeling total captures as a function of fishing effort and dolphin population dynamics. A sensible model would use a binomial distribution with parameters $N_s$ and $v_g$, with $v_g$ subsequently modeled as a function of spatially-resolved fishing effort and dolphin density (but not $p_g$). Within the MCMC algorithm, uncertainty in dolphin abundance is propagated by conditioning on realizations from $[N_s]$. That said, I assume dolphin abundance and density vary over the 21-year period encompassed by the dolphin capture data, yet abundance and density information aren't available annually. I don't recall discussion concerning the assumption that these quantities are static through time. Note that this objective also relies on observer deployments that follow some known sampling design (e.g., random deployments across the entire fishery), which isn't discussed in the report, either. See below for additional comments regarding the authors' model specification for $v_g$.

Aside from the inference concerning non-fishery deaths (i.e., $[\boldsymbol{\rho}|\cdot] \propto [\boldsymbol{\eta}|\boldsymbol{\rho}][\boldsymbol{\rho}]$), everything else in the SEFRA approach is an exercise in stochastic simulation. It's not estimation or statistical modeling, Bayesian or otherwise. It's incorrect to label quantities calculated during this simulation exercise as posterior distributions (e.g., $R$ and $N$ on pg. 7).

This simulation exercise occurs independent of the MCMC algorithm(s). It includes samples from the posteriors of $\psi_g$, $v_g$, and $\boldsymbol{\rho}$, as well as random deviates from the presumed distributions for the quantities $N_s$, $\omega_g$, $r_s^{max}$, $S_s^{1+}$, $p_g$, and $k_g$. Because the quantities $N_s$, $\omega_g$, $r_s^{max}$, $S_s^{1+}$, $p_g$, and $k_g$ are not directly modeled as a function of dolphin data, the output of such a simulation is only as good as the many presumed distributions. I appreciate the authors' effort to define distributions that reflect the current state of knowledge for some quantities (e.g., $N_s$); however, there are others that are only assigned vague distributions due to limitations in available information (e.g., $p_g$ and $\omega_g$; note that neither parameter is actually estimated in the SEFRA approach). The assumed distribution for $p_g$ seems particularly tenuous and, as the authors note on page 71, the results are sensitive to this quantity. Within this simulation exercise, non-fishery mortalities are a function of commercial fisheries bycatch (pg. 28). Therefore, the quality of output concerning deaths due to toxoplasmosis, predation, etc. is closely linked to (and only as good as) the methods used to generate commercial fishery deaths (in addition to presumed detection probabilities for non-fishery deaths).

## Additional comments

*Methods*

1. I found no information in the report on observer coverage (i.e., the proportion of the fishery that's observed for bycatch) or how observers were deployed. Without this information, it's impossible to know the quality of the available data. In particular, if observers weren't deployed under some design (e.g., randomly), the potential for quantifying the relationship between dolphin bycatch and covariates (e.g., fishing effort and dolphin density) may be limited. Comparing observed and estimated mortalities in the commercial fisheries, observer coverage must be extremely low (13 observed mortalities over 21 years vs. an estimated ~40 deaths per year; pgs. 42 and 50).

2. The authors describe vulnerability ($v_g$) as the "probability of capture or death" (pg. 22). Aren't the probability of capture and the probability of death two different things (hence $\psi_g$)? Furthermore, probabilities have support over $[0, 1]$, but the support for $v_g$ is $[0, \infty)$. Consequently, the parameter either isn't described correctly as a probability or the prior distribution is inappropriate. Considering units of measurement alone (the units for $\lambda$ is number of dolphins), I don't believe $v_g$ can actually be a probability. It's also likely that $v_g$ is structurally non-identifiable given the existing model specification (see above).

3. I don't quite follow how the surfaces representing predicted coastal and harbour abundance were merged (pg. 19). The process of scaling the surfaces, merging, and then rescaling again seems ad hoc. Did the authors consider melding density surfaces in a model-based fashion (Conn et al. 2022)? Such an approach may allow for the inclusion of additional information concerning the spatial density of dolphins, if available.

4. There appears to be some confusion between Bayesian modeling and stochastic simulation (pg. 23 and elsewhere). Everything in a Bayesian analysis is a random variable; however, only a small component of this work is actually Bayesian. The rest is post hoc simulation. Moreover, many of the parameters described as "priors" aren't priors in the Bayesian sense of the word (pgs. 23–27). Presentation of the methods would be much clearer if the terminology was cleaned up and the statistical models were described separately from the simulation exercise.

5. Why is cryptic mortality $(k_g)$ equal to the inverse of detection probability (pg. 23)? An event is either observable or it isn't, so shouldn't these two quantities sum to 1?

6. Shouldn't the number of deaths resulting from commercial fisheries depend on $p_g$ and $k_g$? Neither of these quantities enters the equation for $D_{sg}$ at the top of page 28.

*Results*

1. The dolphin habitat model containing turbidity and sprat is clearly the AIC-best model (assuming d-AIC represents $\Delta$AIC in Table 7, pg. 32). Why was an inferior model (based on AIC) selected instead? The quality of Figure 9 is poor, but the "ahuru" model appears to have substantial lack of fit.

2. The unrealistic dolphin abundance estimate for the WCNI undermines the associated methods (pg. 34). I also don't understand why this high abundance estimate is later referred to as a "carrying capacity" (pg. 62). Isn't the intent to estimate current abundance, not carrying capacity?

3. Why does the estimated density for dolphins around the northwest corner of the South Island differ between Figures 17 and 18 (pgs. 40–41)?

4. Human density was used as a surrogate for toxoplasmosis risk (pg. 44). I would guess detection of dolphin strandings is also related to human density, which could bias the sample of necropsied dolphins in favor of toxoplasmosis mortalities. I acknowledge stranding recovery rates are somewhat arbitrary (e.g., the predation sensitivity scenario), but I wonder why a positive bias in toxoplasmosis mortalities wasn't investigated.

5. Several observed captures appear to occur in areas of low risk (pg. 48 and Figure A15-3), though poor figure quality makes it difficult to assess the agreement between the distribution of observed and predicted captures. Were any formal goodness-of-fit diagnostics performed? Visual inspection of maps, etc. is a good starting point for checking model adequacy, but it's not a substitute for formal diagnostics. See Conn et al. (2018) for methods to assess model fit.

6. There's reference to vessel-reported dolphin captures on pg. 48. Is it possible to use these additional data in models of dolphin bycatch?

7. Assessments of risk through time (overlap, deaths, and risk ratio; pg. 56) require dolphin abundance and density surfaces that vary annually. As far as I understand, those data are not available at an annual timescale?

*Discussion*

1. Why is it necessary to assign an arbitrary number of animals to areas without a known, permanent dolphin population (pg. 59)? I also don't understand how number of deaths is insensitive to the arbitrary choice in population size.

2. I agree that the SEFRA approach relies on good knowledge about key quantities (pg. 71); however, I don't follow how $v_g$ can "soak up any mistakes" made in the prior specification for $p_g$. These parameters are "estimated" jointly and the data contain no information concerning $p_g$. This seems problematic because predicting deaths relies "heavily" on the prior specification for $p_g$.

## References

Conn, P.B., Johnson, D.S., Williams, P.J., Melin, S.R. and Hooten, M.B. 2018. A guide to Bayesian model checking for ecologists. *Ecological Monographs* 88(4): 526-542. [Available at: https://doi.org/10.1002/ecm.1314].

Paul B. Conn, Jay M. Ver Hoef, Brett T. McClintock, Devin S. Johnson, Brian Brost. 2022. A GLMM approach for combining multiple relative abundance surfaces. *Methods in Ecology and Evolution* 13(10): 2236-2247. [Available from: https://doi.org/10.1111/2041-210X.13948]

# Appendix 8

# Additional Information Provided by the New Zealand Government at the Pre-Meeting Workshop

The New Zealand Government

## 1. RELEVANT MAPS



Fig. 1. Current set net fishing restrictions and closed areas around the North.

## West Coast North Island
## Trawl Fishing Restrictions

Maunganui Bluff

4 Nautical Miles
from Shore

Kaipara Harbour

Manukau Harbour

Waikalo River

4 Nautical Miles
from Shore

Raglan Harbour

Aotea Harbour

Kawhia Harbour

Auckland and Keimadoc Fishing Regulations Area

Central Fishing Regulations Area

Tirua Point

4 Nautical Miles
from Shore

Waiwhakaiho River

New Commercial
Trawl Closure (2020)

Existing Commercial
Trawl Closure

Marine Reserve

Submarine Cables and
Pipelines Protection Area

Fig. 2. Current trawl fishing restrictions and closed areas around the North Island.

Fig. 3. Set net restrictions and closed areas around the South Island, including the December 2022 expansion around Banks Peninsula.

| Point | Latitude | Longitude |
|-------|----------|-----------|
| A | 43°33.340'S | 173°18.178'E |
| B | 44°01.929'S | 172°37.374'E |
| C | 43°56.275'S | 172°45.382'E |
| D | 43°49.986'S | 173°11.770'E |
| E | 43°49.986'S | 173°16.201'E |
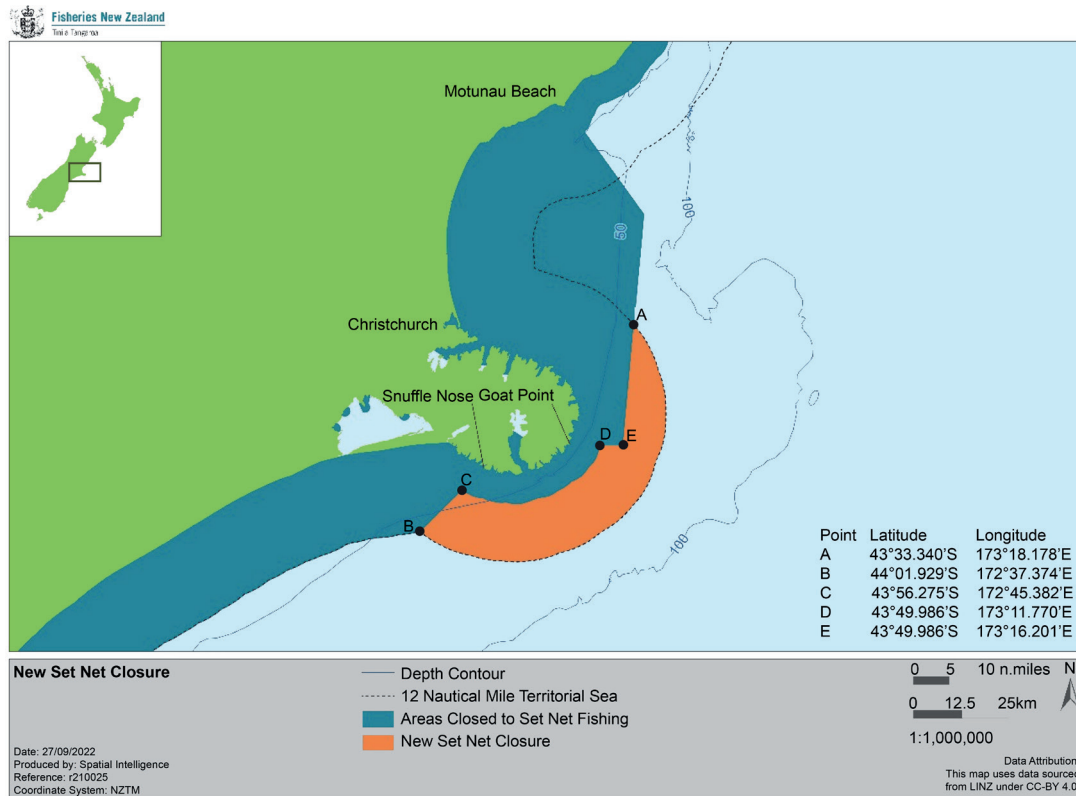
Fig. 4. Commercial and recreational set net closure implemented around Banks Peninsula in December 2022.
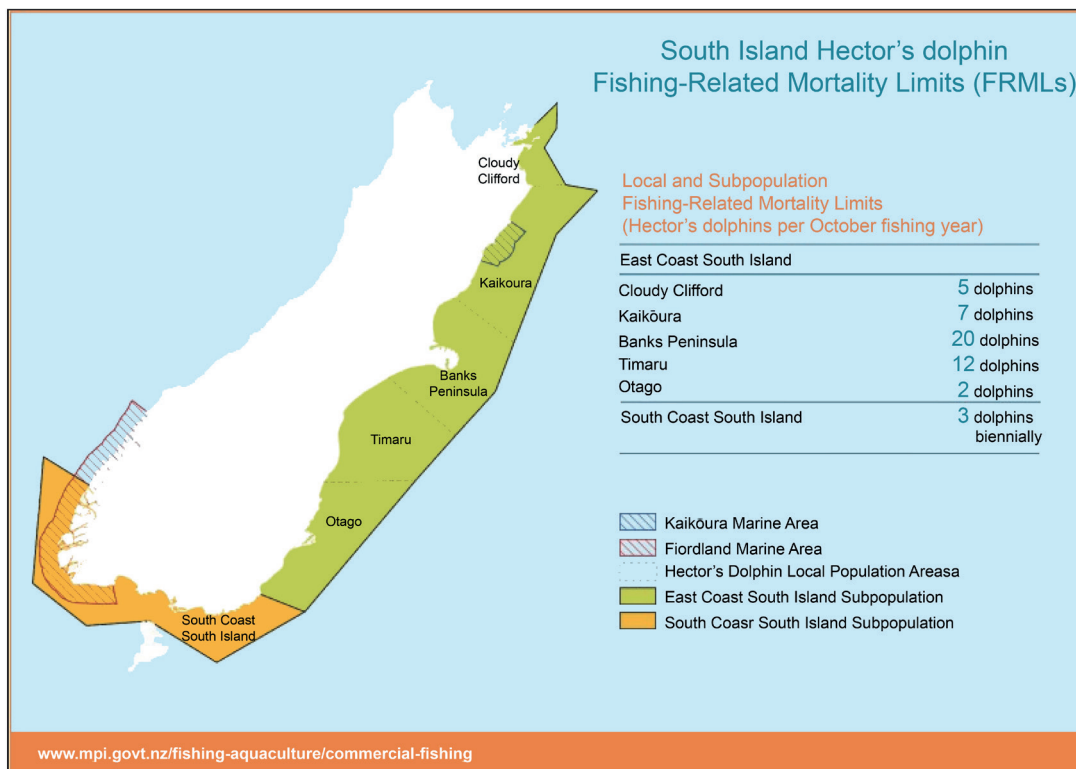


Fig. 5. Fishing-related mortality limits for Hector's dolphins around the South Island.
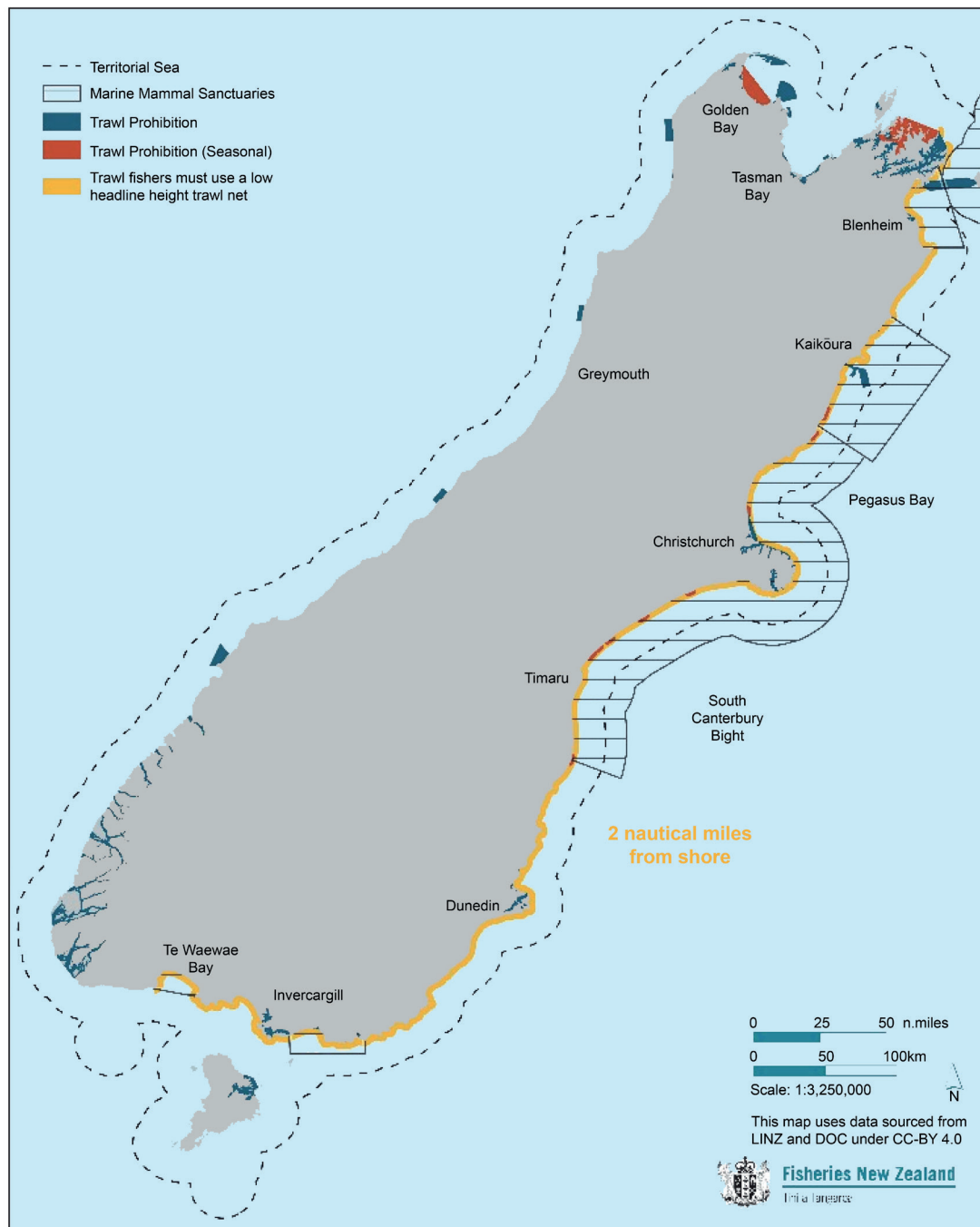
Fig. 6. Current trawl restrictions and closed areas around the South Island.
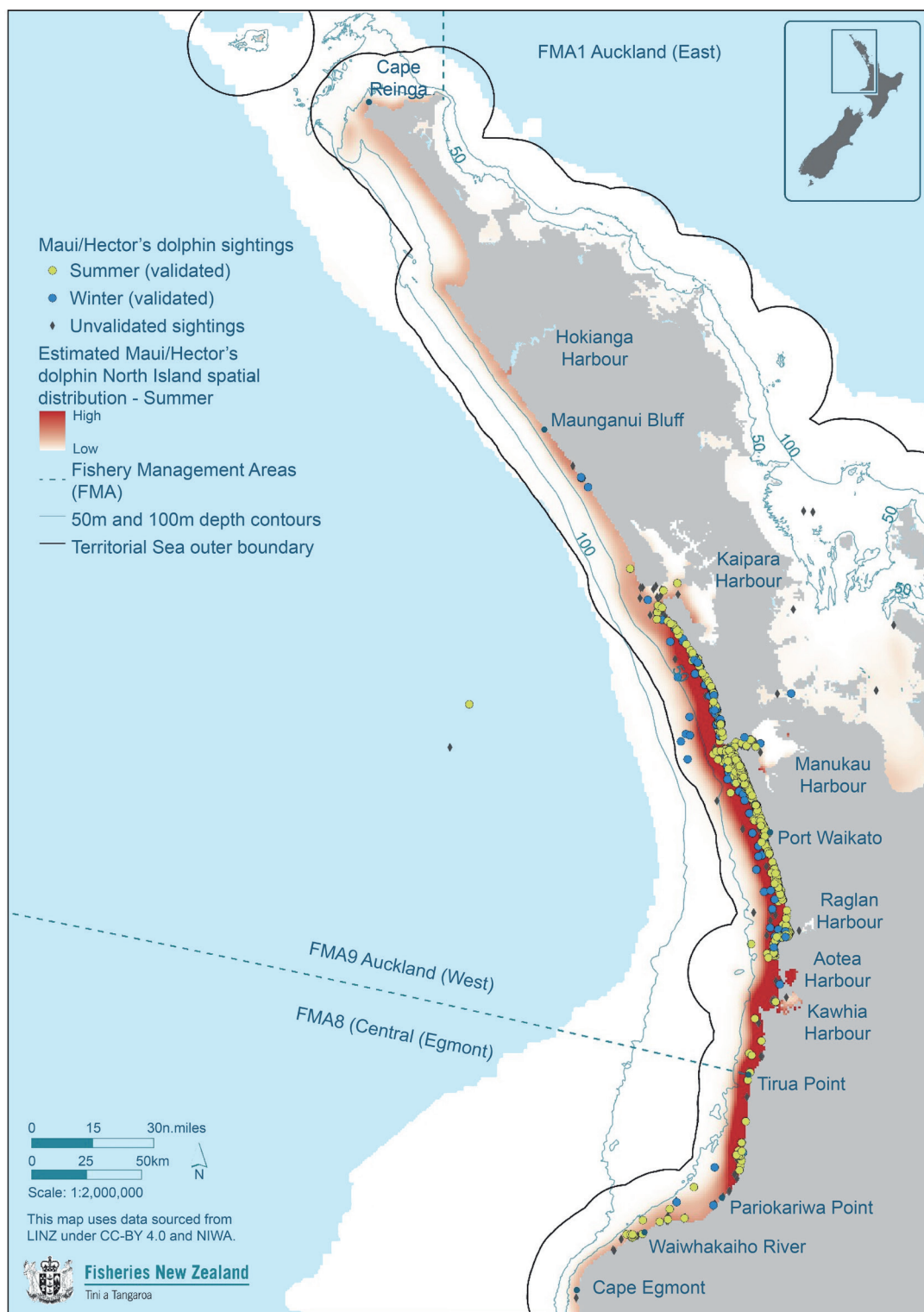
Fig. 7. Estimated (summer) spatial distribution of Māui dolphins from Cape Reinga to Cape Egmont. Locations of all reported public sightings of Māui/Hector's dolphins across the estimated Māui dolphin distribution. Yellow=validated summer sighting; Blue=validated winter sighting; Grey cross=unvalidated sighting. The 50m and 100m depth contours are also shown.
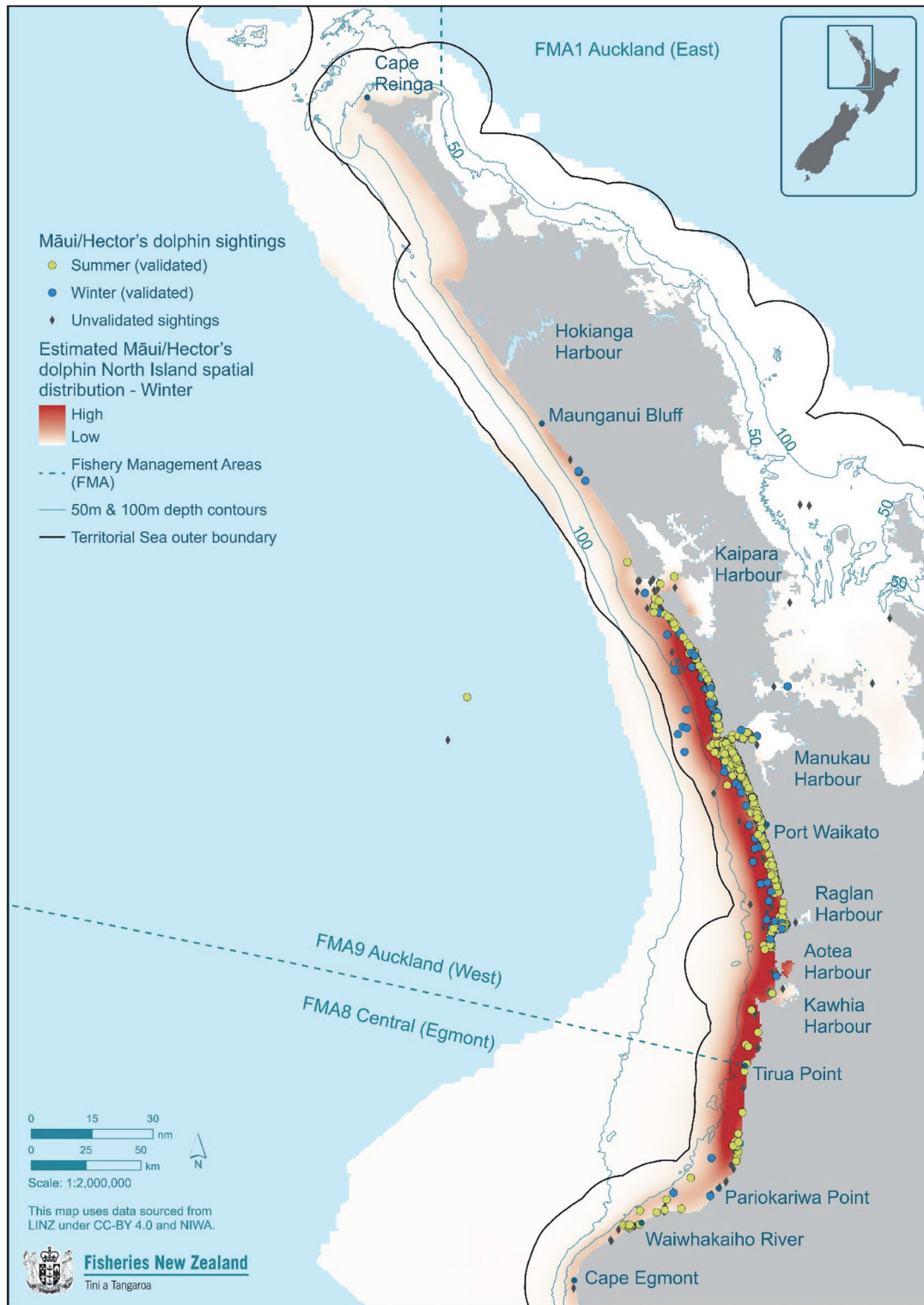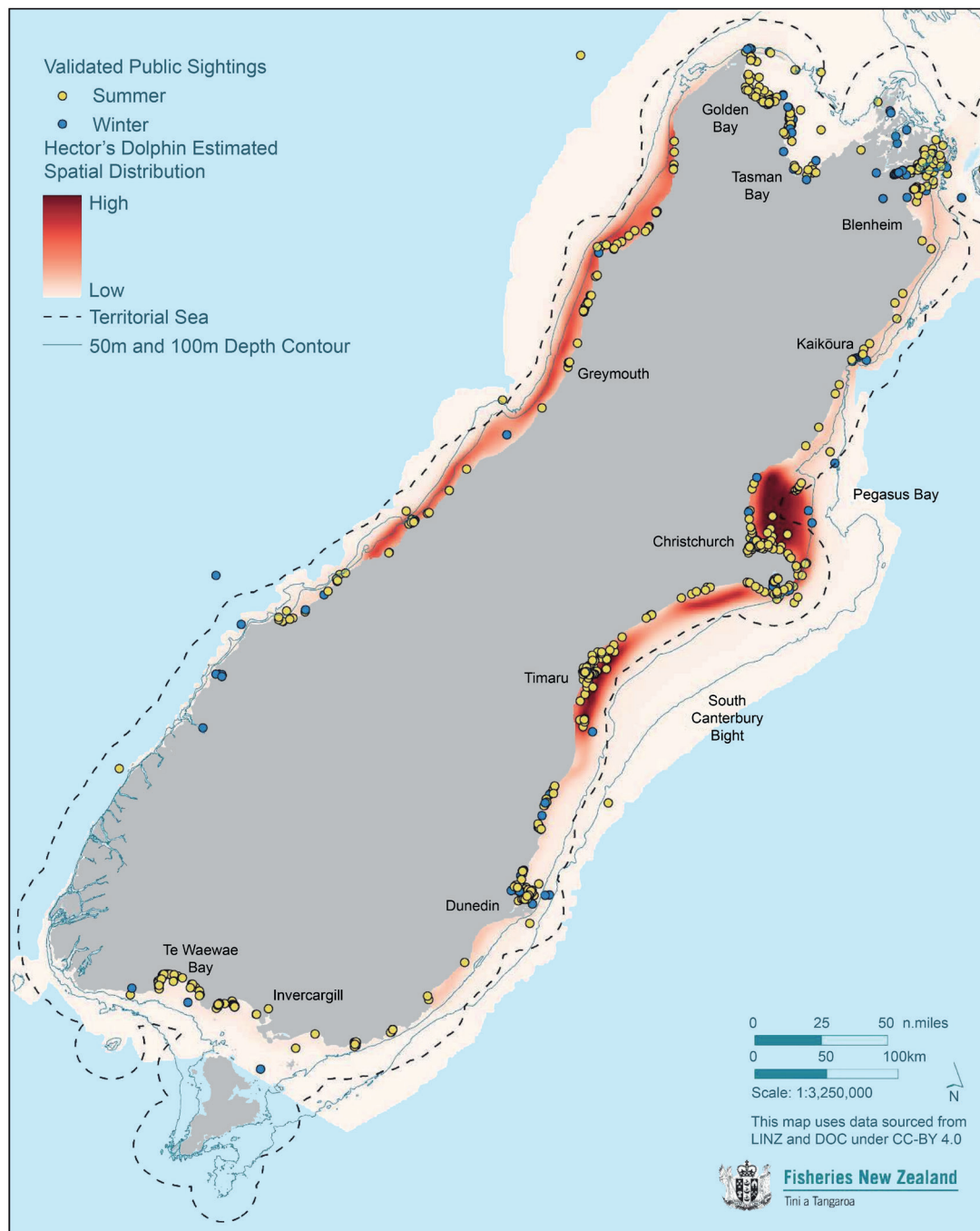
Fig. 8. Estimated (winter) spatial distribution of Māui dolphins from Cape Reinga to Cape Egmont. Locations of all reported public sightings of Māui/Hector's dolphins across the estimated Māui dolphin distribution. Yellow=validated summer sighting; Blue=validated winter sighting; Grey cross=unvalidated sighting. The 50m and 100m depth contours are also shown.

Fig. 9. Estimated (winter) spatial distribution of Hector's dolphins, including validated public sightings (summer sightings in yellow, winter sightings in blue). The 50 and 100m depth contours are also shown.

## 2. ESTIMATION OF $R_{MAX}$

This document provides additional information and suggestions in response to document SC/69A/MD/WP/01, entitled 'Considerations regarding an estimate of $r_{max}$ for Hector's and Māui dolphin', authored by Doug DeMaster, in support of an independent review by the IWC of a spatial risk assessment of threats to Hector's and Māui dolphins (Roberts *et al.*, 2019).

The approach used by Roberts *et al.* (2019) to estimate $r_{max}$ was an update of Edwards *et al.* (2018), which was a modification of the approach developed by Dillingham *et al.* (2016). Briefly, this approach uses a matrix population model to generate estimates of λ when drawing from user- selected 'optimal' distributions of key demographic rates, and then utilises an invariant observed across all vertebrates between $r_{max}$ and optimal generation time (*Topt*) to filter out combinations of demographic inputs that would be implausible given the invariant. The implementation of this approach by Edwards *et al.* (2018) was favourably reviewed by Peter Dillingham (Appendix 1 of this report).

Below, we provide some additional information and highlight some key points for discussion at the workshop in Bled. For clarity, $r_{max} = \ln(\lambda_{max})$, and $r_{max} = \lambda_{max} -1$ (noting that $r_{max}$ and $\lambda_{max}$ appear to have been mixed up in parts of the review):

(1)  The stated overall objective of the review was 'to provide a summary of approaches that could be used to estimate $r_{max}$ in the management of the Hector's and Māui dolphin, and to provide a recommendation as to which approach or approaches are most appropriate from the perspective of stewardship towards this population'. Notably, this part of the review does not list the use of invariants as one of the available options for estimating $r_{max}$ (e.g., (Dillingham *et al.*, 2016), despite its use by Roberts *et al.* (2019), other marine mammal assessments for NZ species (see table 36 of MacKenzie *et al.* (2022)), and some US marine mammal species (Moore, 2015; Moore *et al.*, 2018).

**We suggest that invariant approaches for estimating $r_{max}$ should also be discussed by the workshop in consideration of what is most appropriate for Hector's and Māui dolphins.**

(2)  The review sought to 'pay special attention to the estimation procedure (i.e., the demographic rates used) and application of $r_{max}$ in the Roberts *et al.* (2019) publication'. We found no explicit appraisal of the approach or the inputs used in the review. For example, there is no discussion or critique of the derivation of age at first breeding (a key input for this approach), despite the use of a novel approach for estimating an input distribution for this parameter.

**We suggest that workshop participants review the approach used by Roberts *et al.* (2019), including the relevant input parameters (see next point), in order to have an informed discussion about these methods and their use for Hector's and Māui dolphins. This will be critical to achieving the overall objectives of the review, above.**

(3)  The review summarised demographic input values in lines 135-143 and Appendix 1, but there seems to be some confusion about the values used. For clarity, the parameter distributions used are described below, with information of how they were derived. Note that these are 'optimal' distributions, as would occur when populations are growing at $r_{max}$:

Non-calf annual survival - used a beta distribution with mean=0.923 and sd=0.033. This was derived from a multi-area demographic assessment of Banks Peninsula Hector's dolphins by DuFresne (2004), which obtained higher estimates of annual survival when movement between areas was accounted for.

Calf annual survival - there are no available estimates of survival for calves to age 1 for this species. Hence, this was set equal to non-calf survivorship times a calf survivorship multiplier sampled from the uniform distribution bounded at 0.75-0.90 (as implemented by Taylor *et al.* (2007).

Longevity - no maximum longevity was specified by the base runs of either Edwards *et al.* (2018) or Roberts *et al.* (2019), although a maximum longevity of age 30 was trialled as a sensitivity by the latter of these, affecting a very small increase in estimated $r_{max}$ (caused by a minor shortening of the optimal generation time). For context, this compares with a maximum published age for Hector's dolphins of 26. Also, captive Commerson's dolphins have been observed living to age 33.

Age at maturation - estimated a posterior with mean=age 6.92 (95% CI=6.06-7.85), using existing necropsy-based age-maturation stage observations from the species and a novel prior derived from a meta-analysis of dolphins and porpoises of female length at sexual maturity as a proportion of asymptotic length (see Edwards *et al.* (2018) for a description of this).

Inter-birth interval - assumed to be uniformly distributed between 2 (the minimum - equivalent to becoming pregnant every year following a one-year gestation) and 2.5 (the approximate mode estimated by Gormley (2009) from calving rate observations of 48 female Hector's dolphins).

**We note that in order to fully review the method used by Roberts *et al.* (2019) to estimate $r_{max}$, it is necessary to understand how these input values were derived. Therefore, we suggest workshop participants read the descriptions of these derivations in advance so the workshop can focus on any suggested improvements.**

(4)  The review provides a summary of $r_{max}$ values for Hector's and Māui dolphins and other small cetacean species from the literature (see Table 1 of the review). However, of the examples listed in Table 1: only three of the studies directly

sought to estimate $r_{max}$ for the species (Edwards *et al.*, 2018; Roberts *et al.*, 2019; Slooten and Lad, 1991); two studies estimated something akin to $r_{max}$ using population models including estimated fishery captures, although both relied on assumed values of carrying capacity, about which the data were not informative (Cooke *et al.*, 2019; Davies *et al.*, 2008); and the remainder recycled estimates from other studies or used sensitivity values that were not estimated. The list also omits some other studies that independently estimated $r_{max}$ for this species, e.g., MacKenzie *et al.* (2022) estimated $r_{max}$=0.045 (95% CI=0.024-0.070) for Hector's and Māui dolphin using the approach of Dillingham *et al.* (2016), and Abraham *et al.* (2017) estimated 0.023 (95% CI=0.015-0.034) for Māui dolphin based on expert opinion. We also note that Table 1 presents a limited subset of the available studies estimating $r_{max}$ for other small cetacean species. For example, see tables 33-36 of MacKenzie *et al.* (2022), which show the derivation of $r_{max}$ for all NZ marine mammal species, and also table 37, which shows the US stocks for which non-default values were used (all tables are copied at this end of this text).

**Since comparison is used for justification of some of the main conclusions of the review, we have provided here data from other studies which independently estimated $r_{max}$. We also provided a range of other $r_{max}$ estimates for other small cetaceans for comparison.**

(5) A key difference between the PST approach (used by SEFRA) and the PBR approach relates to the treatment of uncertainty and bias. The PST approach represents uncertainty in input parameters, leaving managers to decide on an acceptable probability of it being exceeded, given estimated deaths. By comparison, the PBR approach uses point values of input parameters (rather than distributions), and relies on the selection of precautionary values to account for uncertainty and potential bias in their estimation.

(6) Typically, PBR assessments will use precautionary values of $r_{max}$ for a stock, e.g., the maximum *observed* growth rate for a population, if the default values of Wade (1998) are not used. By comparison, the PST approach requires a *distribution* of $r_{max}$ that is representative of uncertainty of its true value at both the lower and upper ends. Note that, in a SEFRA assessment, if management is made based on an upper percentile of risk, then this will largely be decided by the shape of the lower tail of the $r_{max}$ distribution.

(7) The review advises against using an $r_{max}$ value [actually $\lambda_{max}$] of 1.049 or higher for Hector's or Māui dolphin populations, since this would require 'aggressive estimates of age-specific reproductive and survival rates'. A theoretical rationale was also given that lower values may be applicable for Māui dolphin, given their location at the northern distributional limits for the species and the apparent relative lack of prey across their range (Roberts *et al.*, 2019). The text under point 5 is relevant here also.

(8) The review states that 'in recommending an $r_{max}$ value [actually $\lambda_{max}$] between 1.03 and 1.04 for the purpose of evaluating conservation strategies to recover populations of Hector's or Māui dolphin' that this range is 'plausible given the life history of this species and other odontocete species'. This was presumably based on the comparison of $r_{max}$ estimates presented in Table 1, which we have suggested amendments to above.

**We note the advice above is generated subjectively rather than quantitatively. Therefore, we suggest some discussion may be warranted to assess the general level of support for this advice among workshop participants, in light of the information presented in the review and suggestions made above.**

### TABLES OF LIFE HISTORY PARAMETER INPUTS, SOURCES, AND ESTIMATED $R_{MAX}$ VALUES FOR NEW ZEALAND MARINE MAMMAL SPECIES, FROM MACKENZIE *ET AL.* (2022)

Please refer to Mackenzie *et al.* (2022) for the following relevant tables:

Table 33. Estimates and standard errors for adult and first-year survival used to define the 'prior' distributions.

Table 34. Estimate and standard errors for age at first reproduction (AFR) and inter-birth interval (IBI) used to define the 'prior' distributions.

Table 35. Literature sources used to derive 'prior' distributions of demographic parameters in Tables 33 and 34.

Table 36. Posterior and 95% credible interval for $r_{max}$ obtained in this analysis and estimate and 95% credible interval elicited in the second round by Abraham *et al.* (2017).

The National Oceanic and Atmospheric Administration (NOAA) in the USA currently use a default $r_{max}$ of 0.04 for cetaceans, apart from those species shown in Table 37. Also shown are the values used for five species of pinnipeds. The posterior median and 95% credible intervals for $r_{max}$ that were obtained in this analysis for orca, humpback whale and (southern) elephant seal are given as a comparison.

Table 37. $r_{max}$ values used by NOAA for specific species rather than a default value. The posterior median and limits of the 95% credible interval obtained in this analysis for comparable species are also given (e.g. southern elephant seal for norther elephant seal).

## 3. SPATIAL DISTRIBUTION OF HECTOR'S AND MĀUI DOLPHINS

This document provides additional information and suggestions in response to document SC/69A/MD/WP/04, entitled 'Review of spatial risk assessment for Hector's and Māui dolphins in New Zealand: spatial distribution', authored by Megan C. Ferguson, in support of an independent review by the IWC of a spatial risk assessment of threats to Hector's and Māui dolphins (Roberts *et al.*, 2019).

Briefly, the approach used by Roberts *et al.* (2019) used Generalised Additive Models (GAMs) fitted to different sources of spatial information to predict: the seasonal (summer and winter) coastal density of Hector's and Māui dolphins (fitted to count-based aerial survey data from the South Island of New Zealand, where Hector's dolphin occur); and the density of dolphins in Harbours and coastal region of the West Coast North Island (fitted to a dataset of validated boat- based public sightings combined with the locations of recreational fishing boats). Both these models used habitat-based covariates of sightings rate. The resulting model predictions were then merged, yielding seasonal spatial densities for coastal and West Coast North Island harbour areas. Finally, these layers were rescaled according to regional population rescaling strata, which used the latest estimated population sizes in each region. The rescaled seasonal layers were then used by the spatially-explicit fisheries risk assessment (SEFRA) model. All parts of the spatial risk assessment and its inputs are described in Roberts *et al.* (2019).

Megan Ferguson provided a detailed, insightful, constructive and well-constructed review of the spatial distribution modelling undertaken, for which we are extremely thankful. Unfortunately, it was not possible to reply to all of the questions and points raised by the review in the time before the workshop. Hence, we have focussed on the main points, noting that other points not responded to can also be discussed at the workshop.

For convenience, we have used the same headers (a)-(h) that the review was partitioned into.

### 3.1 Choice of habitat variables

(a) With respect to the research trawl survey used to generate prey presence layers, there were no winter surveys in some regions. Hence the decision to pool the data from the two seasons, resulting in aseasonal prey layers. We agree that the spatial distributions of some prey (especially red cod and southern arrow squid), will vary seasonally, as Miller showed at a local-scale. To some extent, we expect the seasonal turbidity relationship will have aliased for seasonal prey movements.

(b) We fitted to presence-absence of prey in trawls (rather than catch rate) because we were integrating observations from two different research vessels with probable differences in catchability for each prey species. Furthermore, it was not possible to include vessel as a term in this model due to an extremely limited spatial overlap between the two survey vessels. Thus presence-absence was all that was really possible.

(c) We agree that alternative smooth types could have been trialled using GAMs, including: tensor, Gaussian Process, soap smooths, and others.

(d) We agree that the bulk of the dolphin sample available for dietary analysis by Miller will have been recovered in summer months, and this is a potential bias in our dietary information. We acknowledge a potential conflict with the isotopic information from Miller (a good spot by the reviewer), although would be unsure of how to rectify this without knowing the relevant epipelagic species (which we may well lack spatially comprehensive info for).

(e) Regarding specific questions on page 3, we agree that this part of the modelling could be improved, e.g., using bivariate smoothers, flooding of habitat layers, so that, for example, lots of the SCSI data does not have to be omitted from the analysis. A significant revision of the spatial distribution modelling for the South Island has recently been completed (Roberts and Webber, 2023), which used additional sources of spatial information, used bivariate smoothers, and considered a number of other candidate temporally dynamic habitat covariates (e.g., SST, wind speed, sea level anomaly), although none of these were accepted into the final model structure. The review mentioned some other candidate habitat covariates that could be looked at in the future.

(f) We note the recommendation from the reviewer that the authors should conduct a sensitivity analysis to investigate the effects of making predictions from the model based on climatological vs. contemporaneous covariates. For context, we suggest that, due the small home range size of individual Hector's and Māui dolphins and preference for a turbid habitat, which does not change very much in its spatial distribution through time (compared with seasonal variation), we believe that this species is not so highly mobile as many other coastal cetaceans. Previous boat- based and aerial surveys (compare summaries by Clement (2005) with aerial survey data shown in Roberts *et al.* (2019) indicate that there are quite minimal changes in their spatial distribution by year, compared with seasonal changes, which are likely to be more pronounced.

### 3.2 Choice of dolphin occurrence data for model construction

(a) We agree that the aerial data from the Mackenzie surveys are the best single source of spatial information for fitting Hector's dolphin spatial distribution models. We augmented this with additional sources of sightings-based spatial information in Roberts and Webber (2023), which were not so rigorous in their sampling design and often had patchy information about locations with effort but no sightings, although were helpful for refining model predictions at specific, key locations.

    (b)    Acoustic data have also been collected that could potentially have been used in spatial modelling, although were not. This might have been particularly useful for modelling the relative abundance of Māui dolphin in WCNI harbours, based on the relative detection rate inside harbours versus coastal areas.

### 3.3 Model selection criteria and fitting

    (a)    We can confirm that multiple previous surveys (e.g., (Mackenzie and Clement, 2016; Rayment, 2008) have found that very few Hector's and Māui dolphins are observed over waters deeper than 100m depth. Hence, the domain of model fitting (from 0-250m depth) should have been sufficient.

    (b)    We acknowledge that the 1km spatial resolution of the data fitted to by the spatial distribution models is fine. Future modelling could assess the effects of trialling coarser resolutions.

    (c)    The response variable for the spatial distribution modelling was the observed count for each survey line segment adjusted for helicopter-based estimates of availability for the respective area and season (Mackenzie and Clement, 2016). Uncertainty about availability was not included in the spatial distribution model, although this is a good suggestion by the review.

    (d)    We trialled alternative error structures, including Tweedie and others, when updating the South Island Hector's dolphin modelling (Roberts and Webber, 2023). However, negative binomial consistently came out as the best in terms of AIC. It is possible that the standard qq.gam() function is not appropriate when specifying a negative binomial distribution, and that this may be the source of the apparent misfit. This is something we can look into.

    (e)    Confirming that the spatial distribution models used a log link function.

    (f)    We agree that seasonality in some of the other candidate predictors could have been explored. In the updated Hector's dolphin modelling (Roberts and Webber, 2023), seasonality in distribution was represented using a seasonal (bivariate) surface smooth, in addition to non-seasonal habitat covariates. There are many ways of modelling seasonal movements of the dolphins and we agree this is worth exploring.

    (g)    We agree that it would have been better if regional variation in dolphin density (not explainable by the habitat covariates) had been accounted for in some way. In the updated Hector's dolphin modelling (Roberts and Webber, 2023) this was achieved using the seasonal surface smooth, which allowed much lower densities to occur in the south and north of the South Island relative to the east and west coasts.

    (h)    We agree that it would be better if uncertainty in the spatial distributions was propagated through to the SEFRA risk model. Ideally, the spatial distribution model should be a subcomponent of the risk model, although this would have considerably slowed the optimisation of the overall model, so this was not done at the time. Subsequent SEFRA model development (for selected shark species) is looking into this at the moment.

### 3.4 Combination of models for merging coastal and harbour predictions

    (a)    The relative lack of winter sightings of dolphins along the West Coast of the North Island is most likely primarily driven by low human activity during this period. This was apparent from the outputs of the recreational fishing boat survey, which was run in the summer and winter periods (not shown). Dolphins do move further offshore in winter also, though this is unlikely to be so influential on seasonal sighting rate.

    (b)    We agree that using recreational boat locations as pseudo-absences to accompany the validated boat-based public sightings has some issues. Taken together, these data do indicate that the dolphins rarely, if ever, use the harbours. There is also fairly good agreement with other sources of spatial information we have from this coastline and the harbours, including: sighting rate information from boat-based mark recapture studies (Constantine *et al.*, 2021), and the relative rate of acoustic detections from C-PODs etc (summarised on p.63 of Roberts *et al.*, 2019), none of which were used by the spatial distribution modelling, except to provide informal validation.

    (c)    Ongoing research is using drones to monitor the coastal distribution of the dolphins. But the relative occurrence of dolphins in harbours is also important given fishing effort in the harbours. Better usage could be made of the acoustic information and boat-based research survey data.

    (d)    Confirming that we used a point-based model for the harbours, rather than working with gridded data. We appreciate suggestions made in the review for how this analysis could have been done differently.

### 3.5 Model validation and interpretation of results, including a review of model estimates at the scale of small or hypothesized local dolphin populations

    (a)    We agree that model cross validation would have been highly desirable, though this was not pursued at the time, due to the fast pace of the overall assessment project. This has subsequently been done by a project updating the spatial distribution modelling for South Island Hector's dolphins (Roberts and Webber, 2023). For this, two alternative cross-validation approaches were trialled: (1) a random k-fold approach, which favoured the most complex candidate model; and: (2) a source-based cross validation approach, iteratively withholding different sources of information as the test data, which favoured models with only a couple of habitat terms. We deemed the second of these approaches to be best for guiding the selection of models with good predictive power. Notably, this still resulted in surface turbidity and prey presence (although a different prey species) being the best habitat-based covariates of dolphin sighting rate across the various data sources.

**3.6 Evaluation of spatially-resolved predicted bycatch compared to known records, including beach-cast carcasses and fisher-reported catches from vessels without observers and Evaluation of spatially explicit estimates and associated uncertainty of dolphin deaths from commercial fisheries compared to estimates from simpler models**

The reviewer comments/recommendations from both F and G are both replied to here.

    (a) Confirming that capture events could involve multiple dolphins captures in the same event, hence the use of a Poisson distribution to model observed captures.

    (b) The number of individuals captured by the single observed trawl capture was artificially doubled in order to account for the vessel-reported trawl capture data, for which a mean of around 2 dolphins were captured per event.

**3.7 Review of estimated patterns of changing fisheries risk over time**

The review strongly recommended assessing the sensitivity of SEFRA estimates of dolphin captures, deaths, PST, and risk ratios for scenarios in which the true values of the key model parameters (e.g., vulnerability to capture and cryptic mortality) and spatial distribution vary over time. We suggest that the catch rate information is probably too thin to allow for a meaningful time blocking of the model parameters, although this could potentially be explored. Re spatial distribution, there is limited evidence of any large scale changes through time, although model predictions of spatial distribution could be made using habitat variables averaged across different time periods.

## 4. SEFRA MODELLING

This document provides additional information and suggestions in response to document SC/69A/MD/WP/02, reviewing the spatially-explicit fisheries risk assessment (SEFRA) approach, authored by Brian Brost (NOAA). This review was contributing to the wider independent review by the IWC of a spatial risk assessment of threats to Hector's and Māui dolphins (Roberts *et al.*, 2019).

We would like to thank the reviewer for working through the assessment and acknowledge the difficulty associated with understanding models via a report. This document provides additional information on the main points raised by Brian Brost. ***Text from the review is highlighted in blue italics***. Our replies are below in black.

### 4.1 Identifiability of model parameters

*'The Poisson likelihood presented for the bycatch portion of the SEFRA approach is a function of 4 parameters: dolphin abundance ($N_s$), vulnerability ($v_g$), the probability bycatch is observable ($p_g$), and the probability that a captured dolphin is alive ($\psi_g$). The dolphin bycatch data contain no information about $N_s$ and ($p_g$), so it's no surprise their posterior and priors match (p.48). Any claim to 'estimate' these parameters (e.g., pp.22-23) is misleading because they aren't identifiable.'*

It is true that some of the model parameters are not informed by the fishery observer capture data, including $\psi_g$ (probability bycatch is observable), and we made no secret of this. For these parameters, we expected that the posteriors would look exactly like the priors, which is what occurred. Bayesian inference for models that include non-identifiable parameters is still possible so long as priors are specified for all non-identifiable parameters. Such inference will still result in a valid posterior, but the posterior is strongly affected by the prior. We do not think this is an issue so long as these assumptions are declared.

A major source of residual uncertainty with respect the estimation of commercial set net deaths is the proportion of deaths that are observable ($p\_g$), as highlighted in the executive summary of Roberts *et al.* (2019). The posterior for this parameter was effectively decided by the selection of the prior, which was derived by a literature review in Appendix 10 of Roberts *et al.* (2019). This review was largely based on proxy species, since we lacked the types of observations we needed from Hector's dolphins.

Technically, it is not true that the data contain no information about $N_s$, because we know that the number of dolphins caught as bycatch and that subsequently die cannot exceed the population size (i.e., $N_s$). Therefore, including $N_s$ in the model indirectly provides an upper bound for the number of estimated deaths. These types of constraints become increasingly important in data poor models and this was the impetus for combining the estimation of captures and deaths into an integrated model. It is true, however, that in the case of the Hector's and Māui dolphin model, the priors for $N_s$ were not updated, although we do not consider this an issue.

### 4.2 Confounding of model parameters?

*'Even if the data were informative, there are some structural issues with the SEFRA model specification. Consider y = α x β as an example. No matter how much data (y) are available, it's impossible to distinguish between α and β - any given change in α can be offset by a change in β. It's likely that $v_g$ and $p_g$, in particular, are similarly confounded in this model.'*

We know that it makes no difference to the model outputs if $p_g$ is included in the calculation of captures or not. Let's take a simple example that deals with a single species, a single fishery group, and ignores live captures (i.e., assumes all captures result in death). The number of observed captures and total number of deaths can be defined as $C' \sim \text{Poisson}(O' \times N \times v \times p^{obs})$ and $D = O \times N \times v$, respectively. In these two equations O' is the observed overlap (i.e., observed fishing intensity multiplied

by relative density), O is the total overlap (i.e., including observed and unobserved fishing effort), *N* is the population size, *v* is the vulnerability, and $p^{obs}$ is the probability that an event is observable. The same result can be obtained by dropping $p^{obs}$ from the definition of captures and instead writing $C'$~Poisson(O' x *N* x *q*) and *D* = O x *N* x *q* x 1/ $p^{obs}$, where the vulnerability parameter (*v*) is replaced with a catchability parameter (*q*). If the prior and posterior distributions for $p^{obs}$ are the same (as was the case for Roberts *et al.*, 2019), then these two model forms are equivalent.

However, we agree that it is non-intuitive to have observed captures be contingent on the probability of captures being observable ($p^{obs}$), as was the case for the SEFRA assessment of Roberts *et al.* (2019). Subsequent SEFRA models (e.g., MacKenzie *et al.*, 2022) have moved $p^{obs}$ outside of the estimation of observed captures and catch rate is parameterised as 'catchability' instead of 'vulnerability'. This is a more intuitive (and followable) parameterisation, although has no effect on model outputs, including estimated deaths and risk.

### 4.3 Estimating vulnerability and probability of live capture

*'Estimating $\psi_g$ and $v_g$ are separate objectives and use the bycatch data in different ways. It's better practice (and more transparent) to tackle each objective in turn using appropriate error structures. For example, estimating involves a comparison of dead vs. live captures. Such a model is very easy to implement. In fact, given the prior specification in Table 5, obtaining 10,000 samples from the posterior distribution for in the set net fishery is a single, short line of R code: rbeta(10000, 3 + 1, 12 + 3). This model does not use a Poisson likelihood or involve $\rho_g$.'*

We believe that this approach is merely a different way to achieve the same thing. However, we do not believe the suggestion to use something as simple as '*rbeta(10000, 3 + 1, 12 + 3)*' is correct. Here, the review appears to have mixed the specification of an uninformative model prior (i.e., the Beta(1, 3) prior that was substantially updated by the data) with the data itself (i.e., the 3 caught alive and the 12 caught dead). A better way to implement the reviewer's suggestion would be to rearrange the equations and use a binomial distribution. For instance, the number of observed captures (alive and dead) could be defined as $C'$~Poisson(O' x *N* x *v* x $p^{obs}$) and the number of live captures only could be estimated using *C*live'~Binomial(*C,,ψg*). Although untested, it is likely that either approach would achieve the same result.

*'Estimation of vulnerability ($v_g$), on the other hand, involves modeling total captures as a function of fishing effort and dolphin population dynamics. A sensible model would use a binomial distribution with parameters $N_s$ and $v_g$, with $v_g$ subsequently modeled as a function of spatially-resolved fishing effort and dolphin density (but not $p_g$). Within the MCMC algorithm, uncertainty in dolphin abundance is propagated by conditioning on realizations from [$N_s$]. That said, I assume dolphin abundance and density vary over the 21-year period encompassed by the dolphin capture data, yet abundance and density information aren't available annually. I don't recall discussion concerning the assumption that these quantities are static through time. Note that this objective also relies on observer deployments that follow some known sampling design (e.g., random deployments across the entire fishery), which isn't discussed in the report, either. See below for additional comments regarding the authors' model specification for $v_g$.'*

We do not think that a binomial model makes sense for the main model likelihood, since more than one dolphin could be caught per fishing event.

The reviewer is correct in assuming that dolphin abundance and density may vary over time, but in Roberts *et al.* (2019) a simplifying assumption was made that this was static, mostly because of lack of concrete information about potential changes in population size over time. Note that in subsequent SEFRA models (e.g., an assessment of New Zealand sea lions at the Auckland Islands by Large *et al.* (2019), time-varying estimates of species abundance were used, which was deemed necessary in that case, due to protracted periods of changing population size. This is something that can easily be brought into SEFRA models, if needed and the data exist.

### 4.4 Stochastic simulation v Bayesian modelling

*'Aside from the inference concerning non-fishery deaths (i.e., [$\rho$]/.] $\alpha$ [$\eta/\rho$]), everything else in the SEFRA approach is an exercise in stochastic simulation. It's not estimation or statistical modeling, Bayesian or otherwise. It's incorrect to label quantities calculated during this simulation exercise as posterior distributions (e.g., R and N on p.7).'*

It is true that the model did not update the priors for population size and $p^{obs}$ and the probability of live capture in trawls (there was only one observed capture for this fishery group). For all the other model estimated parameters, the prior was updated, including the vulnerability priors as well as the non-fishery death parameters. We consider it straightforward to determine which parameters are and are not estimable, when models are structured in this way. Furthermore, we do not believe it matters if not all priors are updated, so long as there is sufficient justification/discussion of the generation of those priors (e.g., Appendix 10 of Roberts *et al.*, 2019).

*'This simulation exercise occurs independent of the MCMC algorithm(s). It includes samples from the posteriors of $\psi_g$, $v_g$ and $\rho$, as well as random deviates from the presumed distributions for the quantities $N_s$, $\omega_g$, $r_{max}$, $S^{1+}$, $p_{g, and} k_g$. Because the quantities $N_s$, $\omega_g$, $r_{max}$, $S^{1+}$, $p_{g, and} k_g$ are not directly modeled as a function of dolphin data, the output of such a simulation is only as good as the many presumed distributions. I appreciate the authors' effort to define distributions that reflect the current state of knowledge for some quantities (e.g., $N_s$); however, there are others that are only assigned vague distributions due to limitations in available information (e.g., $p_g$ and $\omega_g$; note that neither parameter is actually estimated in the SEFRA approach). The assumed distribution for $\rho_g$ seems particularly tenuous and, as the authors note on page 71, the results are sensitive to this quantity.'*

Contrary to the statement above, the priors for the $\psi_g$, $v_g$ and $\rho$ parameters were clearly updated by the data (see a plot of priors and posteriors reproduced at the end of this text). An informative prior for the $N_s$ parameter was defined and, if this prior was updated given the available data, it would indicate a potential issue with the model. Vague distributions were assigned to other random variables (e.g., $p_g$ and $\omega_g$) to represent our current understanding (or rather lack of) for these variables and the uncertainty associated with these variables is propagated throughout all model results.

We were very clear in stating that the uncertainty in $p_g$ is a major source of uncertainty with respect to the estimation of fishery risk and that the results are sensitive to this quantity. We agree with this statement.

For clarification, $\rho_g$ and $v_g$ are essentially the same thing as $p_g=1/k_g$. The parameter $k_g$ was only derived from the model estimate of $p_g$ and was not used in the subsequent estimation of deaths or risk. Really it was just for reporting a cryptic mortality multiplier, for readers that might prefer to think in those terms.

*'Shouldn't the number of deaths resulting from commercial fisheries depend on $\rho_g$ and $k_g$? Neither of these quantities enters the equation for $D_{sg}$ at the top of p.28.'*

Given our implementation it should not. Please see the explanation above for the different parameterisations that are possible. Including then $\rho_g$ parameter in the estimation of captures, rather than simulating from the $\rho_g$ prior and including it in the derivation of deaths, only results in the same outcome. However, recent iterations of SEFRA models (e.g., the latest Marine Mammal Risk Assessment and Seabird Risk Assessment) use the latter model construction as we have identified this part of the model as somewhat confusing. Nevertheless, the model results are the same either way.

*'Within this simulation exercise, non-fishery mortalities are a function of commercial fisheries bycatch (p.28). Therefore, the quality of output concerning deaths due to toxoplasmosis, predation, etc. is closely linked to (and only as good as) the methods used to generate commercial fishery deaths (in addition to presumed detection probabilities for non-fishery deaths).'*

We don't see this is as a particular problem – this is a feature of many models. In this case, estimated annual deaths will only have comprised a small fraction of estimated annual deaths, such that the estimation of fisheries risk would have an extremely minor effect on the other deaths that were subsequently divided up into lethal non-fisheries threats. In any case, there are clearly other more significant uncertainties with respect to the estimation of non-fisheries deaths based on necropsy data of beach cast dolphins.

*'The authors describe vulnerability ($v_g$) as the 'probability of capture or death' (p.22). Aren't the probability of capture and the probability of death two different things (hence $\psi_g$)? Furthermore, probabilities have support over [0, 1], but the support for $v_g$ is [0, ∞). Consequently, the parameter either isn't described correctly as a probability or the prior distribution is inappropriate. Considering units of measurement alone (the units for $\lambda$ is number of dolphins), I don't think $v_g$ can actually be a probability. It's also likely that $v_g$ is structurally non-identifiable given the existing model specification (see above).'*

The reviewer is correct that the vulnerability parameter does not represent the probability of capture or death, it represents the vulnerability to capture only. It would only represent the probability of deaths if $p_g=1$, which it was not. Hence this statement may be a hangover from prior SEFRA modelling not including cryptic mortality for some fishery groups. However, the statement 'It's also likely that $v_g$ is structurally non-identifiable given the existing model specification' is incorrect as this parameter is well-defined (see the prior versus posterior plot) and it does not matter if the $p_g$ parameter was introduced during or after inference as either way the same result would be obtained (see the explanation provided above).

### 4.5 Additional comments

*'I found no information in the report on observer coverage (i.e., the proportion of the fishery that's observed for bycatch) or how observers were deployed. Without this information, it's impossible to know the quality of the available data. In particular, if observers weren't deployed under some design (e.g., randomly), the potential for quantifying the relationship between dolphin bycatch and covariates (e.g., fishing effort and dolphin density) may be limited. Comparing observed and estimated mortalities in the commercial fisheries, observer coverage must be extremely low (13 observed mortalities over 21 years vs. an estimated ~40 deaths per year; pp.42 and 50).'*

Observer coverage was presented spatially in figure A15-1, although, admittedly, with no plots by year or season, which may have been helpful for the reader. We acknowledge that it would have been useful for readers and reviewers to have produced a temporal characterisation of observer coverage. The review is correct that the rate of observer coverage is low (around 1% across all commercial set nets and about the same coverage rate for inshore trawls), with better coverage in regions of highest perceived risk for Hector's and Māui dolphin. There is also considerable inter-annual variability in observer coverage rate. This was one of the issues focused on by Hélène Peltier's review, which is probably the better place to discuss the coverage of fisheries observer data.

*'Several observed captures appear to occur in areas of low risk (p.48 and Figure A15-3), though poor figure quality makes it difficult to assess the agreement between the distribution of observed and predicted captures. Were any formal goodness-of-fit diagnostics performed? Visual inspection of maps, etc. is a good starting point for checking model adequacy, but it's not a substitute for formal diagnostics. See Conn et al. (2018) for methods to assess model fit.'*

*A priori* one would assume that some observed captures would happen in areas of low risk. We provided maps illustrating the agreement between the spatial distribution of observed captures and predicted captures (Figure A15-3). We also showed the observed versus predicted number of captures (Figure A13-3). At its heart, SEFRA is not truly a spatial model since it does not fit spatially to the catch rate information. So, we do not think that formal spatially explicit statistical goodness-of-fit diagnostics, as suggested by the reviewer, are appropriate.

*'There's reference to vessel-reported dolphin captures on p.48. Is it possible to use these additional data in models of dolphin bycatch?'*

Unfortunately, this is not possible, because there is no way to establish the reporting rate of commercial vessels. The only viable possibility would be to overplot vessel reported captures on to model predictions (they are plotted in Figure A7-2), with the caveat that reporting rate may well vary considerably in space and time.

*'Assessments of risk through time (overlap, deaths, and risk ratio; p.56) require dolphin abundance and density surfaces that vary annually. As far as I understand, those data are not available at an annual timescale?'*

This would be desirable, though it is not a fundamental requirement. Although dolphin abundance and the density surfaces were assumed to be static through time, fishing effort changes through time, enabling some inference about the time-varying nature of model outputs to be explored. While we lack good information about potential changes in abundance through time, other aerial and boat-based surveys (unused by our assessment) indicate that the spatial distribution of dolphins has changed little through time, except seasonality (summer and winter), which our risk model represented.

*'Why is it necessary to assign an arbitrary number of animals to areas without a known, permanent dolphin population (p.59)? I also don't understand how number of deaths is insensitive to the arbitrary choice in population size.'*

If no dolphins are assumed within an area, then captures, deaths, and risk collapses to zero for that area. Including some arbitrarily small number of dolphins in fringe area allows the occasional dolphin to extend its range beyond the norm and better admits our uncertainty related to their distribution. It also allows the estimation of risk, if dolphins did occur in these regions, which is useful for managers (e.g., in the scenario where dolphins temporarily transit through areas where breeding populations are not known to occur).

*'I agree that the SEFRA approach relies on good knowledge about key quantities (p.71); however, I don't follow how $v_g$ can 'soak up any mistakes' made in the prior specification for $p_g$. These parameters are 'estimated' jointly and the data contain no information concerning $p_g$. This seems problematic because predicting deaths relies 'heavily' on the prior specification for $p_g$'*

The review may have taken this statement out of context. The relevant part of the paragraph reads: '*The first part of the SEFRA method, in which numbers of observed captures are estimated, is the most certain component of the model because the estimated vulnerability parameters can soak up any mistakes made in the specification of the probability that an event is observable prior.*' The 'soaking up of mistakes' is only relevant to the prediction of the number of observed captures part of the model and does not apply when discussing predicted numbers of deaths (remember that including the probability that an event is observable parameter can be done during the estimation of captures or afterwards).

**Plot of Prior and Posterior Distributions from Roberts *et al.* (2019)**
See figure 10 on opposite page.

## 5. FISHERIES INFORMATION FOR SEFRA MODELLING

This document provides additional information and suggestions in response to document SC/69A/MD/WP/03, reviewing the bycatch estimates, available data, and their likely implications, authored by Hélène Peltier (La Rochelle University). This review was contributing to the wider independent review by the IWC of a spatial risk assessment of threats to Hector's and Māui dolphins (Roberts *et al.*, 2019).

We would like to thank the reviewer for working through the assessment and acknowledge the limited information in Roberts *et al.* (2019) on New Zealand's national observer programme. This document aims to provide additional information relevant to the main points raised by Hélène Peltier. *Text from the review is highlighted in blue italics*. **Our replies are below in black.**

### 5.1 Temporal and spatial distribution of coverage

We acknowledge the low and highly variable observer coverage that has been achieved in New Zealand inshore set net and trawl fisheries. Figures 1 and 2 in Peltier show this coverage for inshore trawls and set nets across the EEZ (i.e., including effort where Hector's and Māui dolphins are not found). These numbers don't reflect the higher levels of coverage that are achieved in the targeted areas/seasons of highest dolphin density or where the residual risk is higher. Further detail on New Zealand's observer programme and examples of targeted observer coverage areas are provided in Appendix 1 below.

*'This very irregular temporal distribution of observer coverage is an issue as soon as it is not taken into account by the modelling process. This work considers bycatch time series from 1995/96 to 2016/17, which could suggest a relative stability in observer cover. But a deeper exploration of available data seems to demonstrate important gaps and heterogeneity in observer programmes. As bycatch data are used in the modelling process as non-spatialized input data summed over the whole period, the absence of observer onboard and the absence of bycatch are mixed up.'*

Plot of prior and posterior distributions from Roberts *et al.* (2019)
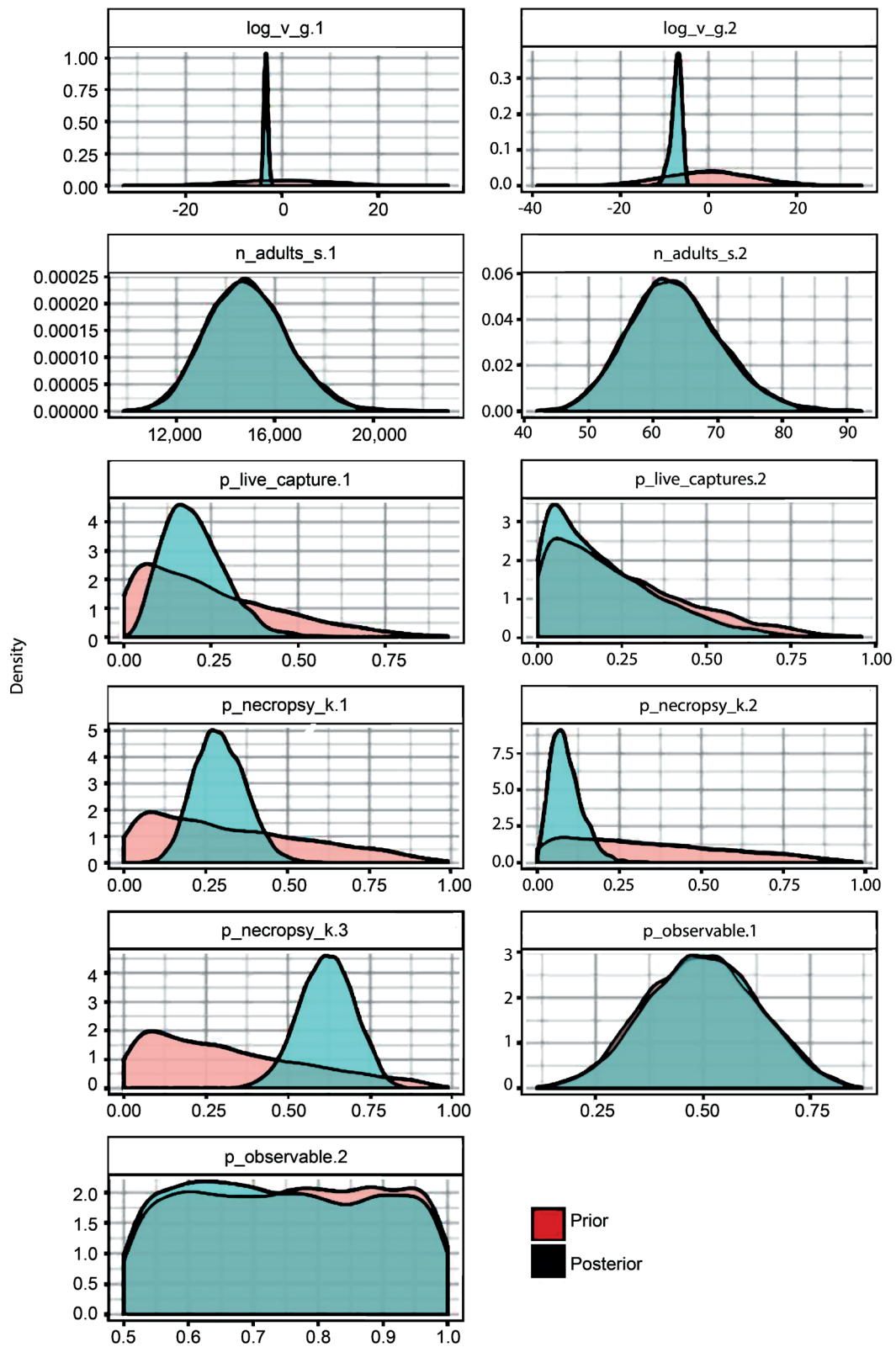
The SEFRA method was designed specifically to address the needs of fisheries managers in low information fisheries (that is, most inshore fisheries) where fisheries observer coverage is too low, and protected species capture rates are too rare, to inform statistically robust capture estimates using more traditional 'spatially blind' capture models.

The SEFRA method recognises that fisher distributions and animal distributions are non- uniform and variable in space and time. By estimating the likelihood of an encounter between protected species and fishing events as a function of their overlap in time and space (rather than assuming that every observed event constitutes an independent sample from a homogenous pool, as in spatially blind methods), the SEFRA method greatly reduces the level of observer coverage required to obtain a statistically robust captures estimate, and neutralises the effects of spatial bias arising from spatially non-random or unrepresentative fisheries observer coverage.

The SEFRA method begins with a spatially resolved estimate of the spatial distribution (i.e., a digitised map) of the species in question. The rate or probability of encounter between animals and fishing is estimated by multiplying this map by a corresponding spatially resolved estimate of the fishing effort distribution (at the scale of each fishing event, not summarised into cells). In this way the likelihood of encounter can be expressed at the scale of each event or summed or disaggregated as required. The map or quantity obtained by multiplying these two distributions is called spatial overlap and is a proxy for the rate or likelihood of encounter with fishing effort.

The likelihood of capture per encounter, i.e., 'catchability' can then be estimated empirically using fisheries observer data from that portion of the overlap for which the fishing event was observed. Because overlap itself is variable in space and time, spatially non-representative observer coverage does not bias the estimation of catchability. The estimated catchability can then be multiplied by total overlap (including fishing events without a fisheries observer) to estimate total captures at the scale of the fishery.

The SEFRA method acknowledges that some animals may die as a consequence of their interaction with fishing gear but nonetheless remain unobserved (cryptic mortality), even on fishing events where a fisheries observer is present.

By using the SEFRA method the observer coverage required to obtain a statistically robust captures estimate is focussed on areas where the risk is most likely to occur. This neutralises the effects of spatial bias arising from spatially non-random or unrepresentative fisheries observer coverage.

The SEFRA estimation of annual captures, deaths, and risk for Māui dolphins from commercial fisheries includes the spatial overlap of fishing events with both the estimated summer and winter spatial abundance of Māui dolphins to estimate the encounter rates with fisheries. Fisheries observer data is combined with these overlap metrics to estimate the probability of death per encounter within the SEFRA model.

For Māui and Hector's dolphins, the SEFRA model was fitted to all fisheries observer-reported capture rate information for commercial set net and inshore trawl fisheries up to 2017. This equates to 21 years of observer data. For fisheries where there is a high probability of not observing any bycatch, there is a potential for models to suffer from 'small sample bias' where risk is underestimated. For set nets, the number of observed deaths (15 across 21 fishing years) indicates that there was unlikely to be any negative small-sample bias.

Small-sample bias is more applicable to the inshore trawl fishery (total of 1 observed capture across 21 fishing years). However, the estimate of entanglement (capture) in trawl fisheries is not based on the single observed capture. The spatial risk assessment model was fit to all observed fishing events over the 21 years. Additionally, within the spatial risk assessment model, a model sensitivity was used rather than the base case model run in Roberts *et al.* (2019), which doubled the observed mortality rate for the inshore trawl fishery to account for the risk of multiple capture events. This is a precautionary approach that doubles the estimate of trawl fisheries risk relative to what was observed.

Unobserved or cryptic mortality has also been accounted for in the SEFRA model. For the Māui dolphins, in lieu of any information about whether observers count dolphins dropping out of gear before brought on deck, a cryptic mortality multiplier was applied. For set net fishery deaths, the cryptic multiplier increased modelled deaths by 33-300% above the observed value.

## 5.2 Spatial distribution

*'It seems that setnets equal or below 8m in length were never observed by observers (Fisheries New Zealand, 2022). Fishing areas of <8m setnets seemed to have a great overlap with both Hector's and Māui dolphins in Northern Island and South Eastern Island (fig. 3 and 4). The absence of bycatch in setnets in Northern Island could therefore be due to the absence of at-sea monitoring in this area.'*

We note that Figures 3 and 4 in Peltier do not provide sufficient resolution to characterise the fishing areas of set net vessels equal to or below 8m in length relative to the estimated distribution of Hector's and Māui dolphins. For example, the effort shown on the East Coast South Island occurs within Lake Ellesmere that has no natural outlet to the sea and does not overlap with Hector's dolphins. The SEFRA model accounts for the potential presence of Māui dolphins in West Coast North Island harbours despite extremely low observed occurrence.

While there has been very little observer coverage on commercial set net vessels ≤8m length, the lack of observer coverage does not prohibit estimation of estimate total deaths and risk ratio commercial set net vessels may cause. The

model applies the estimated catchability of Hector's and Māui dolphins by set net across the areas where fishing effort and dolphin distribution are estimated to overlap. Set net effort (for vessels ≤8m length, including observed and unobserved events) was included in the data set used by the SEFRA model to estimate total deaths and risk ratio for commercial set nets. Note that this assumed that the capture rate (per km of set net) was the same for the small and large vessels, given the same spatial overlap with Hector's and Māui dolphins.

*'As fisheries observer data are used as input data, the spatial heterogeneity of observer covers and total absence of monitoring on 70% of vessels can artificially highlight and/or turn off some bycatch risk areas. The observer effort, in addition of being very low, is not randomly distributed across the fishery, both in terms of vessels and also fishing areas. If these conclusions are common with most of countries, they must be considered in analytical choices in order to integrate these biases in final estimates (Authier et al., 2021; Rouby et al., 2022).'*

One of the main drivers for developing the SEFRA approach was to account for potential differences in the spatial pattern of observed and unobserved fishing events. However, the small total number of observed captures is a constraint for considering alternative vulnerability groups (e.g., based on target species or depth of effort), and can affect our ability to account for potential non-spatial biases relating to observer coverage.

### 5.3 Observer biases

*'Two main biases were identified in observer programmes: (i) the deployment effect, or non-random assignment of observers to vessels and ports due to the fact that accepting an observer on board is at the vessel master's discretion; and (ii) the observer effect, i.e. a change in fishing practices when an observer is present (Amandè et al., 2012; Benoît and Allard, 2009; Faunce and Barbeaux, 2011; Stratoudakis et al., 1998). These biases are likely to hinder the detection of bycatch by observers.'*

We would like to clarify that placement of observers on fishing vessels in New Zealand is not at the vessel master's discretion. Fisheries New Zealand determines observer placement considering health and safety, monitoring objectives and operational matters (as noted above).

Notwithstanding, we acknowledge the potential observer effect in detecting bycatch. One of the ways we try to help address this bias is to target higher proportional coverage for particular areas/seasons. Fisheries New Zealand has also undertaken some work to assess changes in vessel behaviour when observers are onboard (Tremblay-Boyer and Abraham, 2020). This study demonstrated that on-board cameras can effectively remove any potential bias. With the upcoming expansion of on-board cameras in New Zealand's inshore fisheries this is likely to further reduce observer effects.

### 5.4 Fisheries

*'By construction, the analysis presented in the report considers a static and constant fishing effort since 1995/96, including fishing practices and gear dimensions. As no information is available in the report about gear types (pelagic/bottom trawls? Gillnets/trammel nets?) and their temporal evolution, it remains delicate to assume consequences of such analytical choices, like using 2014/15 to 2016/17 as fishing effort paragon.'*

To clarify these points, the SEFRA model assumes that the dolphins' spatial distribution is temporally static (except by season) but that fishing effort in the assessed fisheries is not assumed to be constant. The SEFRA model posteriors, are estimated when fitting to the observed portion of effort. The number of fishery deaths is then calculated for all effort by the respective fishery groups. The spatial distribution and total magnitude of effort varies considerably through time, particularly with the evolution of fishing area restrictions (see Appendix 16 of the report).

No other fishery groups (gear types other than set net or trawl) were considered because there were no observed captures reported for any other groups. Including these other groups would result in near zero deaths estimated for these groups.

*'On Fisheries New Zealand website[1], some available information on fishing effort showed around 50% reduction of net length between 2002/03 and 2019/20. Is this reduction related to changes in net dimensions or number of vessels? If related to changes in net dimensions, is this length reduction followed by an increase of net height, as observed in other countries? If so the catchability of dolphins could raise and consequences on dolphin bycatch could be worrying. If this decrease is related to a diminution of fishing vessels, it supposes that greater net length operated in New Zealand waters when the fisheries wasn't involved in observer programmes. This may have resulted in an important ignorance and underestimation of former bycatch levels. The same pattern was detected in trawl fisheries, that would also require more information to describe the underlying process and its potential consequences on cetacean bycatch.'*

A large amount of the reduction in total net length (see Appendix 16 of Roberts *et al.,* 2019) will have been a response to the implementation of additional fishing area restrictions through time, including new areas closed to set net fishing in 2003, 2008, 2010, 2011, 2012, and 2013. In some regions, we will have had reductions in the number of vessels operating as well as total net length because of spatial management. We note your suggestion to look for potential changes in net height coincident with the reduction in total net length. This is something we will look into.

---

[1] *https://protectedspeciescaptures.nz/PSCv6/.*

**Plot of prior and posterior distributions from Roberts *et al.* (2019)**
See Fig. 10.

**References**

Abraham, E.R., Neubauer, P., Berkenbusch, K. and Richard, Y.  2017. Assessment of the risk to New Zealand marine mammals from commercial fisheries. New Zealand Aquatic Environment and Biodiversity Report 189. Ministry for Primary Industries, Wellington 6140, New Zealand. ISSN 1179-6480 (online), ISBN 978-1-77665-718-6 (online). [Available from: http://www.mpi.govt.nz/news-and-resources/publications and http://fs.fish.govt.nz, go to Document library/Research reports].

Amandè, M.J., Chassot, E., Chavance, P., Murua, H., Molina, A.D. and de Bez, N. 2012. Precision in bycatch estimates: the case of tuna purse-seine fisheries in the Indian Ocean.  69,. *ICES J. Mar. Sci. J. Cons.* 69: 1501–10. [Available at: https://doi.org/10.1093/icesjms/fss106].

Authier, M., Rouby, E. and Macleod, K. 2021. Estimating cetacean bycatch from non-representative samples (I): A simulation study with regularized multilevel regression and post-stratification. *Front. Mar. Sci.* 8. [Available at: https://doi.org/10.3389/fmars.2021.719956].

Benoît, H.P. and Allard, J. 2009. Can the data from at-sea observer surveys be used to make general inferences about catch composition and discards? *Can. J. Fish. Aquat. Sci.* 66: 2025-39. [Available at: https://doi.org/10.1139/F09-116].

Clement, D.M. 2005. Distribution of Hector's dolphin (*Cephalorhynchus hectori*) in relation to oceanographic features. PhD thesis, University of Otago. 292pp.

Conn, P.B., Johnson, D.S., Williams, P.J., Melin, S.R. and Hooten, M.B. 2018. A guide to Bayesian model checking for ecologists. *Ecological Monographs* 88(4): 526-42. [Available at: https://doi.org/10.1002/ecm.1314].

Constantine, R., Steel, D., Carroll, E., Hamner, R.M., Hansen, C., Hickman, G., Hillock, K., Ogle, M., Tukua, P. and Baker, C.S. 2021. Estimating the abundance and effective population size of Maui dolphins (*Cephalorhynchus hectori maui*) in 2020-2021 using microsatellite genotypes, with retrospective matching to 2021. . *NZ Department of Conservation*: 66pp.

Cooke, J., Constantine, R., Hamner, R.M., Steel, D. and Baker, C.S. 2019. Population dynamic modeling of the Māui dolphin based on genotype capture-recapture with projections involving bycatch and disease risk. New Zealand Aquatic Environment and Biodiversity Report No. 216. 38 pp. [Available at: *www.mpi.govt.nz/news-and-resources/publications*].

Davies, N.M., Bian, R., Starr, P., Lallemand, P., Gilbert, D.A. and McKenzie, J. 2008. *Risk analysis for Hector's dolphin and Maui's dolphin subpopulations to commercial set net fishing using a temporal-spatial age-structured model*. Ministry of Fisheries, Wellington, New Zealand. [Available at: www.fish.govt.nz/en-nz/Consultations/Hector+new/default.htm].

Dillingham, P.W., Moore, J.E., Fletcher, D., Cortés, E., Curtis, K.A., James, K.C. and Lewison, R.L. 2016. Improved estimation of intrinsic growth rmax for long-lived species: integrating matrix models and allometry. *Ecological Applications* 26(1): 322-33

DuFresne, S.J. 2004. Conservation biology of Hector's dolphin. PhD Thesis, University of Otago. 152pp.

Edwards, C.T.T., Roberts, J.O. and Doonan, I.J. 2018. Estimation of the maximum rate of intrinsic growth for Hector's dolphin. Final Research Report available from Fisheries New Zealand. 22pp.

Faunce, C.H. and Barbeaux, S.J. 2011. The frequency and quantity of Alaskan groundfish catcher-vessel landings made with and without an observer. *ICES J. Mar. Sci. J. Cons.* 68: 1757–63. [Available at: https://doi.org/10.1093/icesjms/fsr090].

Gormley, A.M. 2009. Population modelling of Hector's dolphins. PhD thesis, University of Otago, Dunedin, New Zealand. [Available at: www.otago.ac.nz/].

Large, K., Roberts, J., Francis, M. and Webber, D.N. 2019. Spatial assessment of fisheries risk for New Zealand sea lions at the Auckland Islands. *New Zealand Aquatic Environment and Biodiversity Report* 224: 85pp.

Mackenzie, D. and Clement, D. 2016. Abundance and distribution of WCSI Hector's dolphin. *New Zealand Aquatic Environment and Biodiversity Report No. 1168 to the Ministry for Primary Industries*. 79pp. [Plus supplementary material].

MacKenzie, D.I., Fletcher, D., Meyer, S. and Pavanato, H. 2022. Updated spatially explicit fisheries risk assessment for New Zealand marine mammal populations. *New Zealand Aquatic Environment and Biodiversity Report* No. 290: 218pp. [Available at: https://www.mpi.govt.nz/dmsdocument/53239/direct].

Moore, J.E. 2015. Intrinsic growth (rmax) and generation time (T) estimates for the genera *Sousa*, *Orcaella* and *Neophocaena*. *NOAA Technical Memorandum* NOAA-TM-NMFS-SWFSC-550: 14pp.

Moore, J.E., Martin, A.R. and da Silva, V.M.F. 2018. Intrinsic growth (Rmax) and Generation Time (T) estimates for the *Inia geoffrensis*, in support of an IUCN Red List Re-assessment. *NOAA Tech. Mem.* [Available at: https://doi.org/10.7289/V5/TM-SWFSC-596].

Rayment, W.J. 2008. Distribution and ranging of Hector's dolphins: implications for protected area design. PhD thesis, University of Otago. 223pppp.

Roberts, J.O. and Webber, D.N. 2023. Local-scale spatial risk assessment of inshore commercial fisheries on Hector's dolphins. *New Zealand Aquatic Environment and Biodiversity Report* 308(April 2023): 104pp.

Roberts, J.O., Webber, D.N., Roe, W.D., Edwards, C.T.T. and Doonan, I.J. 2019. Spatial risk assessment of threats to Hector's and Māui dolphins (*Cephalorhynchus hectori*). *New Zealand Aquatic Environment and Biodiversity Report No.* 214: 175pp.

Rouby, E., Dubroca, L., Cloatre, T., Demaneche, S., Genu, M., Macleod, K., Peltier, H., Ridoux, V. and Authier, M. 2022. Estimating Bycatch From Non-representative Samples (II): A Case Study of Pair Trawlers and Common Dolphins in the Bay of Biscay. *Frontiers in Marine Science* 8: 795942. [Available at: https://doi.org/10.3389/fmars.2021.795942].

Slooten, E. and Lad, F. 1991. Population biology and conservation of Hector's dolphin. *Canadian Journal of Zoology - Revue Canadienne de Zoologie* 69: 1701-07.

Stratoudakis, Y., Fryer, R.J. and Cook, R.M. 1998. Discarding practices for commercial gadoids in the North Sea. *Can. J. Fish. Aquat. Sci.* 55: 1632–44. [Available at: https://doi.org/10.1139/f98-029].

Taylor, B.L., Chivers, S.J., Larese, J. and Perrin, W.F. 2007. Generation length and percent mature estimates for IUCN assessments of cetaceans. 24pp. US NMFS Administrative Report LJ-07-01, Southwest Fisheries Science Center, La Jolla, CA, USA.

Tremblay-Boyer, L. and Abraham, E.R. 2020. Increased fisher-reporting of seabird captures during an electronic monitoring trial. . *New Zealand Aquatic Environment and Biodiversity Report* 238: 32pp. [Available at: https://www.mpi.govt.nz/dmsdocument/39863-AEBR-238-Increased-fisher-reporting-of-seabird-captures-during-an-electronic-monitoring-trial].

Wade, P.R. 1998. Calculating limits to the allowable human-caused mortality of cetaceans and pinnipeds. *Marine Mammal Science* 14(1): 1-37.

## Adjunct 1

### New Zealand's observer programme

*'The sampling strategy of national observer programme is not described. The randomization of sampling strategy should include as many vessels as possible, ideally through relevant stratification.*

The national observer programme is established for the purposes of collecting reliable and accurate information on:
- fisheries research, management, and enforcement,
- vessel safety and employment on fishing vessels, and
- compliance with maritime rules relating to pollution and the discharge of waste material from vessels.

Decisions on <u>whether a fleet will be monitored</u> using observers are made annually. Each year, Fisheries New Zealand and the Department of Conservation plan the number of observer days needed at sea for different inshore, deepwater and highly migratory fisheries while considering the capacity of the programme and how and where resources need to be deployed.

Inshore fishing within the New Zealand Exclusive Economic Zone (EEZ) is immensely diverse, with large amounts of variation in individual practice and effort, spatially, temporally, and between methods. Characterising the inshore sector is difficult, making it challenging to draw conclusions about levels of observation both in terms of coverage and stratification of sampling without an understanding of the objectives of that monitoring in a given year and how it is being delivered.

Planning for inshore observer coverage is prioritised based on monitoring objectives for compliance needs, fish sampling requirements for stock assessment, fish catch composition, monitoring of threatened species interactions and abundance[2], and other commitments. These monitoring objectives inform the method, area, time and in some instances targeted fish species that will be observed.

Observer coverage is planned in 'seadays', which reflects days that observers are at sea. Coverage is planned for a whole year by method and area, for example 'West Coast North Island Inshore Trawl'. The seadays required to meet objectives are estimated based on the relevant effort in previous years incorporating any new fishery management changes, such as alteration to a fishing closure. Coverage is then allocated by month to assist in forward planning. Based on monitoring of coverage, the Seadays Plan can be amended and reconsulted during the year if required.

The Seadays Plan is developed and prioritised in collaboration between Fisheries New Zealand and the Department of Conservation. The size of the plan is based on the capacity of the Observer Programme to deliver the coverage goals.

Using seadays, an estimate of the total percentage of effort likely to be observed is calculated. Once a Seadays Plan has been approved for the year, there are operational decisions that are made by managers on how to deploy those days across the vessels that operate in the fishery/area and whether a vessel will carry an observer. We agree, in general, that observer coverage should cover as many vessels as possible rather than a high level of coverage on a few vessels. While we endeavour to do that, health and safety consideration (such as vessel size and watch keeping capability) or other operational matters (such as observer availability and the amount of effort each vessel contributes to a fishery and when and how frequently it operates) make that difficult to achieve across all fisheries.

For Hector's and Māui dolphins monitoring objectives we also consider:
- Targeting observer coverage in areas/seasons of highest dolphin density, i.e., where the estimated probability of capture is highest per fishing event.
- Reducing biases arising from observer effects on fisher behaviour by targeting higher proportional coverage for particular areas/seasons where dolphin density or risk of interactions is higher and no coverage in others, rather than uniformly low coverage across areas and seasons when fishers are much less likely to interact with a dolphin and observers may not be present.

General descriptions of the annual observer coverage plans and objectives, which relate to understanding the nature and extent of protected species interactions with commercial fishing activities can be found online.

---
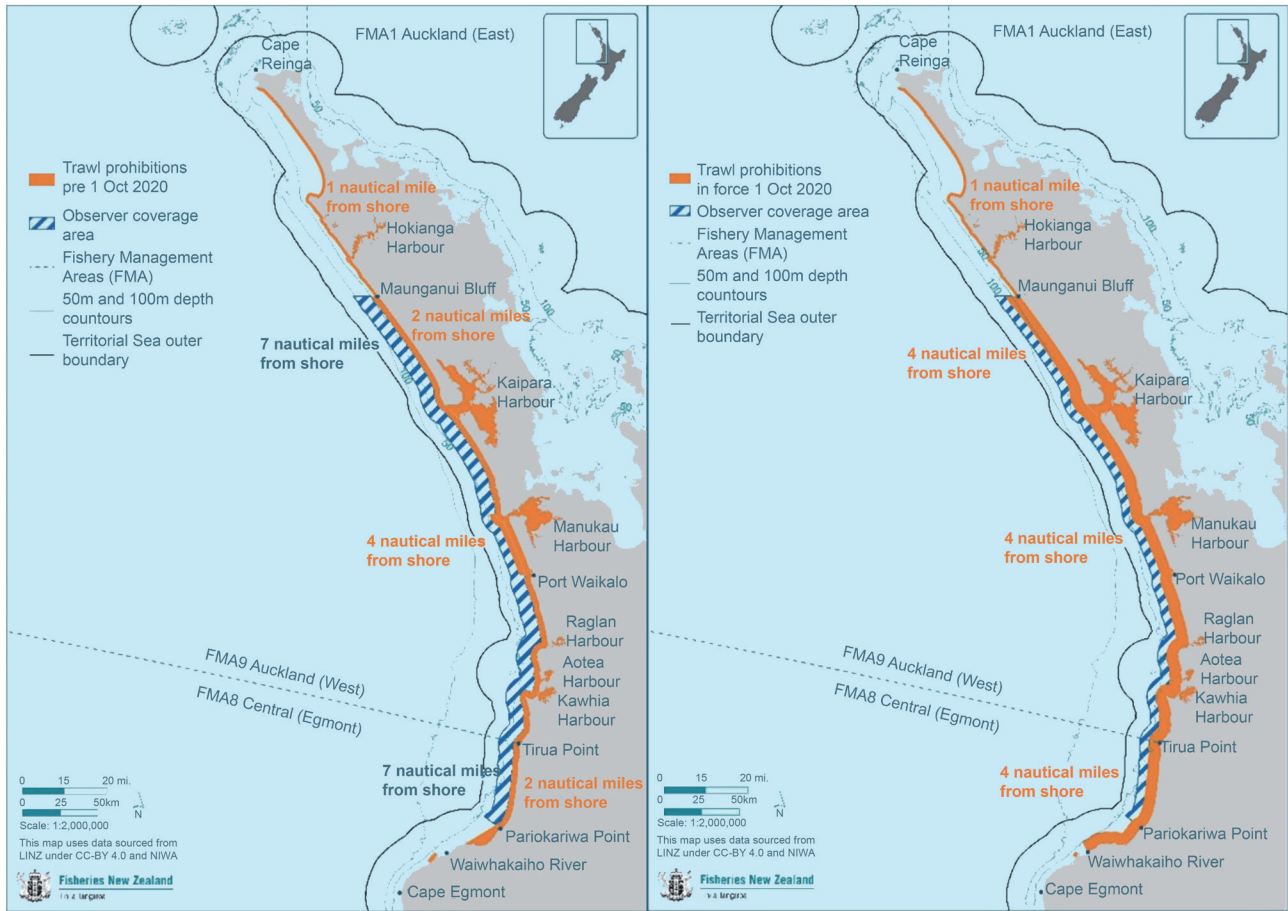
[2]Seabirds around fishing vessels (*dragonfly.co.nz*).

Fig. 1. Fishing prohibition and the targeted monitoring coverage areas for trawl within the core Māui dolphin distribution area pre-1 October 2020 (left) and from 1 October 2020 (right). Targeted monitoring coverage (delivered solely by observers before 1 November 2019) has focused on the core distribution area of Māui dolphins (Maunganui Bluff to Pariokariwa Point and offshore to seven nautical miles).
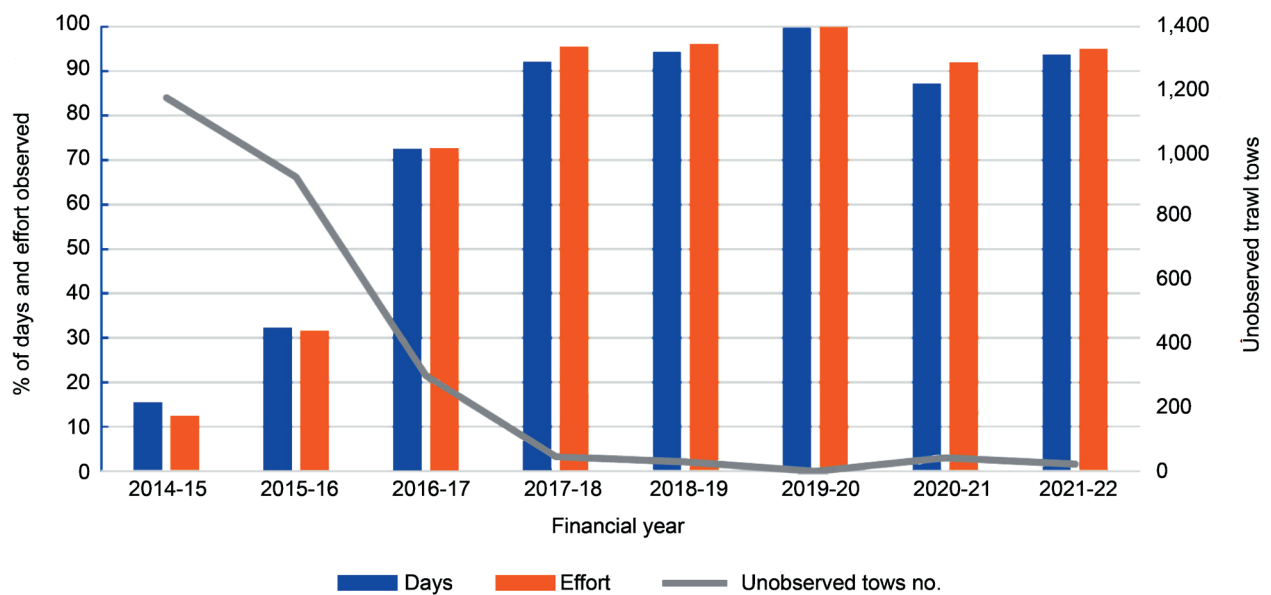


Fig. 2. Monitoring coverage (via observers and on-board cameras) of trawl vessels in the core distribution area between Maunganui Bluff and Pariokariwa Point to seven nautical miles offshore. Coverage is shown as a percentage of fishing days and effort (tows) observed (primary y-axis) relative to the number of unobserved tows (secondary y-axis). Financial year refers to the period from 1 July to 30 June.
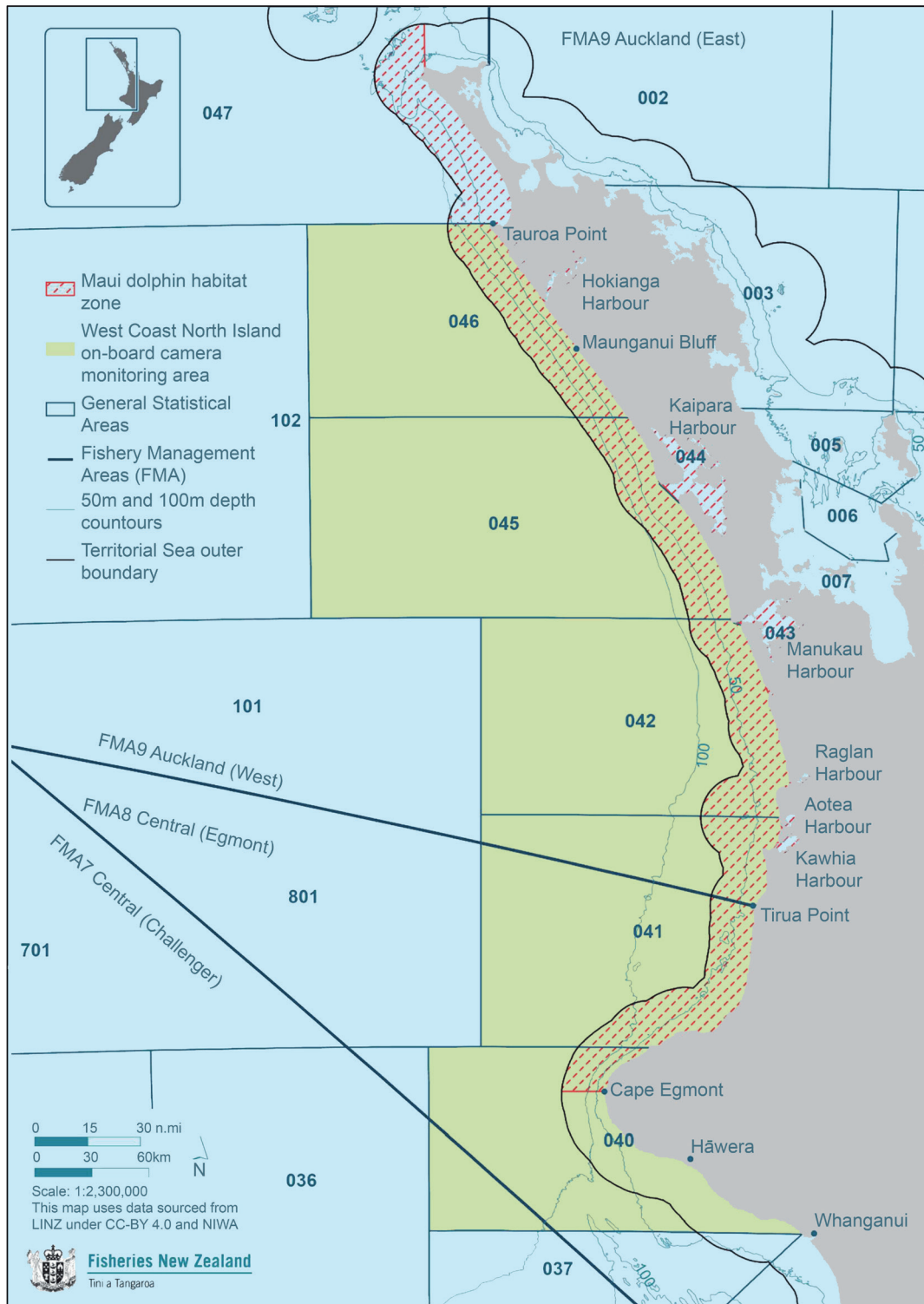
Fig. 3. West Coast North Island proof-of-concept on-board camera monitoring area (fisheries statistical areas 040, 041, 042, 045 and 046) relative to the Māui dolphin habitat zone.
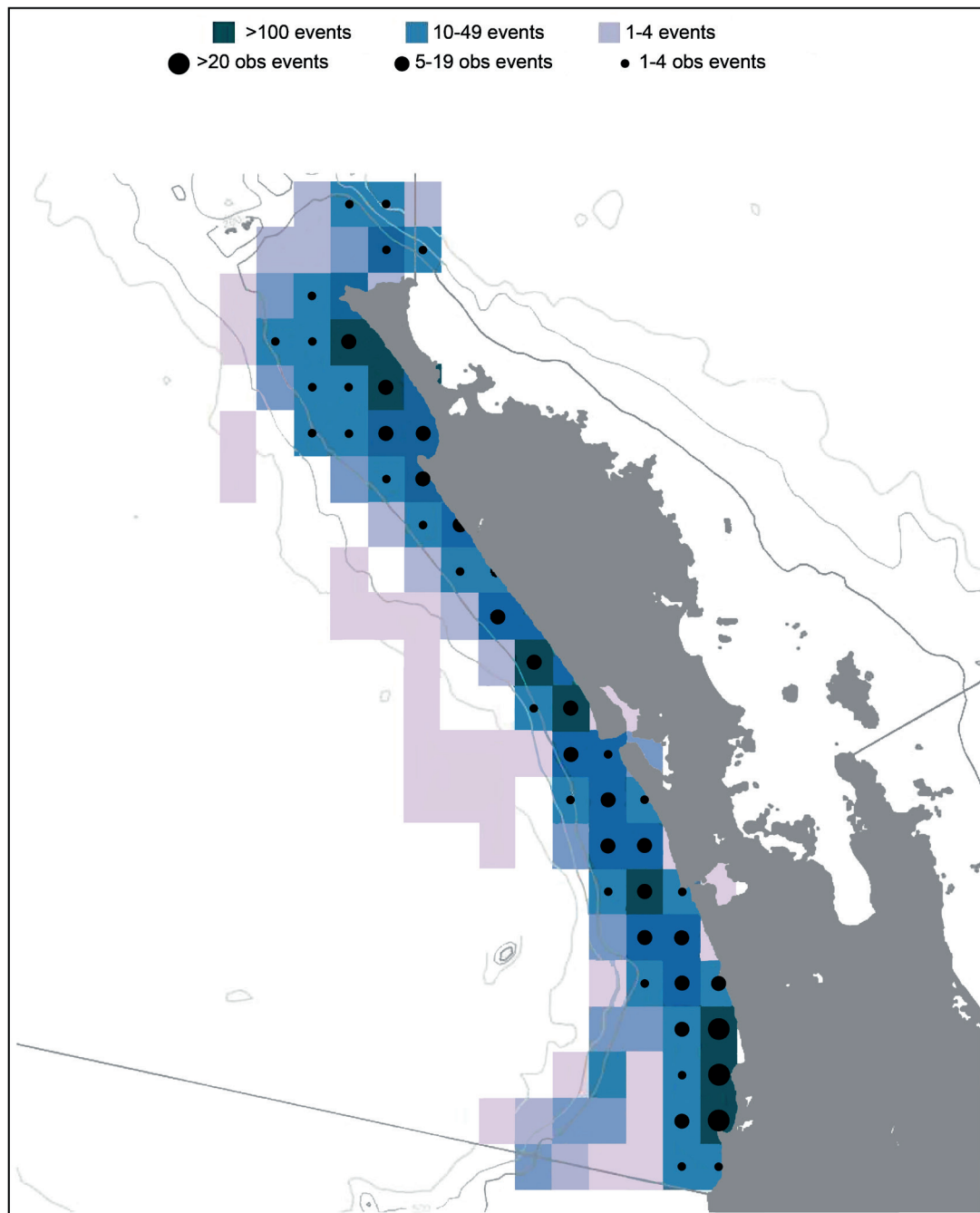
Fig. 4. Map of trawl fishing effort 2002 to 2020. Fishing effort is mapped into 0.2-degree cells, with the colour of each cell being related to the amount of effort. Observed fishing events are indicated by black dots. (source: Fisheries New Zealand, https://protectedspeciescaptures.nz/PSCv6/released/ hectors-dolphin/inshore-trawl/all-vessels/west-coast-north-island/2002-03-2019-20/summary_map.html)

**Examples of monitoring coverage in target areas and future monitoring**

Following a review of the Māui dolphin portion of the Hector's and Māui dolphin Threat Management Plan in 2012-13, the subsequent monitoring programme had two focal areas.

(1) Ramping of observer coverage on trawl vessels in the core Māui dolphin distribution area (from Maunganui Bluff to Pariokariwa Point and between 2 and 7 n.miles offshore) to reach approximately 100% coverage within four years.

(2) Mandatory observer coverage on set net vessels operating from Pariokariwa Point to Hawera between 2 and 7 n.miles. (This coverage ceased when the area was closed to set net fishing in October 2020).

Monitoring was targeted in these areas where the residual risk of a fishing-related mortality was estimated to be greatest and was delivered only using observers until 2019 when on- board cameras were regulated and required on some inshore trawl and set net vessels.

Approximately 75.9% (summer) and 62.3% (winter) occupancy of the Māui dolphin distribution is estimated to fall within the area from Maunganui Bluff to Pariokariwa Point and offshore to seven nautical miles (Figure 1). This core distribution area has been subject to increased fishing prohibitions since 2003 and a significant ramping of observer coverage on trawl vessels beginning in 2014, with over 90% coverage each year since 2017 (Figure 2).

The implementation of a proof-of-concept (POC) on-board camera programme in November 2019 has provided for more extensive (alongshore and offshore) monitoring coverage compared to the focal areas above. However, the POC programme was restricted to vessels that met the regulatory requirements:

• any set net or trawl vessel (≥8 and ≤29 metres in registered length) that:

  • operated in the defined monitoring area that is fisheries statistical areas 040-042, 045 and 046; and

  • operated in that area between 1 October 2017 and 30 September 2018.

The POC programme did not cover all set net and trawl vessels that could operate in the defined monitoring area (fisheries statistical areas 040-042, 045 and 046, Fig. 3). Consequently, formal percent monitoring targets in this broader area were not set but it has enabled additional monitoring information to be gathered from vessels operating in areas deemed to be of even lower risk to Māui dolphins.

The focal areas above remained the primary objectives where full coverage was targeted, and the target coverage on trawl vessels in the core Māui dolphin distribution area shifted to being delivered using a mix of on-board cameras and observers. The use of observer coverage has continued alongside camera deployment to test the effectiveness of on-board cameras and their ability to detect protected species interactions. No captures or interactions have been observed by observers or in footage review.

The Protected Species Capture website also provides a map of observer coverage for the west coast North Island trawl fishery and shows a good level of spatial coverage along the coast (refer to Fig. 4 below). However, we note that the monitoring coverage from on-board cameras is not reflected in this data, the cumulative effort as shown does not reflect the expansion of trawl closures during this period.

*Expansion of the on-board camera programme*

On 17 June 2021, New Zealand's Minister for Oceans and Fisheries announced that up to 300 inshore fishing vessels will be fitted with cameras by the end of 2024. The following vessels will be covered under the expanded programme:

• trawl vessels less than or equal to 32 metres in overall length (except those targeting scampi),

• set net vessels greater than or equal to 8 metres in overall length, and

• vessels using surface longline, bottom longline, purse seine and Danish seine fishing methods.

The expansion of the programme applies to the whole of New Zealand and roll out has been prioritised based on risk to protected species. Information on the timing of this programme, and when certain vessels will have cameras installed can be found on the MPI website[3].

The data from the POC on-board camera programme has not been used to gather data on mortality and injury with which to inform bycatch estimates in the Roberts *et al.* (2019) SEFRA model, but with the ongoing expansion of the programme data from the on-board camera programme will be used in future. This data should also improve our ability to address the main biases in observer programmes.

---

[3]https://www.mpi.govt.nz/fishing-aquaculture/commercial-fishing/fisheries-change- programme/on-board-cameras-for-commercial-fishing-vessels/.