

Seahorses (*Hippocampus* spp.) as a case study for locating cryptic and data-poor marine fishes for conservation

L. Aylesworth¹, T.-L. Loh^{1,2}, W. Rongrongmuang^{1,3} & A. C. J. Vincent¹

¹ Project Seahorse, University of British Columbia, Vancouver, Canada

² Daniel P. Haerther Center for Conservation and Research, John G. Shedd Aquarium, Chicago, IL, USA

³ Sarasin Divers, Phuket, Thailand

Keywords

detection; *Hippocampus* spp.; sampling protocol; relative abundance; cryptic marine species; data-poor species.

Correspondence

Lindsay Aylesworth, Project Seahorse, University of British Columbia, 2202 Main Mall, Vancouver, Canada V6T 1Z4.
Email: l.aylesworth@oceans.ubc.ca

Editor: Nathalie Pettorelli

Associate Editor: Lynne Shannon

Received 31 May 2016; accepted 12 December 2016

doi:10.1111/acv.12332

Abstract

When seeking to conserve data-poor species, we need to decide how to allocate research effort, especially when threats are substantial and pressing. Our study provides guidance for sampling marine fishes that are particularly difficult to find – those species that are cryptic or rare and/or where little information exists on local distribution (data-poor). We used our experience searching for seahorses (*Hippocampus* spp.) in Thailand to evaluate two search strategies for marine conservation: (1) determining relative abundance and (2) searching for presence/absence with detection probabilities. Our fieldwork indicated that using the presence/absence framework was more likely to lead to inferences that seahorses could be found in the site than when using the relative abundance framework. This realization would support a commonsense approach, where presence/absence with detection probabilities is centrally important to marine conservation planning for cryptic and/or data-poor marine species.

Introduction

Cryptic data-poor species are not well served by classic ecological techniques for determining relative abundance. In the absence of baseline information, pilot studies are often conducted (Fairweather, 1991) to determine relative abundance to inform research or to create management plans. But what happens when pilot studies fail to find the species of interest, which may commonly happen for cryptic or depleted species (Engler, Guisan & Rechsteiner, 2004; Durso, Willson & Winne, 2011)? Such failures to locate cryptic species may result from limited data, linked to funding or logistic restraints resulting in a type II error (Underwood, 1996) or because the species was not detected or is truly not there (Mackenzie *et al.*, 2002). Type II errors are problematic in the conservation sense because when populations and their habitats are not identified, it leads to a biased understanding of the status of the resource and the need for management or protection. We need to deploy rigorous yet realistic methods to assess conservation status and enact management for data-poor species.

On land, researchers are finding ways to cope with data collection for relative abundance of cryptic, data-poor species. To overcome difficulties of terrestrial sampling, scientists have used camera traps, scat sampling and genetic analysis to make population or relative abundance estimates for species in need of conservation (Scotts & Craig, 1988;

Foran, Crooks & Minta, 1997; Kelly *et al.*, 2013). Most of these methods rely on count statistics (e.g. number of animals caught, seen and heard) to make inferences about the species of interest (MacKenzie *et al.*, 2006). Count statistics are generated by the product of the quantity of interest (e.g. population/abundance at a site) and the probability of detection (e.g. how likely you are to count it, if present) (MacKenzie *et al.*, 2006). Many relative abundance methods do not explicitly account for the probability of detection and assume that by using standardized survey methods, detection probabilities are similar among sites or equal to 1 (MacKenzie *et al.*, 2006). However, if detection probability is less than 1 or varies among sites, and is unaccounted for, false inferences about the species of interest may be made (MacKenzie *et al.*, 2006).

Studies aimed at estimating species occupancy (psi or presence and absence), given the challenges of finding the species of interest (detection or P), have increased considerably in the terrestrial ecology and conservation literature over the past 12 years (Mackenzie *et al.*, 2002; Hurme *et al.*, 2005; Mazerolle *et al.*, 2007; Petracca, Ramirez-Bravo & Hernandez-Santin, 2013; Robley *et al.*, 2014). Occupancy studies provide a useful alternative to estimates of abundance when there may be challenges detecting the species of interest such as for rare, cryptic or depleted species, because they explicitly account for detection probabilities (MacKenzie

et al., 2006). Such studies estimate the likelihood that the species was present at a site, given the data, which typically consists of counts of individuals of the given species that include zeros or unequal sampling effort (Mackenzie *et al.*, 2002). This can be useful because one can estimate the likelihood that the species was present at a site where it was not observed (MacKenzie *et al.*, 2006). Pilot studies that fail to find the species of interest would be uninformative for determining sites for future research if unable to account for detection probabilities. In addition, investigations determining occupancy (presence/absence) across a landscape typically require less resources and time than estimating density, making them useful for conservation (Gaston *et al.*, 2000).

Studies using presence/absence with detection probabilities have generated sound science for management decisions on a range of terrestrial species. Guidance on the implementation of management measures such as habitat corridors (Petracca *et al.*, 2013), predator control (Robley *et al.*, 2014) and invasive species remediation (Britton, Pegg & Gozlan, 2011) stem from the results of detection research. Management decisions regarding the evaluation of monitoring programs or procedures (de Solla *et al.*, 2005; van Strien *et al.*, 2010), conservation planning (Hurme *et al.*, 2005; Weller, 2008) and effectiveness of conservation measures (Zipkin *et al.*, 2010) have been supported by studies using presence/absence with detection probabilities. The increasing use of studies using presence/absence with detection probabilities to inform terrestrial management activities may represent a paradigm shift or simply a response to the challenges associated with data-poor species (Underwood & Chapman, 2003; Mackenzie *et al.*, 2005; McNie, 2007).

In the ocean, methods for assessing data-poor species are still in their infancy compared to those established for terrestrial counterparts (Moblely, Small & Jones, 2011; Katsanevakis *et al.*, 2012). The aquatic environment is more complex because researchers cannot easily see what lies beneath the surface (Monk, 2013), and the technology for sampling and surveying is less efficient and more costly than in terrestrial systems (Katsanevakis *et al.*, 2012). This adds important concerns about estimates of population density, abundance and simple presence/absence data (Issaris *et al.*, 2012; Monk, 2013). When animals are rare, cryptic or depleted, revealing their presence requires large amounts of effort (Maxwell & Jennings, 2005; Guillera-Arroita & Lahoz-Monfort, 2012). The use of some terrestrial methods is becoming more common in the ocean. Examples include underwater videos at baited sites (Goetze & Fullwood, 2012), scat sampling (Bowles & Trites, 2013) and genetic analysis (Moblely *et al.*, 2011). However, these methods are not widely employed for cryptic data-poor marine species (Katsanevakis *et al.*, 2012). Although there is literature on the use of occupancy modeling with detection probabilities in freshwater environments (Anderson *et al.*, 2012; Rodtka *et al.*, 2015), its use in marine sampling is uncommon even though many aspects of the methods are transferable. Only one study has compared presence/absence with detection probabilities in relation to two diver-based methods – stationary census, where a diver remains stationary and records species within a given distance, and belt transects, where a

diver swims along a transect at a constant speed recording species within a fixed distance of the transect (Green *et al.*, 2013). The results indicated that both diver-based methods had poor detection abilities (Green *et al.*, 2013). With continuing marine species decline (McCauley *et al.*, 2015), managers urgently need new ways to handle data scarcity and facilitate effective decision making for the ocean.

Seahorses (*Hippocampus* spp.) are such a challenging fish genus to locate that they are good representative candidates for a detection study. They are difficult to find in the wild because they camouflage well, have a patchy distribution and are found in low numbers (Foster & Vincent, 2004). Seahorses have small home ranges (Perante *et al.*, 2002) and live in shallow, coastal habitats, including coral reefs, seagrass beds, mangroves and sandy soft bottoms (Lourie *et al.*, 2004). Worldwide there are 41 species of seahorses (Lourie, Pollom & Foster, 2016). The IUCN Red List of Threatened Species (www.redlist.org) lists 12 seahorse species as threatened with 28 classified as 'Data Deficient'.

We here conduct the first study to focus on using presence/absence with detection probabilities for cryptic, data-poor marine fishes. Other species-specific detection research has focused on comparatively large marine animals including marine mammals and sharks, or lionfish (Williams & Thomas, 2009; Green *et al.*, 2013). Both studies found that methods incorporating detection probabilities improved their ability to make inferences about rare (marine mammals) or cryptic (lionfish) species (Williams & Thomas, 2009; Green *et al.*, 2013). As seahorses can also be considered both rare and cryptic (Foster & Vincent, 2004), we suspected that using such methods would improve our ability to study these data-poor fishes. We here simultaneously explored efforts to locate seahorses and contribute to the literature on search methods and detection for cryptic and data-poor marine fishes. First, we evaluated the success of traditional searches aimed at determining relative abundance, which do not account for imperfect detection, to locate our species of interest. Next, we compared two methods – belt transects, where the search area is fixed along a transect, and timed swims, where the search area is unrestricted – for their efficiency to locate cryptic, data-poor marine fishes. We then re-evaluated our datasets comparing belt transects and timed swims, with presence/absence models with detection covariates that enable us to determine if our original estimates of occupancy were biased. Finally, we explored how studies using presence/absence with detection probabilities can be incorporated into future marine research for data-poor species with simulations, sensitivity analyses and a cost–benefit analysis.

Materials and methods

Data collection

Determining relative abundance assuming detection probability = 1

Our search to determine relative abundance of seahorses began on the west (Andaman Sea) coast of Thailand using common underwater survey protocols based on Reef Check (Andrew &

Mapstone, 1987; Hodgson, 2004). Our surveys took place from February to May 2013 and covered 19 sites, across all six Andaman provinces (Fig. 1), in mangrove, seagrass and coral reef habitats. We could not select sampling sites in the national parks because of permit restrictions. Therefore, sites were chosen based on the best available habitat outside national parks, according to local expert opinion (from marine biologists, fishers and dive operators). We conducted searches with belt transects at 10 sites and timed swims at nine sites (Table S1). Our multiple searches (mean 8, range 1–20) using snorkel and scuba at each site covered a mean area of *c.* 1100 m² but ranged between 500 and 2000 m² (Table S2).

The challenges encountered on the Andaman coast (see Results) led us to alter the methods for surveying the east

(Gulf of Thailand) coast. We began by making a concerted effort to extract seahorse sightings from all available information including, but not limited to, Internet resources, popular media, survey reports, published papers and direct contact with scuba divers and dive shops. We also encouraged people to report their observations to the global database of seahorse sightings, iSeahorse.org. We used the responses from this outreach to select 35 sites in the Gulf of Thailand. Our search efforts on the eastern coast took place in September and October 2013 still with the aim of determining relative abundance of seahorses at sites using Reef Check protocols (Hodgson, 2004). We conducted only timed swims for each site. The number of repeat searches per site varied from one to six (mean of two) based on logistic

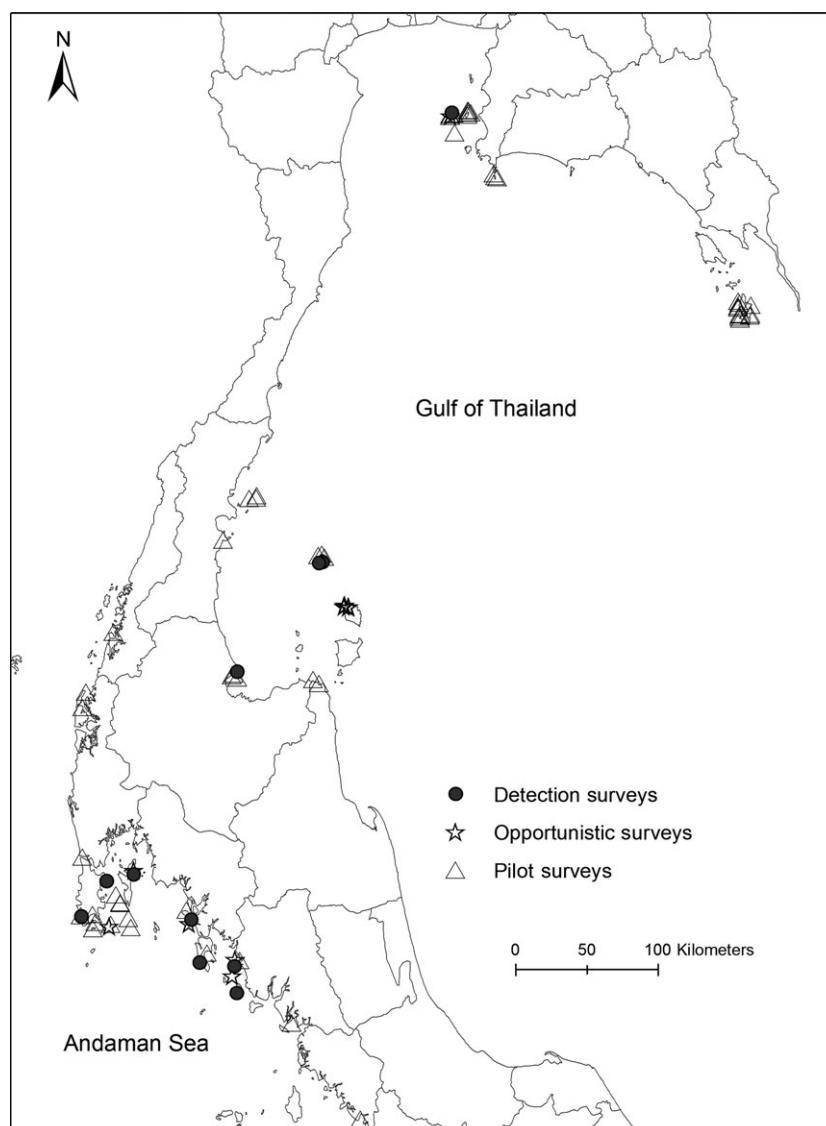


Figure 1 Survey efforts on Thailand's Gulf and Andaman coasts to determine relative abundance or presence, absence and detection. Sites visited opportunistically for presence, absence and detection are also included.

constraints and the need to visit as many sites as possible for wider geographic coverage (Tables S1 and S2).

Comparing methods assuming detection probability <1

Based on the results of our attempts to determine relative abundance, we focused our second field season on comparing the relative efficiency of belt transects (searching in a fixed area) and timed swims (searching in an unrestricted area) – controlling for search time in both – as underwater search methods to locate seahorses (Andrew & Mapstone, 1987). We chose research sites based on the knowledge derived from the aforementioned outreach efforts (Fig. 1). Our challenges in locating our species of interest with Reef Check search protocols based on relative abundance in the first season highlighted the importance of defining the size of our search site. We, therefore, defined each site as an area roughly 75 000–80 000 m² (c. 400 m × 200 m, or 8 ha). Fieldwork occurred from March to July 2014 at 10 sandy, soft bottom sites – seven on the Andaman Coast and three on the Gulf Coast. We chose sandy, soft bottom sites (known habitat for five of the seven Thai seahorse species; Lourie *et al.*, 2004) to maximize our chances of locating a seahorse; seahorses are easier to find in this type of habitat because it has low relief and rugosity and few places to hide. We conducted five replicates of each method per site. An additional 10 sites (five on each coast) were visited opportunistically, but not surveyed fully because of logistic constraints (Table S1).

We surveyed sites over a 3- to 4-day period to minimize the probability of changes in seahorse occupancy (presence or absence) over the sampling period. For transects and timed swims, we generated a random start point, using a random number table, within the site for each of the five replicate searches. All replicates per site were conducted for the same amount of total search time (ranging from 50 to 75 min depending on site depth and allowable dive time), with divers swimming in the same general direction. For example, if a site had a depth that allowed a dive time of 60 min, regardless of where the starting point was located at site X, all timed swims headed in a northern direction and lasted for 60 min. We measured the distance covered by timed swims with calibrated fin kick cycles (to determine length) and assumed a search width of 2 meters. Transect surveys consisted of as many 2 × 50 m transects as dive time would allow (ranging from two to five). We laid out transects perpendicular to the water current direction, in alternating directions, with 50 m in between each transect. For each method, upon finding a seahorse, the species, time and distance covered were recorded.

Data analysis

Determining relative abundance assuming detection probability = 1

First, we determined the proportion of sites where seahorses were located, for both the Andaman and Gulf Coasts. This calculation assumed that any zeros represented a true absence of

seahorses. Next, we identified the number of sites on each coast where local fishers or divers reported seahorses as common (based on fieldwork notes) but where we observed no seahorses in our search efforts. These sites would suggest that seahorses were indeed present, but our search efforts were not robust or sensitive enough to detect them.

Comparing methods assuming detection probability <1

We analyzed our data based on search times, search distance and survey effort with homoscedastic t-tests, to determine which method was more effective at finding the first and second seahorses at a site (Zar, 1999). For this analysis, we used the data from our second field season, but only from sites where seahorses were observed. We also conducted a post hoc power test to determine if we had enough power to determine if there was a difference in the proportion of surveys with seahorses per site between timed swims and transects for a small effect size (0.2, Cohen, 1992) at the $\alpha = 0.2$ level (Zar, 1999).

Re-evaluating with occupancy models when detection probability <1

We re-evaluated our data from relative abundance surveys and from comparing belt transects and timed swims, with occupancy models to explore how our results might differ if detection probabilities were accounted for in our analysis. We defined detection probability as the probability of observing a seahorse in any given swim/transect given that it occurred at a site. For our relative abundance dataset (initial assumed detection probability = 1), we used single season occupancy models to generate estimates of detection probabilities <1 (Mackenzie *et al.*, 2002). This procedure uses the observed data from multiple dives at a site to generate a probability estimate of detection (Mackenzie *et al.*, 2002). Next, we compared our estimates of seahorse occupancy across the seascape (proportions of sites where we located seahorses) to identify how site-specific occupancy varies with and without accounting for detection probabilities. Occupancy models were run with the software Presence (Hines, 2006).

We also explored the possibility that survey method, visibility or time of day influenced our detection probability of seahorses by creating single season occupancy models using each of these factors as detectability covariates. We hypothesized that these three variables would be most likely to affect the efficiency and reliability of finding a small cryptic fish, given our decision to search in sandy soft bottom habitats. We also created multistate, single season models to explore the relationship between detection probabilities and relative abundance (high vs. low) of seahorses per site (Royle & Nichols, 2003). We defined a site as high abundance if we observed four or more seahorses during any given search effort (timed swim or transect). This is because seahorses are commonly found in pairs and rarely found in groups of three or more (Perante *et al.*, 2002). A site was defined as a low abundance site when three or fewer

seahorses were observed during any given search effort. We conducted all analyses with two datasets from 2014 – research sites only ($n = 10$) and research sites plus opportunistically visited sites ($n = 20$) (Table S1). The presence/absence of seahorses at each site was determined by the pooled data from all surveys with both protocols.

Sensitivity analyses and future study design assuming detection probability <1

We explored how an occupancy and detection framework could be incorporated into planning for future marine research. We ran a cost–benefit analysis with several sampling design scenarios to determine the most cost-efficient study design for a data-poor species. There are two main study designs for occupancy and detection models (Mackenzie & Royle, 2005). A ‘standard design’ is one where all sites are surveyed the same number of times (Scenario 1), while a ‘removal design’ is one where surveying stops once the species of interest is found or until a pre-determined maximum number of replicates has been reached (Scenario 2 and Scenario 3, respectively). For our simulations, we calculated the number of recommended replicate searches for a 90% probability that any zeros would represent true absence at a site (Mackenzie & Royle, 2005). This procedure required us to input initial estimates of occupancy and detection, and we used estimates from one of our single season models. This calculation resulted in five replicate searches for the standard design and a maximum of seven searches for the removal designs. Scenario 1 is a typical standard survey design where all sites are searched five times. Scenario 2 is a removal design simulation where a small, cryptic data-poor marine species is found within the first day of searching at each site – before our maximum replicate search number is reached. Scenario 3 is a removal design simulation when the focal species is not observed in early searches. It takes 2 or more days to either observe the species or reach the maximum number of pre-determined replicates for all sites. For this scenario, we estimated the number of days it would take (2+) to reach this number (seven replicates) based on observations from our second research season.

To build on the results of simulations and survey design, we wanted to identify the most important factor driving the optimal number of sites to survey. This required us to conduct a sensitivity analysis. We evaluated the change in number of sites to survey with changes in initial occupancy estimates, detection probabilities and desired confidence level

in future estimates of seahorse presence or absence (occupancy) per site (Mackenzie & Royle, 2005).

Results

Determining relative abundance assuming detection probability = 1

Searching to determine relative abundance among sites resulted in low numbers of seahorses during our first field season. A total of 10 individual seahorses was found on the Andaman Coast at 36% ($n = 7$ of 19) of sites surveyed (Table S2). These results produced relative abundances with high levels of variation and failed to identify sites for future research because the most seahorses observed at any site was three and the majority of sites had zero. In comparison, on the Gulf Coast, a total of 39 individuals was found across 38% of the sites visited ($n = 10$ of 35) (Table S2). We gained some ground in that these results identified two sites for future research because two sites had more than 10 total seahorses observed.

When we compared our search results to reports of where seahorses were common (Table S2), we found discrepancies at 26% ($n = 14$ of 54) of the sites we surveyed. At 14 sites (five on the Andaman and nine on the Gulf) where seahorses were not encountered, local fishers or divers reported seahorses as present or common based on capture in fishing nets or tourism experience (Table S2). These results highlighted that at these sites (and potentially others), our search efforts were not sufficient to detect seahorse occupancy. It remained unclear, however, what level of search effort was actually needed.

Comparing methods assuming detection probability <1

On average, timed swims were more effective than belt transects at finding seahorses based on analysis with effort metrics (number of surveys, time searched and area covered) (Table 1). Timed swims found seahorses in fewer surveys and a shorter amount of time than transects (Table 1). However, transects found seahorses after searching a smaller area than timed swims (Table 1). The amount of time and distance to the second seahorse observed was shorter on timed swims than on transects (Table 1). Homoscedastic t-tests found no significant difference in the total number ($P = 0.59$) or proportion of searches that found seahorses ($P = 0.43$) between transects and timed swims. A post hoc

Table 1 Mean number of surveys, minutes and area covered before observing a seahorse (*Hippocampus* spp.) at a research site for belt transect and timed swim search methods

| Method | Mean (SE) # of surveys to 1 st seahorse sighting | Mean (SE) Time to 1 st sighting (minutes) | Mean (SE) distance covered to 1 st sighting | Mean (SE) time of 2 nd sighting (minutes) | Mean (SE) distance covered to 2 nd sighting |
|-------------|---|--|--|--|--|
| Transect | 2.3 (0.5) | 122.5 (33) | 891 m ² (237) | 49.8 (21) | 343 m ² (141) |
| Timed swims | 1.5 (0.3) | 72.0 (24) | 1095 m ² (459) | 17.8 (11) | 184 m ² (145) |

power test indicated that we had 24% power to detect a difference in the proportion of surveys with seahorses per site for each search method.

Re-evaluating with occupancy models assuming detection probability <1

Relative abundance dataset

Our re-evaluation of relative abundance data revealed that estimates of detection probability, the probability of observing a seahorse in any given swim/transect given that it occurred at a site, were <1 (Fig. 2). We found overlapping confidence intervals in detection probabilities – for timed swims (0.23 ± 0.07) and transects (0.12 ± 0.10) on the Andaman coast, and timed swims (0.41 ± 0.16) on the Gulf coast (Fig. 2). Once we factored in a detection probability of <1, our estimates of sites occupied by seahorses were higher than our initial estimates (Andaman coast occupancy $58.5 \pm 29\%$, Gulf Coast $73.9 \pm 26\%$), but still maintained high standard errors.

Comparing methods dataset – belt transects and timed swims

Drawing in prior information (garnered through outreach activity) increased our probability of finding seahorses at

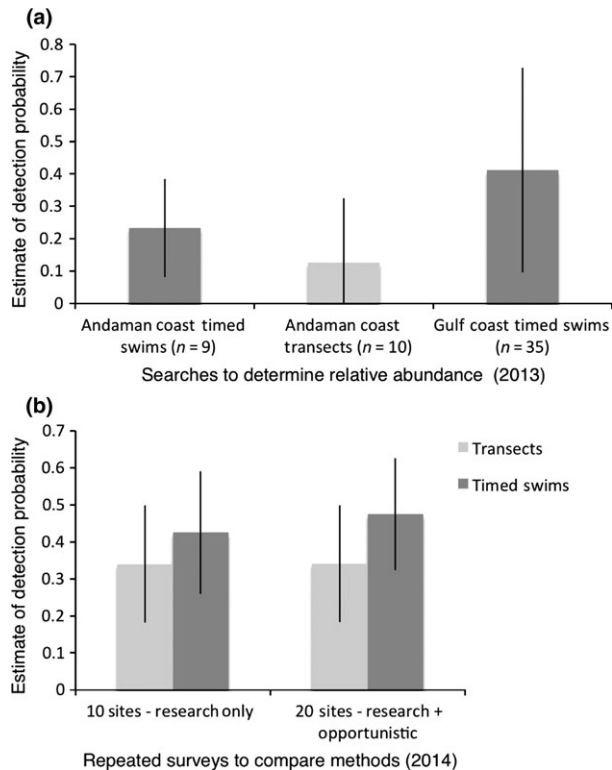


Figure 2 Estimates of detection probabilities ($\pm 95\%$ confidence intervals) of search methods (belt transects vs. timed swims) based on (a) relative abundance (2013) and (b) methods comparison (2014) datasets.

research sites. More seahorses were found during the 2014 research season; a total of 69 seahorses were observed, at 70% (7 of 10) of sites surveyed. By including the sites we visited opportunistically, a total of 98 seahorses at 11 of 20 (55%) surveyed sites were found. When we accounted for detection probabilities <1, our estimates of site occupancy increased with our dataset with 20 sites (65%) and remained the same with 10 sites (70%).

The only factor significantly influencing detection probabilities was abundance as seen in our multi-state single season model (Fig. S1). Confidence intervals of detection probabilities for sites with high and low abundances of seahorses did not overlap when we used our largest dataset (20 sites, refers to research plus opportunistic sites). When we decreased the number of sites in our dataset to 10 (research only), confidence intervals of detection probabilities overlapped for sites with high and low abundance of seahorses. With our single-state, single season models, we found overlapping confidence intervals between timed swims and transects for both our datasets, although timed swims had consistently higher detection probabilities (Fig. 2). Sites with higher total numbers of seahorses had higher detection probabilities for both methods (Fig. 3). Among the single-state, single season models, the model with the greatest support did not include visibility or time of day (Table 2).

Sensitivity analyses and future study design for when detection probability <1

A removal sampling design is the most cost-efficient way to sample for data-poor seahorses (Table 3). The average cost per day for research in the 2014 field season was US\$177.55 (Table 3). The total savings for a removal design with minimal sampling would have provided the opportunity to visit an additional 29 sites. The total savings for a standard sampling design or for a removal design with maximum sampling were close to equal, as were the additional number of sites to visit (Table 3).

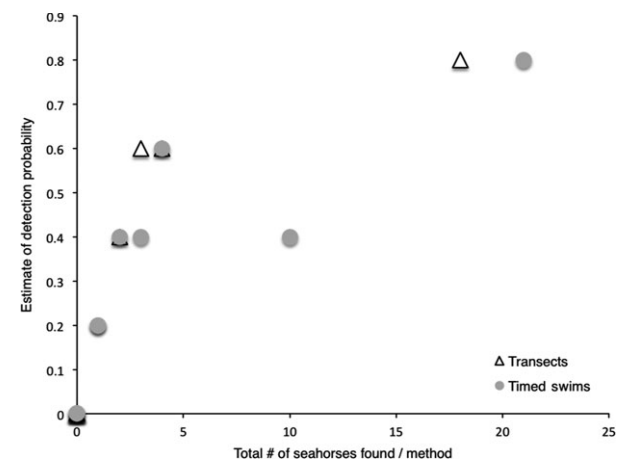


Figure 3 Site detection probability based on the total number of seahorses found per site ($n = 10$) by search method based on 2014 research data.

Table 2 Results of single season occupancy models exploring the effects of method, time of day and visibility on detection probabilities compared with the simplest model with a constant detection covariate

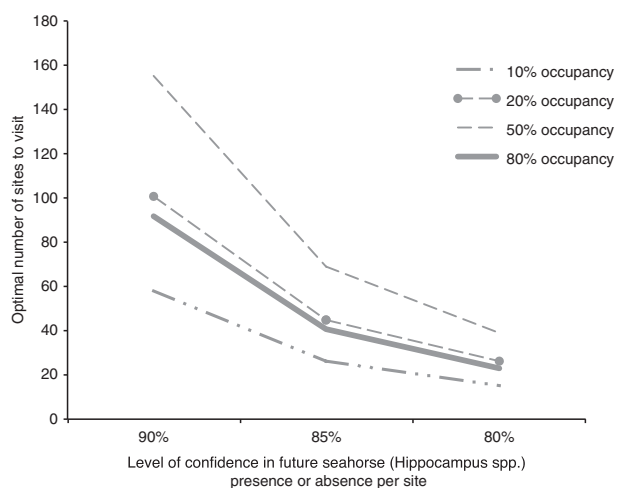
| 2014 research sites | Δ AIC | Support for model (ω) | Number of variables (K) | Occupancy (\pm SE) | Detection rate (\pm SE) |
|---------------------------------|--------------|--------------------------------|-------------------------|-----------------------|--|
| Simplest (null) model | 0.00 | 0.4449 | 2 | 0.7057 (0.14) | 0.38 (0.05) |
| Detection varies by method | 1.46 | 0.2144 | 3 | 0.7055 (0.14) | 0.34 (0.08) Transect 0.42 (0.08) – Timed swim |
| Detection varies by time of day | 1.88 | 0.1738 | 3 | 0.7054 (0.14) | 0.40 (0.08) – AM 0.36 (0.08) – PM |
| Detection varies by visibility | 1.96 | 0.167 | 3 | 0.7053 (0.14) | 0.39 (0.07) – <1 m 0.39 (0.08) – >1 m |

Table 3 Costs and savings of three survey designs scenarios (Mackenzie & Royle, 2005) to search for seahorses (*Hippocampus* spp.) based on per diem costs of 2014 field season

| | Scenario 1 Standard design | Scenario 2 Removal design (species found in 1 st day) | Scenario 3 Removal design (species found in 2 ⁺ days or max. searches reached) |
|--|-------------------------------|---|--|
| Estimated cost per day from 2014 field season | \$178 | \$178 | \$178 |
| Simulated total number of days needed to survey 20 sites | 34 | 20 | 36 |
| Simulated total cost to survey 20 sites | \$6036 | \$3551 | \$6391 |
| Total savings compared to actual 2014 research costs to survey 20 sites | \$2663 | \$5149 | \$2308 |
| Simulated additional sites to visit using total savings | 15 | 29 | 13 |

The standard design (Scenario 1) employs the same number of replicates at all sites, while the removal survey stops surveys at a site once a seahorse has been found or until a pre-determined maximum number of replicates has been reached (Scenario 2/Scenario 3).

Simulations indicated that the factor most influencing the optimal number of sites to visit to understand occupancy (presence/absence at sites) across a seascape was the acceptable level of confidence in the occupancy estimate (Fig. 4).

**Figure 4** Simulation results showing the number of sites to visit increase with increasing confidence in the occupancy estimate of seahorses.

As the acceptable level of confidence increased, the number of optimal sites to visit decreased (Fig. 4). In terms of identifying the optimal number of replicate surveys needed to obtain confidence that zeros represent absence at a site, our simulation at 90% confidence indicated that as detection probability increased, the number of required replicate surveys decreased (Fig. S2).

The presence/absence detection framework was the most informative for effective initial management. Executing timed swims only identified two sites for future research and or management, but could not be used to determine most effective search method or overcome issues of detection. Our search based on relative abundance on the Andaman Coast where we conducted both haphazard transects and timed swims did not yield enough information about seahorses for initial management measures. When we conducted our research to compare methods in a presence/absence detection framework, we identified three sites for future research because of the high number of seahorses observed at these sites, provided supporting evidence that both methods were equally effective at finding seahorses at sites with high abundance, and provided information to deal with issues of non-detection.

Discussion

Our research highlights that using a framework based on presence absence with detection probabilities (Mackenzie

et al., 2002) is more useful for cryptic data-poor marine fishes than one using relative abundance, despite the latter being more widely used in marine conservation research (Andrew & Mapstone, 1987). The presence/absence/detection probability framework enabled us to determine the probability that sites with zeros represented the true absence of species given our sampling efforts, as commonly done in terrestrial detection studies (MacKenzie *et al.*, 2006; Kéry & Schmidt, 2008). This prevents the Type II error of assuming absence of an organism when it is present, a serious matter when one is considering remedial management action for a species of conservation concern.

Gathering local knowledge advanced our understanding of where our cryptic data-poor species could be found, supporting similar findings about the value of local knowledge for research on data-poor species (Thornton & Scheer, 2012). Despite the growing use of local knowledge to inform marine species distributions, its use in a management context remains controversial because of potential data bias (Usher, 2000; Thornton & Scheer, 2012). Indeed, we also initially placed more value on systematic technical approaches than on gathering local knowledge (Turvey *et al.*, 2015), and thus hampered our first attempts to find sites for seahorses. It was only once we began to gather local knowledge that we realized the importance of accounting for detection probabilities and began to alter our methods. Selecting sites based on prior knowledge is entirely reasonable if one is trying to figure out where the animals can be found, in order to mobilize protection as soon as possible. Further research can work out presence/absence more broadly and even address relative abundance eventually.

Our research shows that when surveying data-poor species, factoring in detection probabilities changes the conclusions drawn whether in the ocean (Monk, 2013) or on land (Bailey *et al.*, 2007). Initially, our naive estimates of occupancy were quite low (36% Andaman coast; 38% Gulf coast) when not accounting for detection probability. However, using an occupancy model enabled us to estimate a broader occupancy rate (58% Andaman Coast; 73% Gulf Coast), suggesting that our rare cryptic species occurred in places where they had not been found. Such results have implications for all other seahorse studies relating to density, abundance, distribution and habitat preferences because they have failed to take into account detection probabilities. For example, Yasué, Nellas & Vincent (2012) would have done well to incorporate detection probabilities before declaring that marine protected areas had no significant effect on densities of a seahorse species (*Hippocampus comes*) in the Philippines. This is particularly true as we would expect detection probabilities to decrease in more complex environments, such as seagrass beds and coral reefs (Green *et al.*, 2013), and marine protected areas may often be more complex than heavily fished sites.

Our research highlighting the value of occupancy modeling with detection probabilities has application beyond seahorses to other data-poor marine species, especially those requiring conservation action at national and international levels. Data-

poor marine species may be regulated by national endangered species or fisheries legislations (humphead wrasse), regional commitments (e.g. marine mammals in the Pacific Islands with Convention on Migratory Species; CMS, 1979), or international agreements such as the Convention on International Trade in Endangered Species of Fauna and Flora (CITES, 1973) (e.g. mobulid rays, hammerhead sharks, oceanic white tip sharks, silky sharks and banggai cardinalfish) (CITES, 2016). Implementation of all such legislation relies on baseline information such as species' occurrences, distribution or population size to inform management action (CITES, 2013). In many cases, a lack of these types of data can lead to management inaction, mismanagement or further species declines (Johannes, 1998, 2000). It is both timely and relevant to encourage countries to account for imperfect detection in their sampling protocols, to avoid misestimated distributions and to assist countries to meet their national or international commitments (Issaris *et al.*, 2012; Katsanevakis *et al.*, 2012; Vincent *et al.*, 2013).

Similar to land-based conservation efforts, we found that when planning a detection study for data-poor marine fishes, a removal sampling design is less expensive than the standard design (Mackenzie & Royle, 2005; Bailey *et al.*, 2007). The standard survey design, where all sites are repeatedly sampled the same number of times, is a common choice for monitoring programs because a balanced survey design supports the use of relative abundance analyses and has more flexibility for data analysis (Andrew & Mapstone, 1987; MacKenzie *et al.*, 2006). However, when searching for cryptic or data-poor species, a removal sampling design – one where surveying stops once the species is found or until a maximum number of surveys has been reached – allows for the maximization of resources to visit as many sites as possible (Mackenzie & Royle, 2005). The selection of a standard survey design would be recommended when detection probability is survey specific (e.g. the probability of observing your species of interest varies each time you survey) because the removal survey design can only account for constant detection probabilities. While we have focused on the benefits of occupancy models with detection probabilities, estimating relative or true abundance of an individual species may be preferable for common species, or those species with previously established management goals (e.g. quotas, reducing population decline rates and recovery targets) (Maxwell & Jennings, 2005). When faced with a need to gather baseline data on the presence and absence of a data-scarce species expediently, where detection of individuals is constant for all surveys, marine managers should opt for a removal survey design (Mackenzie & Royle, 2005).

Techniques to assess data-poor species are rapidly evolving for both terrestrial and aquatic systems (MacKenzie *et al.*, 2006; Durso *et al.*, 2011). On land, occupancy and detection work has provided scientific support for decisions about habitat corridors (Zeller *et al.*, 2011), predator control (Robley *et al.*, 2014), invasive species remediation (Britton *et al.*, 2011) and conservation planning (Weller, 2008). Despite such advances on land, we have only begun to incorporate these ideas into ocean conservation and research

(Issaris *et al.*, 2012; Green *et al.*, 2013). Studies using presence/absence (occupancy) with detection probabilities are flexible and practical for science-based decision making (MacKenzie *et al.*, 2006) as compared to the data challenges of comparing densities and relative abundance, which require empirical rigor that is seldom available (Zar, 1999). More advanced methods for linking abundance and occupancy estimates, such as Royle & Nichols (2003), may be explored only once an initial understanding of the relationship between detection probabilities and abundance is reached. Species presence and absence across the land or seascape forms the basis for identifying sites for protection and creating a monitoring protocol, both critical steps during the initial phase of management. Finding a solution to the conservation crisis that is both pragmatic and expedient will involve increased communication among conservation fields to share beneficial lessons such as those to be learned from occupancy and detection.

Acknowledgments

This is a contribution from Project Seahorse. Funding for this work was provided by the Ocean Park Conservation Foundation of Hong Kong, the Riverbanks Zoo and Garden Conservation Fund, The Explorer's Club Exploration Fund, the SciFund Challenge, the John G. Shedd Aquarium, Guylian Chocolates and an anonymous donor. The authors thank the National Research Council of Thailand (permit number 0002/1306), the Thailand Department of Fisheries, Dr. Parichart Laksanawimol and Dr. Suchai Worachananart. They are grateful for the support from numerous dive operators and community groups who facilitated our search for seahorses, including Dive Tribe, Kapoe Dive Club, KiwiDivers.com, Laytrang Divers and New Heaven Dive School. They thank Dr. Sigal Balshine, Dr. Scott Hinch, Dr. Mary O'Connor and Tanvi Vaidyanathan for their manuscript edits. This research was conducted in accordance with UBC Animal Ethics protocols, permit no. A12-0288.

References

- Anderson, G.B., Freeman, M.C., Hagler, M.M. & Freeman, B.J. (2012). Occupancy modeling and estimation of the holiday darter species complex within the Etowah River system. *Trans. Am. Fish. Soc.* **141**, 34–45.
- Andrew, N.L. & Mapstone, B.D. (1987). Sampling and the description of spatial pattern in marine ecology. *Oceanogr. Mar. Biol. Annu. Rev.* **25**, 39–90.
- Bailey, L.L., Hines, J.E., Nichols, J.D. & MacKenzie, D.I. (2007). Sampling design trade-offs in occupancy studies with imperfect detection: examples and software. *Ecol. Appl.* **17**, 281–290.
- Bowles, E. & Trites, A.W. (2013). Faecal DNA amplification in Pacific walrus (*Odobenus rosmarus divergens*). *Polar Biol.* **36**, 755–759.
- Britton, J.R., Pegg, J. & Gozlan, R.E. (2011). Quantifying imperfect detection in an invasive pest fish and the implications for conservation management. *Biol. Conserv.* **144**, 2177–2181.
- CITES. (1973). *Convention on International Trade of Endangered Species of Fauna and Flora*. Washington, D.C.
- CITES. (2013). Compliance and Enforcement: Review of Significant Trade. 62nd Meet. Standing Comm. 1–29.
- CITES (2016). *Convention on International Trade of Endangered Species of Fauna and Flora*. Geneva: Switzerland.
- CMS. (1979). Convention on Migratory Species [WWF Document].
- Cohen, J. (1992). A power primer. *Psychol. Bull.* **112**, 155–159.
- Durso, A.M., Willson, J.D. & Winne, C.T. (2011). Needles in haystacks: estimating detection probability and occupancy of rare and cryptic snakes. *Biol. Conserv.* **144**, 1508–1515.
- Engler, R., Guisan, A. & Rechsteiner, L. (2004). An improved approach for predicting the distribution of rare and endangered species from occurrence and pseudo-absence data. *J. Appl. Ecol.* **41**, 263–274.
- Fairweather, P.G. (1991). Statistical power and design requirements for environmental monitoring. *Aust. J. Mar. Freshw. Res.* **42**, 555–567.
- Foran, D.R., Crooks, K.R. & Minta, S.C. (1997). Species identification from scat: an unambiguous genetic method. *Wildl. Soc. Bull.* **25**, 835–839.
- Foster, S. & Vincent, A. (2004). Life history and ecology of seahorses: implications for conservation and management. *J. Fish Biol.* **65**, 1–61.
- Gaston, K.J., Blackburn, T.M., Greenwood, J.J., Gregory, R.D., Quinn, R.M. & Lawton, J.H. (2000). Abundance-occupancy relationships. *J. Appl. Ecol.* **37**, 39–59.
- Goetze, J.S. & Fullwood, L.A.F. (2012). Fiji's largest marine reserve benefits reef sharks. *Coral Reefs* **32**, 121–125.
- Green, S.J., Tamburello, N., Miller, S.E., Akins, J.L. & Côté, I.M. (2013). Habitat complexity and fish size affect the detection of Indo-Pacific lionfish on invaded coral reefs. *Coral Reefs* **32**, 413–421.
- Guillera-Aroita, G. & Lahoz-Monfort, J.J. (2012). Designing studies to detect differences in species occupancy: power analysis under imperfect detection. *Methods Ecol. Evol.* **3**, 860–869.
- Hines, J.E. (2006). *PRESENCE ver.7.1*. Patuxent: U.S. Geological Survey Patuxent Wildlife Research Center.
- Hodgson, G. (2004). *Reef Check instruction manual: a guide to reef check coral reef monitoring*. Los Angeles: Reef Check, Institute of the Environment, University of California at Los Angeles.
- Hurme, E., Mönkkönen, M., Nikula, A., Nivala, V., Reunanen, P., Heikkinen, T. & Ukkola, M. (2005). Building and evaluating predictive occupancy models for the Siberian flying squirrel using forest planning data. *For. Ecol. Manage.* **216**, 241–256.
- Issaris, Y., Katsanevakis, S., Salomidi, M., Tsiamis, K., Katsiaras, N. & Verriopoulos, G. (2012). Occupancy estimation of marine species: dealing with imperfect detectability. *Mar. Ecol. Prog. Ser.* **453**, 95–106.

- Johannes, R.E. (1998). The case for data-less marine resource management: examples from tropical nearshore fin fisheries. *Trends Ecol. Evol.* **13**, 243–246.
- Johannes, R.E. (2000). Ignore fishers' knowledge and miss the boat. *Fish Fish.* **1**, 257–271.
- Katsanevakis, S., Weber, A., Pipitone, C., Leopold, M., Cronin, M., Scheidat, M., Doyle, T.K., Buhl-Mortensen, L., Buhl-Mortensen, P., D'Anna, G., de Boois, I., Dalpadado, P., Damalas, D., Fiorentino, F., Garofalo, G., Giacalone, V.M., Hawley, K.L., Issaris, Y., Jansen, J., Knight, C.M., Knittweis, L., Kröncke, I., Mirto, S., Muxika, I., Reiss, H., Skjoldal, H.R. & Vøge, S. (2012). Monitoring marine populations and communities: methods dealing with imperfect detectability. *Aquat. Biol.* **16**, 31–52.
- Kelly, M.J., Klenzendorf, S., Vaughan, M.R., Hutajulu, M.B. & Parakkasi, K. (2013). Threatened predator on the equator: multi-point abundance estimates of the tiger *Panthera tigris* in central Sumatra. *Oryx* **47**, 211–220.
- Kéry, M. & Schmidt, B.R. (2008). Imperfect detection and its consequences for monitoring for conservation. *Community Ecol.* **9**, 207–216.
- Lourie, S.A., Foster, S.J., Cooper, E.W.T. & Vincent, A.C.J. (2004). *A guide to the identification of seahorses*. Proj. Seahorse TRAFFIC North Am. Washington, D.C.: TRAFFIC USA.
- Lourie, S.A., Pollom, R.A. & Foster, S.J. (2016). A global revision of the Seahorses *Hippocampus Rafinesque 1810* (Actinopterygii: Syngnathiformes): taxonomy and biogeography with recommendations for further research. *Zootaxa* **414**, 1–66.
- Mackenzie, D.I. & Royle, J.A. (2005). Designing occupancy studies: general advice and allocating survey effort. *J. Appl. Ecol.* **42**, 1105–1114.
- Mackenzie, D.I., Nichols, J.D., Lachman, G.B., Droege, S., Royle, J.A. & Langtimm, C. (2002). Estimating site occupancy rates when detection probabilities are less than one. *Ecology* **83**, 2248–2255.
- Mackenzie, D.I., Nicholas, J.D., Sutton, N., Kawanishi, K. & Bailey, L.L. (2005). Improving inferences in population studies of rare species that are detected imperfectly. *Ecology* **86**, 1101–1113.
- MacKenzie, D.I., Nichols, J.D., Royle, J.A., Pollock, K.H., Bailey, L.L. & Hines, J.E. (2006). *Occupancy estimation and modeling: inferring patterns and dynamics of species occurrence*. San Francisco: Academic Press.
- Maxwell, D. & Jennings, S. (2005). Power of monitoring programmes to detect decline and recovery of rare and vulnerable fish. *J. Appl. Ichthyol.* **42**, 25–37.
- Mazerolle, M.J., Bailey, L.L., Kendall, W.L., Andrew, J., Converse, S.J. & Nichols, J.D. (2007). Making great leaps forward: accounting for detectability in herpetological field studies. *J. Herpetol.* **41**, 672–689.
- McCauley, D.J., Pinsky, M.L., Palumbi, S.R., Estes, J.A., Joyce, F.H. & Warner, R.R. (2015). Marine defaunation: animal loss in the global ocean. *Science* **347**, 1–7.
- McNie, E.C. (2007). Reconciling the supply of scientific information with user demands: an analysis of the problem and review of the literature. *Environ. Sci. Policy* **10**, 17–38.
- Mobley, K.B., Small, C.M. & Jones, A.G. (2011). The genetics and genomics of Syngnathidae: pipefishes, seahorses and seadragons. *J. Fish Biol.* **78**, 1624–1646.
- Monk, J. (2013). How long should we ignore imperfect detection of species in the marine environment when modelling their distribution? *Fish Fish.* **15**, 1–7.
- Perante, N.C., Pajaro, M.G., Meeuwig, J.J. & Vincent, A.C.J. (2002). Biology of a seahorse species, *Hippocampus comes* in the central Philippines. *J. Fish Biol.* **60**, 821–837.
- Petracca, L.S., Ramirez-Bravo, O.E. & Hernandez-Santin, L. (2013). Occupancy estimation of jaguar *Panthera onca* to assess the value of east-central Mexico as a jaguar corridor. *Oryx* **48**, 1–8.
- Robley, A., Gormley, A.M., Forsyth, D.M. & Triggs, B. (2014). Long-term and large-scale control of the introduced red fox increases native mammal occupancy in Australian forests. *Biol. Conserv.* **180**, 262–269.
- Rodtka, M.C., Judd, C.S., Aku, P.K.M. & Fitzsimmons, K.M. (2015). Estimating occupancy and detection probability of juvenile bull trout using backpack electrofishing gear in a west-central Alberta watershed. *Can. J. Fish Aquat. Sci.* **72**, 742–750.
- Royle, J.A. & Nichols, J.D. (2003). Estimating abundance from repeated presence-absence data or point counts. *Ecology* **84**, 777–790.
- Scotts, D.J. & Craig, S.A. (1988). Improved hair-sampling tube for the detection of rare mammals. *Wildl. Res.* **15**, 469–472.
- de Solla, S.R., Shirose, L.J., Fernie, K.J., Barrett, G.C., Brousseau, C.S. & Bishop, C.A. (2005). Effect of sampling effort and species detectability on volunteer based anuran monitoring programs. *Biol. Conserv.* **121**, 585–594.
- van Strien, A.J., Termaat, T., Groenendijk, D., Mensing, V. & Kéry, M. (2010). Site-occupancy models may offer new opportunities for dragonfly monitoring based on daily species lists. *Basic Appl. Ecol.* **11**, 495–503.
- Thornton, T.F. & Scheer, A.M. (2012). Collaborative engagement of local and traditional knowledge and science in marine environments: a review. *Ecol. Soc.* **17**, 8.
- Turvey, S.T., Trung, C.T., Quyet, V.D., Nhu, H.V., Thoai, D.V., Tuan, V.C.A., Hoa, D.T., Kacha, K., Sysomphone, T., Wallate, S., Hai, C.T.T., Thanh, N.V. & Wilkinson, N.M. (2015). Interview-based sighting histories can inform regional conservation prioritization for highly threatened cryptic species. *J. Appl. Ecol.* **52**, 422–433.
- Underwood, A.J. (1996). Detection, interpretation, prediction and management of environmental disturbances: some roles for experimental marine ecology. *J. Exp. Mar. Biol. Ecol.* **200**, 1–27.
- Underwood, A.J. & Chapman, M.G. (2003). Power, precaution, Type II error and sampling design in assessment of environmental impacts. *J. Exp. Mar. Biol. Ecol.* **296**, 49–70.

- Usher, P.J. (2000). Traditional ecological knowledge in environmental assessment and management. *Arctic* **53**, 183–193.
- Vincent, A.C.J., Sadovy de Mitcheson, Y.J., Fowler, S.L. & Lieberman, S. (2013). The role of CITES in the conservation of marine fishes subject to international trade. *Fish Fish.* **15**, 563–592.
- Weller, T.J. (2008). Using occupancy estimation to assess the effectiveness of a regional multiple-species conservation plan: bats in the Pacific Northwest. *Biol. Conserv.* **141**, 2279–2289.
- Williams, R. & Thomas, L. (2009). Cost-effective abundance estimation of rare animals: testing performance of small-boat surveys for killer whales in British Columbia. *Biol. Conserv.* **142**, 1542–1547.
- Yasué, M., Nellas, A. & Vincent, A.C.J. (2012). Seahorses helped drive creation of marine protected areas, so what did these protected areas do for the seahorses? *Environ. Conserv.* **39**, 183–193.
- Zar, J.H. (1999). *Biostatistical analysis, 4/e*. Upper Saddle River: Pearson Prentice Hall, Pearson Education, Inc.
- Zeller, K.A., Nijhawan, S., Salom-Pérez, R., Potosme, S.H. & Hines, J.E. (2011). Integrating occupancy modeling and

- interview data for corridor identification: a case study for jaguars in Nicaragua. *Biol. Conserv.* **144**, 892–901.
- Zipkin, E.F., Royle, J.A., Dawson, D.K. & Bates, S. (2010). Multi-species occurrence models to evaluate the effects of conservation and management actions. *Biol. Conserv.* **143**, 479–484.

Supporting information

Additional Supporting Information may be found in the online version of this article at the publisher's web-site:

Table S1. Total number of sites surveyed for seahorses by search strategy, Thai coast and search method.

Table S2. Rapid Assessment sites surveyed in 2013 on the Andaman and Gulf Coast of Thailand.

Figure S1. Detection probabilities (\pm 95% confidence intervals) of sites with low and high abundance of seahorses at sandy soft bottom research sites in Thailand.

Figure S2. Estimated number of survey replicates per site needed to obtain a 90% confidence that zeros represent true absence of seahorses (*Hippocampus* spp.) at the site.