Review

Comparing methods suitable for monitoring marine mammals in low visibility conditions during seismic surveys

Ursula K. Verfuss[a,*], Douglas Gillespie[b], Jonathan Gordon[c], Tiago A. Marques[d,1], Brianne Miller[e,2], Rachael Plunkett[a], James A. Theriault[f], Dominic J. Tollit[e], Daniel P. Zitterbart[g,3], Philippe Hubert[h], Len Thomas[d]

[a] SMRU Consulting, Europe, New Technology Centre, North Haugh, St Andrews, Fife KY16 9SR, UK
[b] Sea Mammal Research Unit, Scottish Oceans Institute, University of St Andrews, Fife KY16 8LB, UK
[c] Marine Ecological Research, 7 Beechwood Terrace West, Newport-On-Tay, Fife DD6 8JH, UK
[d] Centre for Research into Ecological and Environmental Modelling, The Observatory, University of St Andrews, St Andrews, Fife KY16 9LZ, UK
[e] SMRU Consulting, North America, 1529 W 6th Ave, Vancouver, BC V6J 1R1, Canada
[f] Ocean Environmental Consulting, 9 Ravine Park Cres, Halifax B3M 4S6, NS, Canada
[g] Alfred Wegener Institute Helmholtz Centre for Polar and Marine Research, Am Handelshafen 12, 27570 Bremerhaven, Germany
[h] Prove Systems Ltd, Unit 1 Mill court, Mill lane, Tayport, Fife DD6 9EL, UK

ARTICLE INFO

ABSTRACT

Loud sound emitted during offshore industrial activities can impact marine mammals. Regulations typically prescribe marine mammal monitoring before and/or during these activities to implement mitigation measures that minimise potential acoustic impacts. Using seismic surveys under low visibility conditions as a case study, we review which monitoring methods are suitable and compare their relative strengths and weaknesses. Passive acoustic monitoring has been implemented as either a complementary or alternative method to visual monitoring in low visibility conditions. Other methods such as RADAR, active sonar and thermal infrared have also been tested, but are rarely recommended by regulatory bodies. The efficiency of the monitoring method(s) will depend on the animal behaviour and environmental conditions, however, using a combination of complementary systems generally improves the overall detection performance. We recommend that the performance of monitoring systems, over a range of conditions, is explored in a modelling framework for a variety of species.

1. Introduction

Anthropogenic sound from shipping, pile driving, the use of explosives, high intensity active sonar operations, seismic surveying and many other activities can mask marine mammal communication sounds, cause changes in the behaviour of these animals, exclude them from important habitats and, in extreme cases, induce auditory injury or death (Erbe, 2002; Gordon et al., 2003; Ketten, 1995; Pirotta et al., 2014; Southall et al., 2007). To reduce the risk of potential impacts, those carrying out industrial projects and naval operations offshore are often required to monitor their operational area for the presence of marine mammals, so that mitigation actions can be taken (e.g. ACCOBAMS Scientific Committee, 2004; IAGC, 2011; JNCC, 2010a, 2010b, 2010c, 2017; Martin et al., 2014; Nowacek and Southall, 2016;

Weir and Dolman, 2007). Traditionally, this kind of monitoring involves trained marine mammal observers (MMOs) scanning the sea surface for marine mammals. Visual methods are, however, restricted to daylight hours and relatively good weather conditions. Visual detection is often subjective and happens in an instant, so it is difficult or impossible to confirm or review a detection at a later stage. The effectiveness with which an MMO can visually detect an animal is reduced by weather conditions such as fog, rain, high sea state, sun glare or the lack of light (e.g. Clarke, 1982; Harwood and Joynt, 2009; Palka, 1996; Parente and de Araujo, 2011). Visual detection at night without the aid of additional equipment is impossible. Animal behaviour, such as diving and an undemonstrative presence at the sea surface, can also reduce detection probability.

In recent years, there has been an increased interest in using other

---

technologies to overcome the most obvious limitations of visual monitoring. In particular, the use of passive acoustic monitoring (PAM) during monitoring for mitigation purposes has increased, with some national guidelines encouraging its use and industry efforts focusing on improving existing PAM capabilities. Additional promising approaches that could potentially enhance the detection of marine mammals in low visibility conditions include active acoustic monitoring (AAM), thermal imaging (thermal IR) and radio detection and ranging (RADAR).

Passive acoustic monitoring detects an animal's vocalisations using hydrophones (underwater microphones). Active acoustic monitoring is a method where sound pulses are emitted into the water and acoustic reflections from an animal are received by hydrophones. Thermal IR uses an electro-optical imaging sensor to detect temperature differences between the body or the exhalation of the warm blooded marine mammal and that of the surrounding environment. In RADAR, radio or micro-waves are emitted into the air and echoes from the animal's body, its exhalation, or from disturbance on the sea surface are picked up by an array of receivers. Systems using these modalities usually incorporate software detection and/or visualisation tools, usually supervised by a trained human operator who makes the final judgement on animal detection. All of the methods outlined above can be used for marine mammal detection and have the potential to complement traditional visual monitoring methods. Each method has its strengths and weaknesses, and may be more or less suitable for particular scenarios depending on the species monitored and the environmental conditions in which the monitoring takes place. For example, as with visual monitoring, IR and RADAR techniques have the weakness that they cannot detect submerged animals and may have reduced effectiveness in high sea states, whereas PAM has the weakness that it cannot detect silent animals and may miss animals whose vocalisations are highly directional.

The purpose of this review is to reveal each method's strengths and weaknesses from the perspective of monitoring for mitigation, and to list examples of systems which are currently available. We highlight the factors that need to be considered in order to make well informed decisions on the monitoring method, or combination of methods, to apply and the specific systems to use. We discuss the conditions that are favourable or unfavourable for each method and the strengths and weaknesses of each method in terms of both extrinsic and intrinsic factors:

- **Extrinsic factors** are those factors that cannot be influenced by the monitoring team (e.g. sea state, light conditions, animal behaviour, animal size, etc.).
- **Intrinsic factors** are those properties that can realistically be influenced by the monitoring team (e.g. quality and sophistication of the instruments and associated software, method of deployment, etc.).

In this review we focus on evaluating the monitoring methods when applied to marine mammal monitoring for mitigation purposes on a seismic survey vessel, assuming that monitoring systems will be installed on the main survey vessel or on non-specialised ancillary vessels, which is standard practice. This kind of monitoring is generally conducted in the course of seismic surveys. During seismic surveys, acoustic pulses are generated by the seismic sound source and transmitted through the water column into the sea bed (OGP and IAGC, 2011). Some of the transmitted sound energy is reflected by rock strata and received on hydrophones distributed in very large arrays of sensors in long survey streamers, towed by and behind the survey vessel. Acoustic data from thousands of sensors are processed on board and can be viewed as maps showing the structure and nature of the layers in the surveyed area.

Table 1 provides a summary of systems that are available and suitable for such monitoring based on a questionnaire survey of developers, suppliers and users of such monitoring techniques carried out in

2015, supplemented with publicly available information, the practical experience of the authors and contributions from an advisory panel (see Verfuss et al., 2016 for further information).

To understand which of the monitoring methods and systems may be useful for low visibility monitoring conditions, we first evaluate the requirements for effective monitoring for mitigation during seismic surveys and discuss how monitoring effectiveness can be quantified. We then present the results of our analysis, revealing which intrinsic factors (technical or operational parameters) should be addressed to achieve a high detection performance across a wide range of species and how extrinsic factors (animal behaviour and environmental conditions) influence monitoring effectiveness.

An evaluation of the specific systems reviewed highlights their detection, classification and localisation capabilities, and provides ball park figures on the system costs, their commercial availability and installation requirements.

The review concludes by making recommendations for research to assess and improve the effectiveness of low visibility monitoring technologies.

2. Method description and system overview

Each of the methods described is able to detect and classify cues from marine mammals. This section summarises the principle of operation for each method and synthesises the systems listed in Table 1. For definitions of the technical terms and abbreviations used in this review please see Table 2.

2.1. PAM

Marine mammal monitoring with PAM depends on the animal emitting sounds that can serve as cues for detection. Marine mammals produce sound to communicate with conspecifics (Janik and Sayigh, 2013; Madsen et al., 2002; Schulz et al., 2008), for orientation (Payne and Webb, 1971; Verfuss et al., 2005), to locate and capture prey (Au et al., 2004; Madsen et al., 2005; Verfuss et al., 2009), mate selection and social interactions (Janik, 2009; Quick and Janik, 2012; Smith et al., 2008). Marine mammal vocalisations are often characteristic and loud, providing a suite of acoustic cues that can be used to detect and localise marine mammals.

The PAM systems used for monitoring for mitigation purposes during seismic surveys, and listed in Table 1, are of two distinct types: "ancillary" and "integrated" systems. Ancillary systems involve deploying one, or more, dedicated marine mammal PAM hydrophone arrays (streamers) from either the seismic survey vessel or, more rarely, from some other vessel already operating on site (e.g. a guard vessel). Hydrophones are monitored aurally by human observers and/or using acoustic analysis software, such as the open source PAMGuard software (Gillespie et al., 2008; www.pamguard.org), to detect, classify and localise marine mammal vocalisations in real-time. PAM systems of this type have been used for monitoring during seismic operation since the late-1990s. Ancillary towed hydrophone systems are provided by several companies (Table 1), but share several common features. They typically consist of several hydrophones in a terminal array section towed on between 100 and 400 m of strengthened cable. Hydrophones are typically grouped in matched pairs covering different frequency ranges. A deck lead carries signals from the hydrophone termination on the aft deck of the tow vessel to the instrument room where additional hardware providing signal conditioning (filtering and amplification) and digitisation are housed, along with analysis computers (typically high end laptop PCs). One company (Seiche Ltd.) also offers short PAM hydrophone arrays deployed via the seismic source array. This configuration is intended to avoid some of the entanglement risks associated with streaming cables from the aft deck. The most complicated, and arguably the most capable, hydrophone streamer system for which we were provided information on was the "Delphinus" array being

Table 1

AAM, PAM, RADAR and thermal IR monitoring systems discussed in this publication and the corresponding company names of those developers and suppliers that took part in the 2015 questionnaire survey of Verfuss et al. (2016). The type of system as classified by the authors (see text) in brackets and the website of the company are also provided.

Method	Systems (kind)	Company name	Website
AAM	AN/SSQ-963 (< 20 kHz)	Ultra-Electronics Sonar Systems	http://www.ultra-sonar.com
	CMAS-36/39 OMNI Sonar® (20–50 kHz)	Nautel C-Tech Limited	http://nautelc-tech.com
	Echoscope (> 150 kHz)	Coda Octopus Products Ltd	http://www.codaoctopus.com
	Gemini 720 (> 150 kHz)	Tritech International Limited	http://www.tritech.co.uk
	HFM3 (20–50 kHz), SDSN (50–150 kHz)	Scientific Solutions, Inc.	http://sci-sol.com
	SeaBat (> 150 kHz)	Teledyne RESON UK Ltd	http://www.teledyne-reson.com
	Sentinel (50–150 kHz)	Sonardyne	www.sonardyne.com
	Simrad SX90, SU90 (20–50 kHz), Simrad SH90 (50–150 kHz)	Kongsberg Maritime Subsea	www.km.kongsberg.com
PAM	(Ancillary system)	Abakai International LLC	http://www.manta.com/c/mrs5rjt/abakai-international-llc
	(Ancillary system)	Columbia University	http://www.ldeo.columbia.edu/~shimah
	(Ancillary system)	SANYA Institute of Deep-Sea Science and Engineering, Chinese Academy of Sciences	http://www.sidsse.cas.cn/jgsz/yjxt/shkx/brsw
	(Ancillary system)	Vanishing Point Marine	http://vpmarine.co.uk
	Custom made (ancillary system)	Bio-Waves Inc.	http://biowaves.net
	Custom made (ancillary system)	Passive Acoustic Monitoring Online Services (PAMOS)	http://www.pamos.ca
	Custom made research system (ancillary system)	GeoSpectrum Technologies Inc.	www.geospectrum.ca
	Delphis (ancillary system)	TNO, Acoustics & Sonar Research Group	https://www.tno.nl/en/collaboration/expertise/technical-sciences/acoustics-and-sonar
	FarONE (ancillary system)	NAUTA Ricerca e Consulenza Scientifica	www.nauta-rcs.it
	Night Hawk III (ancillary system)	MSeis Limited	http://www.mseis.com
	Ocean noise console (ancillary system)	Quiet-Oceans	http://www.quiet-oceans.com
	ONCET (ancillary system)	University of Toulon, DYNI LSIS Lab	http://www.lsis.org
	QuietSea (integrated system)	SERCEL	http://www.sercel.com/products/Pages/QuietSea.aspx
	Seiche PAM (ancillary system)	Seiche Measurements Ltd	http://www.seiche.com
	WhaleWatcher (integrated system)	WesternGeco	http://www.slb.com/services/seismic/seismic_acquisition/contact.aspx
RADAR	FMCW surface detection RADAR, Magnetron pulsed surface detection RADAR	Radar Technology AS	http://www.radar-technology.com
	Kelvin Hughes RADAR with WaMoS II digitiser	National Oceanography Centre	noc.ac.uk
	SHN X9 (polarimetric RADAR)	Sea-Hawk Navigation AS	http://www.sea-hawk.com
Thermal IR	(Uncooled planar)	Ocean Life Survey	www.oceanlifesurvey.com
	(Uncooled planar)	Polaris Sensor Technologies, Inc.	www.PolarisSensor.com
	(Uncooled planar)	Toyon Research Corporation	www.toyon.com
	AIMMMS (rotating line scanner)	RDE (Rheinmetall)	http://www.rheinmetall-defence.com
	Gobi (cooled planar)	Xenics	http://www.xenics.com/en
	Hyper-cam (cooled planar)	Telops	http://telops.com
	Night navigator (uncooled, cooled planar)	Current Scientific Corporation	www.currentcorp.com
	RADES (uncooled planar)	Seiche Measurements Ltd	http://www.seiche.com
	SECurus (cooled planar)	Aptomar AS	www.aptomar.com

developed by TNO in the Netherlands (Sheldon-Robert et al., 2008; von Benda-Beckmann et al., 2010). The Delphinus system is currently only being used for research purposes and has not been deployed from a seismic survey vessel. Its main application thus far has been for tracking cetaceans (especially deep diving species such as beaked whales) as part of tagging studies.

Integrated systems, to a large extent, utilise the hydrophone elements in the streamers already towed by seismic vessel. These integrated systems are a relatively recent innovation and currently only two array manufacturers offer this type of system: the WhaleWatcher (WesternGeco) and QuietSea (Sercel). In addition, at least one academic group, at Colombia University, is developing a similar approach utilising the sensors in seismic streamer arrays used for geophysical research to detect baleen whale calls (Abadi et al., 2015).

2.2. AAM

The marine mammals' body and, to a lesser extent, the disturbance (wake) that animals generate in the water can reflect the sound pulses emitted by an AAM system. These reflections are the cues that enable active sonar systems to detect these animals. The degree to which environmental parameters affect the performance of sonar (as well as their detection ranges) is mostly frequency dependent (see Section 4.2) and systems may, therefore, usefully be classified by their transmission

frequency. Generally, the size of the system and the power requirements decrease with increasing frequency. Of the wide range of systems available, those reviewed can be grouped into transmission frequency classes of 20–50 kHz, 50–150 kHz, and > 150 kHz. Table 1 lists the reviewed systems and their frequency range.

2.3. Thermal IR

Thermal IR detection is dependent on the animal surfacing and revealing parts of its warm body or its exhalation (blow). The apparent temperature difference between the marine mammal body or blow and the sea surface provides a thermal contrast, which can be detected using a thermal imaging sensor.

Thermal IR systems are designed to be mounted on a vessel or a plane. In this review, planar and rotating line scanners are considered as they provide the optimal solution for monitoring for the mitigation purposes considered in this study. Rotating line scanners can provide a full 360° view of the ocean by spinning a sensor with a frequency of several revolutions per second, while planar scanners monitor the ocean with a Focal Plane Array (FPA), very much like a digital camera. If a field of view larger than the cameras lens can provide is desired, either more than one camera must be used, or the camera must scan using a pan-tilt unit. The Automated Infrared-based Marine Mammal Mitigation System (AIMMMS) developed by Rheinmetall Defence is, to

Table 2
Terms and corresponding definitions used in this review.

Term	Description
AAM	Active acoustic monitoring
Ambient noise	That part of the total noise background observed with a non-directional hydrophone which is not due to the hydrophone and its manner of mounting or to some identifiable localised source of noise (Urick, 1984)
Apparent temperature	The uncompensated reading from an IR camera without correction of emissivity
Background noise	All acoustic sound detected in the environment at a time, including all sound in the ocean, and excluding the signal of interest, system noise, electrical noise and self-noise
Bit depth	The precision with which a digitiser can measure voltage changes
CCOS	Concurrent ocean coverage
Cue	Signal elicited by the presence of an animal of the target species that potentially triggers detection
Electrical noise	Any electrical interference including that resulting from sources such as ground loops and radio interference which create a noise in electrical systems
Extrinsic factors	Factors that cannot be influenced by the team carrying out the monitoring (e.g., sea state, visibility, animal behaviour, animal size, etc.)
Flow noise	Component of self-noise that results from turbulence as water flows around a hydrophone
HF	High-frequency, ranging up to 40 kHz
Intrinsic factors	Factors that realistically can be influenced by those responsible for monitoring (e.g., quality and sophistication of the instruments and associated software, form of deployment)
In-time detection	Detection of an animal sufficiently early to allow mitigation measures to be implemented before the animal enters the mitigation zone
IR	Infrared
JNCC	Joint nature conservation committee
LWIR	Long-wavelength Infrared, with wavelength ranging from 8 to 12 µm
MF	Mid-frequency, ranging up to 12 kHz
Mitigation zone	Area, within which national regulations or guidelines prescribe that the detection of an animal of a target species should trigger mitigation measures
Monitoring zone	Area that needs to be monitored during monitoring for mitigation purposes
MWIR	Mid-wavelength infrared, with wavelength ranging from 3 to 5 µm
Noise	Any energy (not necessarily of acoustic origin) which is not signal and can potentially interfere with the detection and localisation of signals (i.e. the cues)
PAM	Passive acoustic monitoring
RADAR	Radio detection and ranging
Self-noise	Energy originating from the recording system itself
Signal	Synonym to cue
Signal-to-noise ratio	Ratio of the energy of a signal to non-signal energy over a period of time
Source level	Sound pressure level at one meter distance from the sound source
System noise	The electrical noise which is an inherent part of the properly working system and that which may result from a shortcoming or fault in the system. This is a component of self-noise
Target species	Species for which the monitoring needs to be conducted
Total noise	The sum of all kinds of noise as defined below, i.e. all the energy registered on a system, excluding that of the signal
Transmission loss	Reduction in the amplitude of a signal or cue passing between two points (here: animal to receiver for passive systems, and sender to reflector to receiver for active systems) of a transmission path
UHF	Ultra-high frequency, ranging up to 150 kHz

our knowledge, the only one currently used for marine mammal detection. It produces an image by rotating the sensor perpendicular to the ocean's surface, which is maintained by a gimballed stabilisation. This result in a larger weight and size footprint compared to directional cameras. Planar sensor systems are offered by several companies (Table 1). These systems can be further divided into those that employ cooled sensors and those that use uncooled sensors, or both. Cooled systems generally have a better signal-to-noise ratio but come at a higher cost. Several system suppliers (Ocean Life Survey, Seiche Measurement Limited and Toyon Research Corporation) use cameras that are produced by the company FLIR and state that the cameras, and therefore sensors, can be exchanged.

2.4. RADAR

RADAR (radio detection and ranging) is a system that uses electromagnetic waves travelling through air to determine the range and direction of objects that reflect them. A system consists of a transmitter that emits either microwaves or radio waves that are reflected by the target and detected by a receiver, which often uses the same antenna as the transmitter. Typically, a RADAR antenna scans the area of interest continuously and at a moderate rate. Both the transmitter and receiver arrays are highly directional and the bearing to the target is given by the orientation of the antenna. RADAR can detect marine mammals at the surface from a reflection of the RADAR pulse from the exposed body of the animal, an exhalation, or from disturbance on the sea surface (e.g. awake). Thus RADAR is most effective for detecting larger animals in low sea state conditions. Specific information on four differing

RADAR systems from three companies was retrieved, two of which sell or lease high-end RADAR systems, mainly for ice detection use on vessels working in Arctic and sub-Arctic conditions. RADAR Technology AS provided information on their Frequency-Modulated Continuous-Wave (FMCW) surface detection RADAR, and a less expensive system (Magnetron pulsed surface detection RADAR). The FMCW radiates continuous transmission power and can change its operating frequency during the measurement so that the transmission signal is frequency modulated. This maximises the total power on a target because the transmitter is broadcasting continuously and increases the reliability by providing distance measurement along with speed measurement, which is essential when there is more than one source of reflection arriving at the radar antenna. Sea-Hawk Navigation AS provided information on their advanced polarimetric RADAR systems, specifically the SHN X9, which they considered most suitable for marine mammal detection. The National Oceanography Centre (NOC) provided details of their non-commercial system that used a Kelvin Hughes RADAR and WaMoS II digitiser and proprietary software (called GANNET).

3. Monitoring requirements and effectiveness

In the following section we outline requirements for effective monitoring for mitigation purposes and how this can be evaluated.

3.1. Monitoring requirements

Monitoring is often required as part of mitigation procedures during seismic surveys, particularly in areas with vulnerable marine mammal

populations. Monitoring requirements are usually outlined in national regulations or guidelines. Weir and Dolman (2007) and Martin et al. (2014) provide comprehensive descriptions and comparisons of guidelines from a range of countries. One common feature of these is that a period of monitoring for the presence of individuals of a target species is required before source arrays are activated. This is intended to ensure that no animal is present near the sound source (and thus potentially impacted) when the sound source is first activated. Many guidelines additionally require monitoring when the source is active. Regulations and/or guidelines outline mitigation measures (e.g. a delay in the seismic sound source activation) to be taken if an animal of a target species is detected within a pre-defined area around the sound source. We term this area, within which the national regulations or guidelines specify that the detection of an animal of a target species should trigger mitigation measures, the "mitigation zone".

Some guidelines or regulations only require this mitigation zone to be monitored, while others (e.g. Australia, DEWHA, 2008) require the monitoring of a wider area around the mitigation zone and the tracking of animals before they enter the mitigation zone. We term the wider area that should be monitored the "monitoring zone". In some guidelines or regulations, the monitoring zone matches the mitigation zone (e.g. JNCC, 2010a), while in others this zone can be substantially larger (e.g. DEWHA, 2008). The maximum radius of the monitoring zone can vary from 500 m (e.g. UK JNCC, 2010a) to > 3 km (Australia, DEWHA, 2008).

Marine mammals are difficult to detect, in part because most of the cues used for detection are only available intermittently. Thus, animals cannot be tracked continuously as they enter the mitigation zone, and the probability of detecting an animal before it enters the mitigation zone needs to be considered. Mitigation actions require a certain lead time for them to be implemented. The decision to apply a mitigation measure must therefore be taken with sufficient lead time for it to be initiated before an animal enters the mitigation zone. The timely detection of an animal with sufficient time in hand to allow for mitigation measures to be implemented before the animal enters the mitigation zone is defined as "in-time detection" (Verfuss et al., 2016; Zitterbart et al., 2013) and will require monitoring of an area larger than the mitigation zone.

Determining the optimal size and shape of the monitoring zone is a complex issue and has rarely been considered. Many factors including the speed of the sound source, the patterns of cue production for the animals of concern and the capabilities of the available detection systems in the prevailing conditions are all relevant. For static sound sources, the ideal monitoring zone would be circular. This also applies for the inactive seismic sound source during the monitoring period before it will be activated. In this case, the ideal monitoring zone would be a circle centred at the position at which activation of the source array is planned (Fig. 1A). For moving active sound sources, however, such as an active seismic source array, which typically moves with a speed of 4–5 knots (Fig. 1B), the effective monitoring zone and the distribution of searching effort should be biased forward. This is because animals detected ahead of the source array are more likely to enter the mitigation zone which is moving towards them and will on average do so more quickly than animals in other directions. Ideally, an animal's movements should be tracked when it is detected in the monitoring zone in order to assess the likelihood of it entering the mitigation zone and, if there is a high probability of the animal entering the mitigation zone, appropriate mitigation measures should be taken even if the animal has not been detected in the mitigation zone.

3.2. Detection probability

The probability of detecting marine mammals and how to measure this probability has been the focus of many studies within the framework of population abundance and density surveys (e.g. Borchers et al., 2002; Buckland et al., 2001, 2004; Buckland et al., 2015; Seber, 1986;

Thomas et al., 2010). This work provides a useful framework for assessing detection probability during monitoring for mitigation purposes.

There are, however, some key issues that need to be considered for monitoring for mitigation purposes, which are not relevant for abundance estimates. One key difference is that during population surveys the only requirement is that detection probability and the size of the area being monitored are known (or can be measured). For mitigation, the probability not only need to be known, but should be close to 1 across the entire monitoring zone if all animals potentially at risk are to be detected. Another key difference is that efficient monitoring for mitigation purposes is required in the environmental conditions in which industrial activities might take place, while abundance surveys are generally only conducted under favourable conditions. Furthermore, detections need to be made in near real time for mitigation purposes, which is not required for abundance surveys.
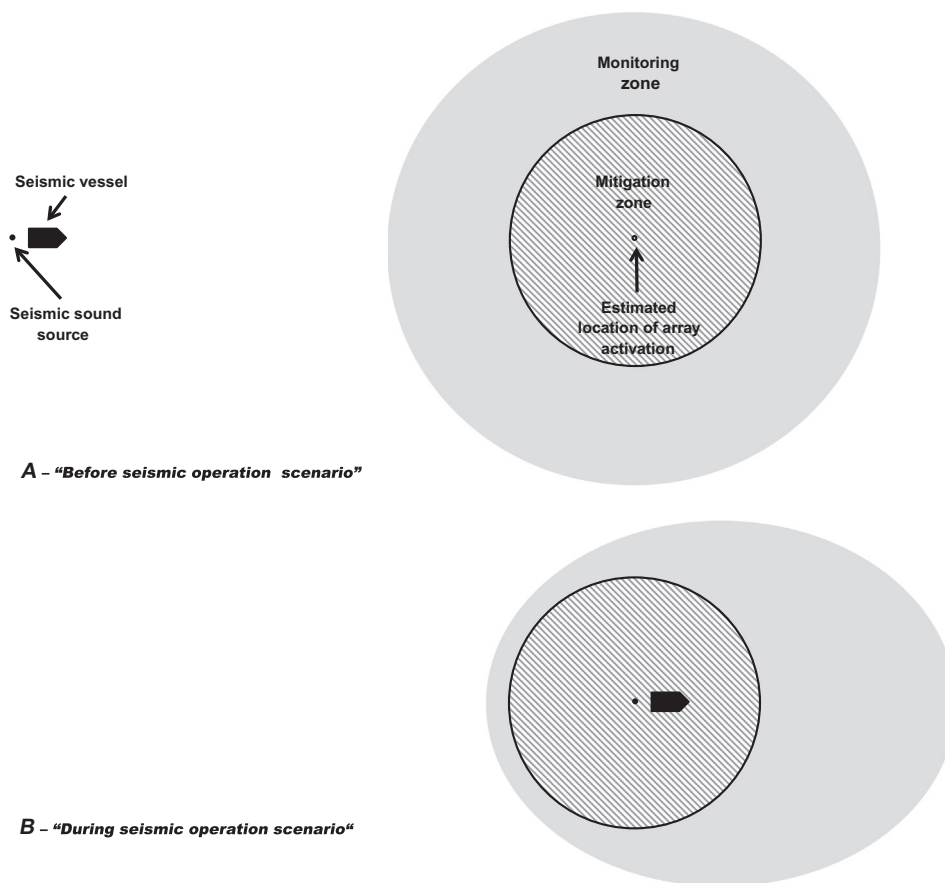
An animal that is present within an area being surveyed by a given monitoring technique might be missed as a result of two different processes termed **availability bias** and **perception bias** by Marsh and Sinclair (1989). **Availability bias** occurs when an animal is missed because it was not available to be detected. For example, an animal might be present but not producing detectable cues. Obvious examples include animals underwater and thus not available to be seen by visual observers, or silent animals that cannot be detected with PAM. **Perception bias** occurs when animals are available for detection (e.g. the animal is at the surface for visual observers or vocal for PAM), but the detection system fails to detect the available cue. In the case of a visual observation this includes instances when an animal is undetected because of poor environmental conditions, or the observer is looking in a different direction or fails to notice the animal for any other reason. Observer fatigue is a well-known source of perception bias. Maintaining visual vigilance is mentally and physically taxing and observer performance diminishes if observers are not sufficiently rested. In addition, weather and other environmental conditions affect detection probability; for example, visual detection becomes increasingly difficult as sea state increases (Marsh and Sinclair, 1989; Palka, 1996).

Monitoring for mitigation purposes can have four possible outcomes, with different consequences:

1.) True positive: an animal of a target species is detected and mitigation measures are taken in time before it enters the mitigation zone.
2.) True negative: no detections are made of an animal of the target species that could enter the mitigation zone, no animal enters the mitigation zone and no mitigation measures are taken.
3.) False positive: mitigation measures are taken based on a detection but no animal of a target species enters the mitigation zone.
4.) False negative: an animal of a target species enters the mitigation zone undetected and no mitigation measures are implemented.

True positives and true negatives are considered mitigation successes. False positives and false negatives are considered as mitigation failures: they will either have potentially negative consequences for the animal or impact the efficiency of the seismic surveys.

False positives have two different origins. The first is when a false alarm is triggered by "noise" that was mistakenly identified (by software and/or a human observer) as a cue of an animal from the target species; detection was then reported and mitigation measures were taken, even though no animal of the target species was present. Noise in this context is defined as any energy which is not signal and can potentially interfere with the detection and localisation of signals (i.e. the cues). It may be the case that either (a) it can be determined that this error occurred (e.g. by post-inspection it was clarified that it was a false alarm), or (b) that it cannot be determined that this error occurred (e.g. no post-inspection is done, or post-inspection does not reveal a false alarm). The second origin is that an animal of the target species was

Fig. 1. The optimal distribution of monitoring effort around a mitigation zone before and during seismic operation. Sizes are not to scale. Adapted from Verfuss et al. (2016).

A – The "Before seismic operation scenario" illustrates the shape and position of the mitigation (hatched area) and monitoring zone (grey area) to be monitored for mitigation purposes in the period before the sound source is activated, with the intended location of array activation given as a small dot, and the survey vessel moving towards the activation location. The monitoring zone is circular around the static mitigation zone. This would also be the pattern for a stationary sound source.

B – The "During seismic operation scenario" illustrates the extent of the mitigation and monitoring zones while the active seismic sound source is towed behind the survey vessel. The grey area illustrates the forward biased monitoring zone and effort as a result of the moving sound source. Animals detected ahead of the source array are more likely to enter the mitigation zone than those in other directions.

correctly detected outside the mitigation zone, a decision was made that it was sufficiently likely to enter the mitigation zone for mitigation measures to be implemented, but the animal did not subsequently enter the mitigation zone as expected. This could be because of errors in an assessment of the animal's range from the sound source, or because it did not move as expected between the time of its last detection and the commencement of mitigation action. As before, it may be the case that either (a) it can be identified that this error occurred (e.g. the animal was afterwards seen in a location that made it impossible for it to have been in the mitigation zone), or (b) it cannot be determined that a false positive occurred (e.g. it is unknown whether or not the animal entered the mitigation zone as the animal was not re-sighted or the position of the re-sighting made it possible for it to have been in the mitigation zone). These types of errors have an inherent time, resource and efficiency cost for a given activity with no associated benefit.

False negatives are instances in which an animal of a target species enters the mitigation zone without being detected. As a result, the animal may be exposed to an unacceptable level of sound. False negatives are regarded as monitoring failures to the detriment of the target species. False negatives are of two different types:

(a) An animal of the target species enters the mitigation zone unnoticed and is subsequently detected (detected false negatives). Mitigation measures might then be taken but it is likely the animal may have already been exposed to the sound.

(b) An animal of the target species enters the mitigation zone unnoticed and remains unnoticed, therefore the animal may be exposed to sound (undetected false negatives).

In terms of the actual impact on the animal, undetected negatives might be more detrimental than detected ones because the acoustic exposure may be longer. However, by definition, only detected false negatives are actually observed in the field. The number of undetected errors can only be calculated indirectly.

There are several issues associated with the quantification of false negatives. There is the risk that mitigation success will be overestimated if undetected false negatives are not considered. Further, in more difficult survey conditions, when the detection systems are less effective, the number of undetected false negatives will increase while the number of detected false negatives will decrease, with a corresponding decrease in the number of reported false negatives. A naïve interpretation of this would suggest that mitigation efficiency has increased when in fact the opposite is the case. It is clear that a simple count of observed false negatives is not a sensible way of assessing the performance of a monitoring technique, since it ignores undetected false negatives.

A good detection system will have a high rate of true positives and a low rate of false positives. In most detection systems, there is a direct trade-off between these two quantities, with systems tuned to have a high rate of detections also having a higher rate of false positives and vice-versa. This review will make recommendations for how to quantify both types of false negatives.

4. Factors influencing the detection probability of marine mammals

In this section we present an assessment of intrinsic and extrinsic factors for each of the low visibility monitoring methods considered. Identifying and understanding the factors that influence the detection probability of marine mammals will assist in choosing the most appropriate method(s) and sourcing appropriate systems for a specific monitoring task.

4.1. Intrinsic factors

An overview of intrinsic factors and their influence on the

Table 3

Schematic and simplified listings of the most important intrinsic factors affecting monitoring with PAM, AAM, vessel mounted thermal IR and RADAR systems. The positive or negative influences of the intrinsic factors lead to strengths (+) and weaknesses (−) of the methodology. Those rows pertaining to a factor that has a strong influence on the detection performance shown in bold.

Method	Internal factor	Property	Strength (+) or weakness (−)
PAM	**Array design**	**Multiple hydrophone wide aperture array (4 or more)**	**Localisation possible (+)**
		Small array (2 to 4)	**Localisation only by target motion (−)**
	Array gain	High	Increases detection range (+)
			narrow receiving beam (animals may be missed) (−)
		Low	Wide receiving beam (+)
			decreases detection range (−)
	Bit depth	High	Increased dynamic range(+)
		Low	Decreased dynamic range (−)
	Deployment depth	**Deeper**	**Generally better signal to noise (+)**
		Shallower	**Worse signal to noise ratio and greater risk of entanglement (−)**
	Detector configuration	**Appropriate**	**Increases signal to noise ratio (+)**
			maintains dynamic range (+)
		Not appropriate	**Decreases signal to noise ratio**
	Filter setting	Optimised	Good signal to noise ratio (+)
		Not optimised	Poor signal to noise ratio (−)
	Flow noise/self-noise	Low	Good signal to noise ratio (+)
		High	Poor signal to noise ratio
	Frequency range	**Covering vocalisation**	**Detection possible (+)**
		Outside vocalisation range	**Detection impossible (−)**
	Sensitivity	Appropriate	Detection possible (+)
		Not appropriate	Detection less possible (−)
	System noise	Low	Good signal to noise ratio (+)
		High	Poor signal to noise ratio (−)
AAM	Motion compensation	Good	Improved performance in high sea state (+)
		None	Limits detection probability in high sea state (−)
	Pulse bandwidth	Broad	High resolution (+)
		Narrow	Low resolution (−)
	Pulse duration	**Long**	**Higher echo strength (+)**
			increases noise (reverberation) (−)
		Short	**Weakens echo strength and detectability (−)**
			reduces noise (reverberation) (+)
	Pulse frequency range	**Low**	**Less absorption of signal (+)**
			makes small animals difficult to detect (−)
		High	**Higher absorption of signal (−)**
			enables finer resolution of small animals (+)
		Inaudible for animals	**Lower likelihood of impact on animals (+)**
		Audible for animals	**Higher likelihood of impact on animals (−)**
	Sonar blind spot	Small	Increased detection probability (+)
		Large	Decreased detection probability (−)
	Source & receiver design	**Favourable**	**Increased detection probability (+)**
		Unfavourable	**Decreased detection probability (−)**
	Source level	**High**	**Increased detection range (+)**
			higher likelihood of impact on animals (−)
		Low	**Decreased detection range (−)**
			lower likelihood of impact on animals (+)
	System noise	Low	Good signal to noise ratio (+)
		High	Bad signal to noise ratio (−)
Thermal IR	Camera band	Wide	Receive cues stronger (+)
		Small	Receive cues weaker (−)
	Concurrent ocean coverage	**Large**	**Wide monitoring angle (+)**
		Small	**Narrow monitoring angle (−)**
	Spatial resolution	**High**	**Cue more accurately displayed (+)**
		Low	**Cue less accurately displayed (−)**
	Stabilization	**Good**	**Wider monitoring angle (larger CCOS) (+)**
		None	**Narrower monitoring angle (smaller CCOS) (−)**
	Thermal resolution	**High**	**Cue more accurately displayed (+)**
		Low	**Cue less accurately displayed (−)**

Table 3 (continued)

Method	Internal factor	Property	Strength (+) or weakness (−)
RADAR	**Antenna height**	**Low**	**Improved range for animals (+)**
			reduced range for ice (−)
		High	**Improved range for ice (+)**
			reduced range for animals (−)
	Antenna type	**Vertical only**	**Improved detection (+)**
		Horizontal only	**Decreased detection (−)**
	Frequency modulation	**Used**	**Improved clutter removal (+)**
		Not used	**Increased clutter (−)**
	Polarimetric filtering/antenna	**Used**	**Improved clutter removal (+)**
		Not used	**Increased clutter (−)**
	Scan rates	**High**	**Higher probability of short signal detection (+)**
		Low	**Lower probability of short signal detection (−)**
	Solid state core	**Used**	**Longer operational life (+)**
		Not used	**Shorter operational life (−)**
	System power transmission	Low	Decreased detection range (−)
			lower power needs (+)
		High	Increased detection range (+)
			higher power needs (−)
	System resolution	**High bandwidth**	**Cue more accurately displayed (+)**
		Low bandwidth	**Cue less accurately displayed (−)**

effectiveness of a monitoring system is provided below and summarised in Table 3.

4.1.1. PAM

A fundamental requirement for any PAM system is that it should have good sensitivity in the frequency range of the sounds emitted by the species of interest. Achieving this may not be straightforward. The huge differences in the peak frequencies (frequencies with the most energy) in the vocalisation of marine mammal species may mean that different types of PAM systems, with differing sensor types and configurations, signal conditioning and digitisation capabilities, and different software detectors and localisation algorithms may be required for the detection of different species. Similarly, array design and gain, depth of deployment, filter settings and sampling rates should all be optimised for the signals of interest and the environmental conditions. With a simple two element array, localisation of a sound source can sometimes be achieved using target motion analysis; i.e. taking multiple bearings to a sound source and estimating their crossing point. However, this method is far from ideal for mitigation since it requires the monitoring hydrophone to be moving past vocalising animals, and assumes that the sound source (marine mammal) is either relatively stationary or moving very predictably. Further, the time taken to collect sufficient bearings to an animal can make it hard to make timely decisions. More sophisticated large aperture arrays with appropriately spaced hydrophones can potentially localise an animal instantaneously from a single vocalisation but these are challenging systems to deploy alongside the seismic gear towed behind the seismic survey vessel.

The ratio of the cue's energy and the energy of the noise (the non-signal energy), the "signal to noise ratio", is critical for cue detection. The total noise affecting detection will come from many sources. The system's electrical self-noise can be reduced by using well designed and high quality equipment. Flow noise resulting from the hydrophones being towed through the water will be more significant at lower frequencies and can be reduced by streamlining. Noise from the towing vessel and its associated machinery (e.g. seismic sound sources, locating pingers, clanking chains) is usually the dominant source of background noise during seismic surveys. The most straightforward way of dealing with ship noise is by positioning the PAM streamer as far from the noise sources as possible, most straight forwardly by towing them on long cables. However, the need for a 'quiet' deployment must be balanced against the need to have a high detection probability close to the sound source itself and deploying long cables in the busy environment at the back of a seismic survey vessel can be challenging due to the risk of entanglement with other deployed gear.

Integrated systems (see Section 2), which tend to use many more hydrophones than ancillary systems, can combine the outputs of those hydrophones using beam forming techniques, thereby providing "array gain". The system then effectively listens in one direction (or a set of directions) and is able to reduce the influence of noise arriving from other directions. Array gain increases the detection range, but will also narrow the receiving beam, leading to a patchy coverage of the monitoring area, and animals may consequently be missed.

The dynamic range (the ratio of the loudest sound to the quietest sound that can be detected) can also be important, particularly in systems spanning a wide frequency range. For example, low amplitude high frequency sounds may be lost in the presence of loud low frequency noise that exceeds the system's dynamic range. A digitiser with a greater bit depth or resolution, e.g. 24 bit as opposed to 16 bit, can contribute to an increased dynamic range.

There have been few attempts to directly measure the efficacy of PAM conducted during seismic surveys, and studies have usually compared visual and acoustic detection rates and have presented contrasting results. An early study (Gordon et al., 2000) found that acoustic detections of odontocetes using a towed PAM system during a seismic survey were an order of magnitude higher than detections made by a visual MMO at the same time (note that in this case the hydrophone was being deployed from a guard boat ahead of the seismic vessel). By contrast, Stone (2015) reviewed detection data from seismic surveys conducted in UK waters between 1995 and 2010, and found that, during periods when both visual and acoustic monitoring were taking place, visual detection rates were higher for all species groups. Variability in such comparisons is perhaps not surprising. The relative performance of these methods will be influenced by a number of factors including target species, environmental conditions, equipment type and how it is deployed and the skill of monitoring personnel. However, it does seem to be the case that PAM is not achieving its potential during typical seismic surveys. A common concern, raised with us by MMOs, is that ancillary PAM hydrophone(s), which are typically fitted at sea "around" existing complex arrays of towed seismic gear, are often deployed on rather short cables and towed close to the vessel where propeller and machinery noise is dominant and compromises PAM performance.

4.1.2. AAM

For AAM, the source level (sound pressure level at 1 m distance from the sound source) of the outgoing sonar pulses, their type (mainly determined by pulse duration and bandwidth) and frequency should be adapted to the size of the target. Generally, raising the source level

increases the detection range, as does lowering the frequency of the pulse (as absorption of the sound energy decreases at lower frequency). Similarly, longer pulses increase the returned energy from a target. However, increased source level or pulse length may also produce more reverberation and therefore increase the background noise. The frequency bandwidth of the emitted pulses influence the resolution of the image of a detected animal – broadband pulse will give a higher resolution than narrowband pulses. Similarly, as the frequency increases, the wavelength decreases, and therefore the resolution improves. Systems with motion compensation automatically correct beam steering directions for ship motion and therefore have improved performance at high sea states. Some AAM systems can have blind spots, i.e. small areas within the field of view where nothing can be detected, either due to acoustic propagation effects, placement of the system on the vessel, or overloading the receiver when the AAM system is emitting (if the source and receiver are next to each other). In some cases, not only are the source and receiver near each other, but the same transducers may be used to both transmit and receive. Generally, the system will not be able to receive while it is transmitting, with the result that targets near the transducers cannot be detected. As with PAM, a low noise system is favourable for the system's performance as it contributes to higher signal-to-noise ratio. Emission of sonar pulses in the hearing range of marine mammals, especially with high source levels, may have an impact on marine mammals (e.g. DeRuiter et al., 2013; Hastie et al., 2014) and should therefore be considered in any impact assessment.

4.1.3. Thermal IR

We define the concurrent covered ocean surface (CCOS) as the field of view covered by all cameras of a system at any given time, with sufficient resolution to be able to detect the target species. The CCOS is largely determined by the system design. Systems with large CCOS are desirable, because they cover a greater monitoring area at any given time, increasing the likelihood of an animal surfacing in a covered zone and hence being detected. Stabilisation of systems improves thermal IR system performance and results in a larger effective CCOS compared to systems that are not mechanically stabilised. In mechanically stabilised systems the vertical field of view is kept constant, therefore the automatic detection algorithm always has the full available data to search for animals. An electronic stabilisation will reduce the CCOS, as the camera is not pointing at the same ocean patch at every given time, and only the section of the ocean monitored for a at least a few seconds consecutively can be analysed. However, it will still perform better than an un-stabilised system. Cameras that can pick up a wider thermal-IR frequency band may receive stronger cues than those with a narrower band, therefore very narrowband IR sensors should be avoided. Camera bands specifically adapted to prevailing atmospheric IR radiation windows provide the best signal-to-noise ratio. Each IR camera/lens combination has to be tuned to the desired wavelength. Proper adaptation is therefore a positive factor. Higher spatial resolution and better thermal resolution result in more accurately displayed cues and usually in a larger detection range. They are therefore highly positive factors. It has to be noted that the positioning of an IR system on board of a vessel has to be very well designed to avoid reflections from vessel decks that appear in the image and hot areas and hot smoke particles can seriously hamper the performance. An ideal placement is as high as possible to reduce obstruction by the vessels superstructure and in front of the exhaust.

4.1.4. RADAR

The ability of RADAR to discern marine mammals amongst background noise is improved if the system has high resolution (high bandwidth provides finer resolution and improves identification), high power transmission (resulting in increased range) and high scan rates (that better detect short surfacing events) (Briggs, 2004). Frequency modulation, where the transmission signal frequency is varied, results in a better ability to simultaneously measure the target range and its

relative velocity, as well as to reduce unwanted return echoes (so called clutter) such as from rain or waves (Leong and Ponsford, 2008). Frequency-Modulated Continuous-Wave (FMCW) RADAR is a special type of RADAR, which radiates continuous transmission power (as opposed to pulsed). FMCW RADAR can also change its operating frequency during the measurement, thereby improving detection success. Systems with solid state cores are reported to have greater running life but are more expensive than magnetron alternatives. While detection range is generally related to antenna height and target height, a relatively low antenna height results in improved detection rates for marine mammals (pers. Comm. S. Wärnfeldt). All these factors should be adapted to the specific monitoring purpose. Many standard RADARs transmit and receive radio waves with a single, horizontal polarization (i.e. the electric field wave crest is aligned along the horizontal axis). Polarimetric RADARs, on the other hand, transmit and receive both horizontal and vertical polarizations. Cross-polarized RADARs transmit and receive polarizations orthogonal to one another (i.e. transmit horizontally and receive vertically, Briggs, 2004). Polarimetric and cross-polarimetric RADAR requires specialised antenna and receiver and display systems, however, polarimetric antenna and filtering raises the detection abilities of RADARs in sub-optimal conditions by removing clutter (e.g. Anderson and Morris, 2010).

4.2. Extrinsic factors

An overview of the animal dependent and environmental extrinsic factors and their properties, and how they influence the probability of detecting an animal is provided below and summarised in Table 4 (animal dependent factors) and Table 5 (environmental factors).

4.2.1. Animal dependent factors

As noted above, any monitoring method is affected by cue availability. For methods relying on the animal being available at the water surface (e.g. thermal IR, RADAR and visual observers) the animal's behaviour, its size, surface behaviour, the strength of its exhalation as well as its diving behaviour and school size are factors that will directly affect detection probability. The bigger the animal's body or its exhalation, the larger the school size, the more energetic its surface behaviour, and the more frequently it surfaces, the more likely it is to be detected (e.g. Silber et al., 2009). An animal's colouration may influence visual detection probability, with contrasting colouring being more readily detected. Animals cannot be detected by thermal IR, RADAR or a visual observer while they are submerged. Visual observers, though, may detect animals close to the water surface if the water is clear. The longer an animal dives, the more likely it is that an animal will enter a mitigation zone without being detected by these methods.

Table 4
Animal dependent factors that may influence (Y) the detection performance of monitoring methods AAM, PAM, thermal IR, RADAR or visual observer, or may have no influence on their detection performance (−).

	AAM	PAM	Thermal IR	RADAR	Visual observer
Animal behaviour	Y	Y	Y	Y	Y
Animal colouring	–	–	–	–	Y
Animal size	Y	–	Y	Y	Y
Strength of exhalation	–	–	Y	Y	Y
Displayed surface behaviour	–	–	Y	Y	Y
Diving behaviour	Y	Y	Y	Y	Y
Movement in relation to monitoring system	Y	Y	Y	Y	Y
Position relative to water surface	Y	–	–	Y	Y
School size	Y	Y	Y	Y	Y
Vocalisation	–	Y	–	–	–

Table 5

Environmental factors that may (Y) or may not (N) influence the detection performance of monitoring methods AAM, PAM, thermal IR, RADAR or visual observer. Those factors that are advantageous for the detection performance when increasing are marked with (+), those that are disadvantageous, i.e. the detection performance is decreasing with increasing factor, are marked with (−).

	AAM	PAM	Thermal IR	RADAR	Visual observer
Aerosols	N	N	Y(−)	N	N
Acoustic background noise	Y(−)	Y(−)	N	N	N
Fog	N	N	Y(−)	Y(−)	Y(−)
Glare	N	N	Y(−)	Y(−)	Y(−)
Light level	N	N	Y(−)	N	Y(+)
Rain	Y(−)	Y(−)	Y(−)	Y(−)	Y(−)
Sea bed properties	Y	Y	N	N	N
Sea state	Y(−)	Y(−)	Y(−)	Y(−)	Y(−)
Snow	N	N	Y(−)	Y(−)	Y(−)
Surface expression of non-targets	N	N	N	Y	Y
Vertical sound speed profile	Y	Y	N	N	N
Water depth	Y	Y	N	N	N
Water temperature	N	N	Y(−)	N	N

For AAM, which detects animals when underwater, sonar target strength (the proportion of energy reflected by the target) is a key determinant of detection probability. This correlates well with body size and may also be influenced by school size. In addition, movement and diving behaviour will also affect detection probability. Animals displaying certain patterns of behaviours, i.e. movements and diving patterns that often take them across the field of detection of the AAM system are more likely to be detected than others, such as deep divers that are below the field of detection. Clutter, created by reflection of the sonar pulses from the water surface, makes animals close to the water surface harder to detect than those in mid water (Urick, 1984).

PAM depends on detecting sounds produced by the animals, and therefore can only detect individuals which are vocalising. Sound production is not obligatory for any species; and for some, acoustic availability bias can be both large and highly variable. The acoustic characteristic of their vocalisation is a key determinant of detectability. The frequency of vocalisations varies enormously between species, from the infrasonic calls of the large baleen whales (e.g. McDonald et al., 2006; Stafford et al., 1998), which can be as low as 10 Hz, to the ultrasonic clicks of dolphins and porpoise at 130 kHz (e.g. Au et al., 1999; Mohl and Andersen, 1973). Similarly, signal duration can vary from > 10 s (blue whale moans, *Balaenoptera musculus*: Linnaeus, 1758) to < 100 μs (harbour porpoise clicks, *Phocoena phocoena*: Linnaeus, 1758). Source levels also cover a huge range and include some of the highest values reported for sound producing animals (e.g. Mohl et al., 2003). Some vocalisations, particularly the echolocation clicks of odontocetes, are highly directional, being projected in a narrow forward facing beam (Au, 2012), while others are less directional (e.g. baleen whale calls, Sirovic et al., 2007). The effect of cue directionality on detection can be complicated. If the acoustic energy of a signal is emitted in a narrow beam, detection range is enhanced for a sensor that happens to be within the beam but diminished for sensors outside it. The probability of detecting sufficient sounds for a mitigation decision may also be very dependent on animal movement, e.g. how often the animal orientates itself towards the receiver. Some species show seasonal or diurnal variability in vocal behaviours, which may also vary with the animal's gender and age (e.g. Dunlop, 2016). Most aspects of vocal behaviour vary between activities such as foraging, travelling or social behaviour, and may be affected by group size, diving behaviour, and the presence or absence of predators or prey. A further complication is that vocalisation rates may be influenced by the presence of the seismic survey vessels themselves and by the sounds of air gun arrays. For example, Blackwell et al. (2015) showed that bowhead

whales (*Balaena mysticetus*: Linnaeus, 1758) change their vocalisation rates in response to seismic survey activities. All of these extrinsic factors have a direct influence on the detection performance of a PAM system. Knowledge on the vocalisation characteristics of the target species is needed to understand the likelihood of detecting those animals and hence the effectiveness of these systems for mitigation purposes.

4.2.2. Environmental factors

Environmental factors can influence a system's performance by degrading a cue while it travels from the sender to the receiver, or by creating noise energy that may mask a cue or trigger false alarms. These factors are summarised in Table 5.

Any cue will lose energy as it passes through the environment from its source to a detector. This is known as transmission loss. Active systems such as AAM suffer this loss twice (i.e. on both the transmit and receive path). For above-water systems, transmission loss is influenced by atmospheric conditions such as fog and rain, and the magnitude of this effect depends on the methodology considered. Both PAM and AAM will be affected by the way sound propagates underwater. This depends on the characteristics of both the propagating sound and the water body it is passing through, including the nature of surrounding boundaries: the sea floor (e.g. bathymetry and bottom type) and water surface (in particular its roughness, which is mainly determined by the sea state). Sound velocity gradients (which usually occur with depth) cause sound refraction and can have strong effects on propagation and detection ranges.

Noise in the environment can both mask cues and trigger false detections. For acoustic systems any natural or anthropogenic acoustic background noise which overlaps the cue in both time and frequency can result in masking. In most monitoring scenarios for mitigation purposes, detection systems must operate in environments where levels of anthropogenic sound (e.g. from vessels, seismic sources and ancillary acoustic sources) may be very high. For AAM systems, the transmitted sonar signals can cause reverberation when scattered by reflective surfaces or objects other than the target species (e.g. the sea floor, a rough sea surface, fish schools). Objects (debris) floating on the sea surface, as well as waves, reflects RADAR pulses and become a source of noise that may lead to false detections. Sun glare can be troublesome for visual and thermal IR methods, especially in the higher frequency IR bands. Thermal IR based camera systems rely on the apparent temperature difference between the whale blow or body surface and a cooler background, which is usually the ocean surface. Sea surface temperature is therefore a crucial variable in determining a thermal imaging device performance. The warmer the water, the higher the noise.

High sea states are characterised by a rough sea surface and breaking waves which are a source of both visual noise above water and acoustic noise underwater. White caps and rough sea surface can hinder the detection of animals by all above-water monitoring methods, and a raised acoustic background noise degrades PAM and AAM performance. Increasing sea state is therefore unfavourable for all detection systems; however, it has less effect on thermal IR and PAM than on AAM and RADAR (Briggs, 2004; Zitterbart et al., 2013).

Fog has a strong impact on detection probability for visual observers and thermal IR, and some influence on RADAR (depending on the type). Low to zero light levels do not have any negative effect on the non-visual methods mentioned here. Thermal IR is typically more effective at night, when there is less reflected radiation from the sun and sky than during the day. Very heavy rain compromises the performance of all the systems. Rain can lead to masking of cues for above water methods (Briggs, 2004) as well as under water methods due to an increase in the underwater background noise levels.

5. System evaluation

We evaluate the performance of the systems listed in Table 1, and highlight factors that need to be considered when choosing the appropriate system for a given monitoring scenario. In addition to the system's ability to detect, classify and localise animals, we also consider system costs, commercial availability, and installation requirements.

5.1. Detection performance and classification and localisation abilities

The detection range of each method depends greatly on the regime that the system is to be used in, environmental conditions and species behaviour. For example, species with high surface activity will much more likely be detected by above water methods than those with low surface activity. Often, a number of different species will need to be detected and a single system may not be optimal for all species of interest during a particular deployment. The decision on which method to use needs to be made for each scenario independently based upon the species of interest and the likely performance of each system in the expected environmental conditions.

5.1.1. PAM

Ancillary PAM system streamers contain several hydrophones which can and should be chosen to have good sensitivity in the frequency band of the range of species of interest. Typically, more than one type of hydrophone will be required to cover the full range of frequencies of interest and often these take the form of pairs of matched hydrophones spaced so that time of arrival analysis can be applied to signals from each pair to calculate bearings to sounds. A typical configuration in the systems reviewed is to have a low to medium frequency pair of hydrophones ($\sim 50\,Hz$–$30\,kHz$) for the detection of whales and a high frequency pair ($\sim 2\,kHz$–$150\,kHz$) enabling the detection of dolphins and porpoises as part of the same streamer section with inter-hydrophone spacing of $\sim 3\,m$ and $25\,cm$ respectively. The Delphinus system was designed to make detections in the mid-frequency (MF, up to 12 kHz), high-frequency (HF, up to 40 KHz) and ultra-high frequency (UHF, up to 150 kHz) bands. The MF and HF section consists of an array of 16 hydrophones while UHF sensing is provided via single sensors. Recent developments in the Delphinus array include using "triplet" hydrophones, which resolve left-right ambiguity and can provide a slant angle to animals vocalising at depth (Sheldon-Robert et al., 2008; von Benda-Beckmann et al., 2010).

For animals that vocalise consistently, and which are relatively slow moving (e.g. sperm whales, *Physeter microcephalus*: Linnaeus, 1758), target motion analysis can be used to determine the most likely location of a vocalising animal based on the changing patterns of bearings as the hydrophone is towed and passes a vocalising animal. In addition, it is possible that a combination of received levels and some summary statistics (e.g. average of maximum whistle intensity) and bearing could potentially provide some indication of range. However, empirical work is required to underpin and test this. Typical maximum detection ranges for some species, such as harbour porpoise, are less than the mitigation zone ranges specified by some regulations. In these cases any detection should trigger mitigation action.

Two systems were reviewed that utilise the hydrophones of the seismic streamers for marine mammal detection. These have the advantage over ancillary streamers of being able to utilise a large number of hydrophones which can be accurately located using the seismic array positioning management systems. This allows for the application of advanced beamforming techniques which can both improve signal to noise ratio and localise detected calls.

The seismic streamer hydrophones of the WhaleWatcher system are spaced $\sim 3\,m$ apart within the array and are sensitive up to $\sim 250\,Hz$, while the point positioning hydrophones (which are primarily used to determine the arrays position), are spaced $\sim 60\,m$ apart and have a higher frequency response, to about 4 kHz. The low frequency detection

component of WhaleWatcher makes use of the large hydrophone number and large aperture available within the seismic streamers to implement beamforming on specified sub-arrays. This should improve signal-to-noise ratio and provide bearing information. Coherent signals, such as vocalisations from baleen whales or dolphin whistles below 4 kHz, are detected by software and localisation is attempted automatically. Spectrograms of these detections are also available to human operators to allow their expert input on classification. The higher frequency point source elements are too sparse to allow beamforming, however, background noise is a lesser issue at higher frequencies. Comparison of the time of arrival of higher frequency signals at the elements in such a large array could potentially provide instantaneous localisations. Locations are plotted as a layer in one of the standard seismic acquisition displays making the data immediately available to the seismic crews as well as to visual observers. One obvious shortcoming with these systems is the limited bandwidth of the available hydrophones. While well suited for baleen whales, detection of some odontocete species will be severely restricted, and many species will not register at all. Less information is available for the Sercel QuietSea system. Like WhaleWatcher, the seismic elements of the QuietSea system are used for low frequency monitoring, while point positioning hydrophones provide coverage at mid-frequencies. Additional UHF elements can be attached to the seismic source array with signals returning via additional cables. The upper frequency limit of these elements is currently $\sim 96\,kHz$, which allows the monitoring of most whale and dolphin species but not porpoises; however, there are plans to improve this to cover the full frequency range of marine mammal vocalisations in the future (Sercel, pers. Comm.). As both of these systems are very recent developments, there are, as yet, few published data on the performance and accuracy of detection, localisation and classification of real marine mammal targets.

If and when integrated PAM systems are equipped with appropriate broadband hydrophones they should, in theory, be able to detect all cetacean vocalisations and classify them to species group. The large number of hydrophones, in a relatively quiet location well astern of the seismic vessel, large array aperture and ability to use approaches such as beam forming to improve signal-to-noise ratio, should mean that these arrays will ultimately be better at detecting and localising baleen whales than the ancillary systems mentioned above. Localising rapid, repeated and highly directional signals such as odontocete echolocation clicks with a small number of widely separated HF or UHF elements will be challenging and is likely to need additional hardware and software development.

Taxonomic classification of signals detected on both types of PAM system should be possible at least to the level necessary for monitoring for mitigation purposes (i.e. the ability to distinguish between classes of marine mammals for which guidelines require different mitigation monitoring actions, such as large whale, dolphin, beaked whale or porpoise).

Acoustic detection ranges will vary greatly between species and scenarios, being affected by both the acoustic characteristics of the sound source (intensity, directionality and frequency) and also the total noise on the detection system (comprising environmental and anthropogenic background noise, flow noise and system noise). Towed hydrophone population surveys provide a realistic indication of the ranges that might be achieved. For example, the effective half strip width (EHSW, a range derived empirically during line transect surveys for which on average as many animals are detected beyond it as are missed within it) for sperm whales of 10 km measured by Lewis et al. (2007) is typical for towed hydrophone surveys of this species under relatively favourable noise conditions. Detection range for harbour porpoise is over an order of magnitude lower with the EHSW of ~ 230–$300\,m$ reported by Cucknell et al. (2016) being typical. Baleen whales can potentially be detected at ranges of many tens of miles; however, the overlap between their low frequency vocalisations and low frequency background and flow noise will typically limit detection range with

towed hydrophones. Abadi et al. (2015) report detecting and localising baleen whales by processing data from multiple hydrophones in a research seismic streamer. They showed that they were able to both locate an artificial sound source accurately and to localise baleen whales calls (thought to be from humpback, *Megaptera novaeangliae*: Borowski, 1781, or blue whales) at ranges as great at 26 km. Ancillary towed systems are typically limited to detecting the higher frequency baleen whale calls such as those in humpback whale (e.g. Norris et al., 1999) and right whale (*Eubalaena sp*.) (e.g. Matthews et al., 2001) repertoires.

The sort of detection ranges and detection probabilities achievable by PAM systems for sperm whales and most small cetaceans should generally be useful for improving the monitoring required by most jurisdictions. It is reasonably likely that vocally active groups of these species could be detected before they entered zones with radii of a few thousand meters, provided hydrophone systems were well deployed. Currently available ancillary PAM systems, however, are often unable to provide adequate information on the detection range for small cetaceans and might be best used in conjunction with other methods such as visual observers or AAM. PAM would primarily be useful in detecting animals and providing a bearing that can be used to direct visual observers or AAM to search in the appropriate direction and determine range.

Potentially, detection range for integrated systems should be greater than those achievable using ancillary streamers because better deployments and the ability to perform beam forming should reduce noise. As discussed above, these systems should also have better localisation capabilities. However, the detection range for some species, such as porpoise, may still remain too short to cover the monitoring zones required by many regulations.

5.1.2. AAM

The detection ability of AAM systems is dependent on a number of parameters. Frequency, source level, beam shape and waveform are four of the important considerations. High frequency systems such as the Echoscope (Coda Octopus Products Ltd) have been used to detect, classify, and localise small animals, such as dolphins, at ranges of 50 to 100 m (Hastie, 2013). Its maximum display range scale is 100 m with a maximum range of 80 m, while the Gemini 720 (Tritech International Limited) has a maximum range of 120 m. Classification is enabled through image and movement recognition. At the other end of the frequency range, the SX90 (Kongsberg Maritime Subsea) has been tested for the detection of animals such as seals up to a maximum of 2 km (Pyć et al., 2016). The HFM3 (Scientific Solutions Inc.) maximum display range scale is stated to be 2 km. The CMAS-36/39 system (Nautel C-Tech Limited) has a maximum display range of 4 km. Though detections to that range have not been demonstrated, the combination of lower frequency and higher source level for this model would, at least in theory, support the greater predicted maximum range.

Similar to detection, classification and localisation ability is dependent on a number of parameters. The travel time from the emitted pulses to return as an echo is converted to a range from the source, and bearing is achieved through the directionality of the transmitter and receiver. However, range errors are also affected by the type of waveform. A frequency modulated (FM) pulse such as used by SSDN, HFM3, SX90, SU90 and CMAS 36/39 provides improved range resolution compared to continuous wave (CW) pulses such as those available on the CMAS 36/39. CW waveforms, on the other hand, can provide additional information regarding animal speed (Urick, 1984). Classification of animal species groups must be achieved by imaging (similar to that used for human ultrasonic echocardiograms), animal behaviour (motion), or in combination with an alternate technology. The HFM3 only uses the echo strength as a classification aid allowing large animals such as humpback whales to be distinguished from small animals such as delphinids.

With maximum detection ranges being < 100 m, the Echoscope and Gemini 720 are considered to be inadequate to meet regulatory

requirements for seismic surveys. With a theoretical detection range of up to 2 km, the SSDN, HFM3, SX90, SU90 and CMAS 36–39 are likely able to meet regulatory requirements with the caveat that there are environmental conditions (e.g. strong downward refracting sound speed profiles) where no system will be able to meet desired detection range.

5.1.3. Thermal IR

As thermal IR methodology is based on the detection of temperature differences, it is generally restricted to the detection of warm blooded animals, in our case marine mammals. There is very little information on the performance of thermal IR for marine mammal detection. To date there are two reports that evaluate the detection performance of the systems on free living whales (Weissenberger and Zitterbart, 2012; Zitterbart et al., 2013). Both evaluated the AIMMMS system for sensor performance and the performance capabilities of auto-detectors. Large whale blows or surface displays were detected at distances up to 5–8 km under ideal conditions in a cold water environment. Smaller cetacean blows or surface displays were detected at distances of approximately 3–5 km for small whales and shorter ranges, of up to 1.5 km, for porpoises and walruses. The performance of NightNavigator was evaluated using fake whale blows that were produced using hot steam and emitted from a barge (Current Corporation, 2011). These fake whale blows, which were 3–6 m in height, were detected at distances of up to 2 km. Tests were conducted without an automatic detection algorithm, therefore evaluation was based on human verification of the video stream.

Information on the detection performance of Toyon, Polaris, RADES, Hyper-Cam and Gobi systems was provided through the questionnaire survey (Verfuss et al., 2016). Polaris provides detection performance estimates that are rather conservative with 400 m detection ranges for large whales, 250 m for medium-sized cetaceans, and 130 m for single animals and groups of small cetaceans and seals. The detection distances that at minimum can be achieved for the RADES system are estimated by Seiche to be 2 km for large cetacean, 1.5 km for medium cetaceans, 1 km for individual small cetaceans, and 2 km for groups of small cetaceans. For seals a maximum detection distance was estimated to be 500 m. Toyon provided detection distances for their land-based setup. These could be > 8 km in ideal conditions for large whales, but generally around 2–5 + km. Detection distances for medium cetaceans would be around 500 m–2 km, and 1–3 km for small cetaceans. They state that these distances are highly dependent on species behaviour and conditions. Overall, a comparable detection performance by human screening (having a human watching the IR video stream and distinguishing whether a whale blow was recorded, as opposed to real-time automated detection) can be expected from other systems that use similar combinations of thermal imaging sensor and lens. There was a consensus that large cetaceans can be detected at ranges of up to 2 km, while the detection distance of small cetaceans is significantly reduced. No information is available on performance in equatorial regions. It should be noted that these detection performance metrics only apply if the animal is available for detection in the field of view and at the surface. The manufacturers did not state if these numbers represent performance of an automatic detection algorithm or a human observer, so it is assumed that they were based on a human screening process. Note too that the stated detection performance from AIMMMS, RADES, NightNavigator and Toyon are based on empirical data while the detection performance of Polaris, Hyper-Cam and Gobi are estimates provided by the system provider.

The overall detection performance is a multi-stage process. While it is possible to infer detection performance from technical parameters (e.g. how likely it is that enough photons emitted from a whale at distance x will be recorded by sensor y if sensor y is pointed towards the whale), the overall detection performance (how likely is it that whale A will be detected by system B) is highly dependent on the whale's surfacing profile and the CCOS. If a planar camera rotates too slowly, it

could miss a whale surfacing if the camera was pointing in a different direction, even if it was within range. This fact is reflected by the CCOS variable, and it is not possible to derive an overall detection probability for any of the planar camera systems without knowing the CCOS.

Systems using a cooled camera will have a higher thermal resolution due to lower system noise but they come at a higher purchase and maintenance costs. The cooling mechanism usually requires to be serviced after several thousand hours of use, as opposed to uncooled camera systems which do not need regular maintenance. Although no values for system noise were specified, thermal resolution for cooled camera systems is typically increased by a factor of 3–5 compared to uncooled systems which ultimately leads to higher detection probability.

It is highly desirable to use automatic detection algorithms that pre-screen the data and subsequently present only short video sections that resemble a whale blow or surfacing animal to the human observer for verification. Processing of the video stream needs to be performed in real time, with a time-lag between surfacing and presentation of potential detections of not more than a few seconds. This is necessary to allow the MMO to perform a timely decision as they would if the animal were detected visually. Additionally, a timely detection makes it more likely that an MMO can be guided to make a direct visual verification that makes an identification to species level feasible. This is a key feature when this technology is used for mitigation monitoring. Automatic detection algorithms are available for the following systems: AIMMMS, Toyon, and RADES. The rate of false positives was only provided for the AIMMMS system and was reported as an average of 6 false alerts per hour in cold water environments. The extent to which these reliably predict performance during any given monitoring exercise will depend on how the conditions they were measured under compare to those during monitoring for mitigation purposes. All system providers state that real-time detection is available, but, studies to measure detection reliability and false alarm rate have only been completed with the AIMMMS system. If no automatic detection algorithm is provided, real-time detection would presumably mean a real-time screening of the video stream by human observers.

There are no data on the reliability of classification of detections into species groups, nor on the quality of localisation; but all systems that either can see the horizon in the image, or know their absolute inclination angle and are mounted some tens of meters above the sea surface should have the potential to localise the animal with a reasonable error of < 10–15% of the distance (Leaper and Gordon, 2001).

Overall, thermal IR based marine mammal detection can be effective at detecting large whales at distances of a few kilometers which should be useful for monitoring for mitigation purposes; provided that the animal is available to be detected (i.e. at the surface), the sensor has a large CCOS in ideal conditions (i.e. no fog, little wind, low sea state) and the camera system is mounted sufficiently high (at least ~15 m). Systems using a Mid-Wave IR (MWIR) camera will be more affected by fog and glare than systems using a Long-Wave IR (LWIR) camera. Given these optimal conditions, medium sized whales will be detectable reasonably well in up to 1.5 km, and larger whales with reliable detection ranges at 2 km. Reliable detection of small whales and dolphins should be possible up to 500 m, and pinnipeds at < 500 m. A reliable detection of pinnipeds and porpoises up to 1 km is unlikely and has not yet been documented (maximum detection range for a walrus is currently 800 m, Weissenberger and Zitterbart, 2012).

These detection distances could be increased by reducing the field of view (i.e. increasing the magnification level of the camera setup) and increasing the number of cameras to attain maximum CCOS. Such a system would, however, be difficult to handle, require stabilisation, produce challenging data volumes and has yet to be developed. Recent development of very small IR cameras for the drone market might make such a system feasible in the near future, without adding much effort to stabilisation and operation personnel required.

5.1.4. RADAR

There is a clear lack of empirical data on the ability of RADAR to detect marine mammals in real world conditions. RADAR has a 360° detection zone, but requires targets to be at the surface for detection. Cue masking, especially when targets have small cue amplitudes and the sea state are higher, is considered the most significant short-coming of RADAR, while the ability to use RADAR detection at night and potentially in fog or rain are clear advantages. Attempts to utilise conventional RADAR systems to detect marine mammals reveal some potential for detection in optimal sea state conditions at short range (< 1 km) with diminishing performance at greater distances (DeProspo et al., 2005; Forsyth, 2011). These studies by Arête Associates used an adapted commercial Furuno marine RADAR and reported, that in optimal sea state conditions, whales can be detected (using automated algorithms) more than half the time at 1 km and were very rarely detected up to 5–6 km. Commercial RADAR systems were reported to be moderately affected by fog, heavy rain and high sea state (depending on the system) and performance was unaffected by darkness or low light conditions (Briggs, 2004). Discussions with manufacturers of high performance RADAR systems lead us to conclude that, for species with high RADAR Cross Sections, more sophisticated RADAR (e.g. surface detection/frequency modulated with polarimetric antenna) might achieve higher detection rates than conventional marine RADAR. High performance RADAR also is reported to be less affected by fog, rain and moderate sea states and has the ability to detect ice. Marine mammals with small RADAR Cross Sections (e.g. seals and porpoise) are poorly detected by any system in anything but optimal conditions. Classification appears to be restricted to coarse inference based on body size. Detection of large marine mammals (polar bears and walrus) on floating ice was reported using RADAR Technology AS systems. Commercial systems have automated vessel detection and tracking capabilities, but these appear unsuitable for marine mammal detections.

Brainlike Inc. and Arête Associates have both worked on using outputs from standard marine RADAR to automatically detect whales (DeProspo et al., 2005; Forsyth, 2011). Both groups highlight issues with detecting and removing false positives that result from non-targets like logs and buoys, as well as clutter from wave crests. Brainlike Inc. tested the detection of grey whales (*Eschrichtius robustus*: Lilljeborg, 1861) using RADAR mounted in aircraft and a near real-time detection system, the Brainlike Processor™. This beta system has been built to combine outputs from cameras and RADAR to identify likely marine mammals for secondary review by a human operator.

The weaknesses of RADAR presently are poor performance in real world conditions, the inability to differentiate between species, the scarcity of empirical detection data (notably for smaller animals/cues) in a range of low visibility field conditions, the lack of any available and tested automated systems to assist in marine mammal detection and identification/removal of false positives. Cue masking appears to be a potentially significant issue when clutter increases due to high sea state (and given signals from marine mammals will be intermittent and variable). Overall, it is currently uncertain if even specialised, and so relatively expensive, RADAR systems could be an effective monitoring tool in low visibility conditions.

5.2. System costs, commercial availability and installation requirements

5.2.1. PAM

Ancillary PAM systems are available "off the shelf" from a variety of suppliers and are now increasingly considered a standard component of monitoring programs for mitigation purposes. In most cases systems are rented for specific projects. Indicative costs for renting a full commercial system, which would normally include several cables and 100% redundancy with spare equipment were between $300 and $500 USD per day (note: some respondents declined to provide rates as they considered it to be commercially sensitive information). Purchase prices ranging between $10,000 and $100,000 USD were quoted by

some respondents. PAM systems are relatively easy to install on vessels. They are routinely shipped to seismic vessels on station and fitted at sea in roughly half-a-day to a day. MMOs can usually achieve this with the help of vessel crew and other seismic personnel, though several equipment providers expressed their preference for sending one of their own engineers to conduct the initial setup. Once deployed, PAM equipment is at risk of becoming entangled with seismic equipment, and this often leads to restrictions on how ancillary streamers can be deployed. Entanglements are not uncommon and usually result in damage to the PAM streamer. We are not aware of any incidents in which seismic streamers or gear (which are much larger and more robust than the PAM systems) have been damaged.

No cost estimates were available for integrated PAM systems, and it is likely that with such an involved commercial product there may be several cost components. The most substantive financial implication of these systems is that they are dependent on the vessel being equipped with the appropriate customised seismic system; changing a complete seismic acquisition system is clearly an enormously significant and long-term decision for any seismic operator. Systems are thus available only if vessels are equipped with the appropriate seismic hardware. The WhaleWatcher system only utilises hydrophones which are already present in the steamers, it is not clear if any additional processing hardware is required on the vessel. Sercel's QuietSea system requires additional high frequency hydrophones which are fitted on the source array and it should be relatively easy to retrieve this to allow additional sensors to be fitted. There is no apparent conflict of these systems with the seismic vessel's equipment or existing survey protocols.

5.2.2. AAM

All of the AAM procurement estimates from providers exceeded $100,000 USD with the exception of the AN/SSQ-963 (Ultra Electronics Sonar Systems), and with the Sentinel (Sonardyne) exceeding $300,000 USD. Each of the systems, other than HFM3 (which is under development), appear to have a high technical readiness level and are currently available for commercial use. Some, but not all, AAM systems require permanent installation which will take in the order of several days and requires the vessel to be in dry dock. The primary conflict of AAM with seismic survey use is that the sound generated from the seismic source array may produce sufficiently large signals to overload the AAM receivers when concurrently operated. Time varying gains included in the specification of SSDN, HFM3, and CMAS 36/39 will, at least partially, address the overload issue. However, as the gain varies, the detection signal may still be lost.

5.2.3. Thermal IR

The price of thermal IR systems can vary from $20,000 USD to over $200,000 USD. Cooled systems are usually much more expensive than uncooled systems and the suppliers of planar camera systems all showed wide variability in the procurement costs ($20,000 USD–$200,000 + USD). If 360° CCOS is to be achieved, several tens of planar cameras would need to be purchased, raising the costs accordingly.

With the exception of the Polaris system, all systems are in routine use, either for scientific or commercial purposes. However, for many thermal IR systems it is unclear how often they are used for scientific versus industry use, and how routinely they are used for marine mammal detection purposes. The ones that are regularly used for marine mammal detection include AIMMMS, Toyon and RADES.

Installation of thermal IR systems can usually be completed within one to two days and is described by all suppliers as being easy. Ideally, thermal IR systems (or any electromagnetic radiation based observation method) should be installed as high up as possible to achieve the maximum detection distance. As a result of this placement, there is a potential conflict with the ship's RADAR. Since navigational RADAR usually has the highest priority, thermal IR systems are usually mounted one or two decks below the RADAR. Being passive by nature,

the operation of IR systems does not interfere with any other ship systems or components.

5.2.4. RADAR

High performance systems cost $100,000 USD–$350,000 USD excluding installation and are readily available. The RADAR Technology AS system is trade restricted in accordance with EU & UN restrictions, while the NOC system is unavailable for purchase. Fitting can take from a few hours to a day or two and requires the vessel to be stationary at berth. There will be potential space conflict with traditional navigation RADAR. Installations can be either permanent or temporary. Currently no known autonomous marine mammal detection systems are available for use commercially, and developers do have target tracking detection software. All commercial systems had 360° coverage, with 10–24° vertical beam width, a variety of polarization and input power modes available and peak power of 25,000 W. Raw output, open format signals were available for connection to a custom processor but software was only available under proprietary license agreements. Systems all had Ethernet interfaces and the ability to log detections. RADAR can also provide important information on ice presence, non-marine mammal target detection including other vessels and debris, as well as information on wave size.

RADAR systems are commonly available for lease and/or purchase and are relatively easily fitted to a seismic vessel.

6. Discussion and recommendations

Effective monitoring for mitigation purposes during seismic surveys requires that animals are detected with high confidence before they enter the area where anthropogenic impacts might occur (the mitigation zone) so that required mitigation actions can be implemented. On the other hand, the frequency with which mitigation measures are activated should be kept to the minimum necessary, and hence the monitoring process should not generate many false positives. Traditionally this task has relied principally on visual observers scanning the sea surface for animals of the species of interest. Anyone who has made observations at sea knows that even under perfect detection conditions it might be possible for a marine mammal to approach a vessel, and hence enter a mitigation zone without being detected visually. Therefore, consideration of additional detection systems to increase the probability of detection and decrease the probability of false alarms is important. PAM has been increasingly used for cetacean detection in the vicinity of seismic surveys since the late 1990's. Other systems, such as RADAR, IR and AAM have not been widely used to date. We presented an overview of the concepts involved in mitigation at sea for seismic surveys, the detection systems, which systems are available and appropriate, their cost, properties and performance characteristics. Some of this information, such as cost and technical specification, is likely to become outdated quickly, but the key concepts involved will remain the same.

As with any decision process, the fact that perfection is not possible must be acknowledged. There will always be poor outcomes: false negatives and false positives, and the best we can hope for is to reduce these to acceptable levels. We recognise that there is no single system that can provide high detection probability over a wide range of species and environmental conditions. On the other hand, there are few circumstances under which most animals cannot be detected in a given time by at least one of the methods. A combination of multiple systems employing different and complimentary methods is undoubtedly the most effective way to conduct monitoring and mitigation, and the key questions then relate to determining the most cost-efficient combination of systems under a given scenario, assuming that cost and vessel space precludes the use of all monitoring methods all the time.

We recommend three areas of research to both understand and improve overall system performance for marine mammal monitoring during seismic surveys:

(1) Evaluating efficiency of individual systems.
(2) Improving the performance of individual systems.
(3) Understanding the performance of combined systems.

6.1. Evaluating efficiency of individual systems

The direct way to evaluate a system is to set up detection trials under real life conditions. Vessels could be equipped with more than one system and each different system could then be used to set up trials for each other. The number of detections (1) common to both systems, (2) from system A that were missed by B and (3) from system B missed by system A provides information about the performance of each system. This has parallels with a methodology proposed by Buckland and Turnock (1992) for dual platform line transect surveys, whereby 'tracker' observers scan far ahead of the survey vessel with powerful binoculars. Animals sighted and followed by the tracker team, which subsequently enter the area being scanned by the 'primary' observation platform, act as trials which are either detected or missed by the primary platform observers, thereby measuring the primary observer's detection efficiency (note that this is a combination of both the availability bias and the perception bias). A key problem for this approach is that detection is affected by many different factors, and consequently a large sample size is required to adequately characterise a system under the range of conditions in which it might be used. System performance is not fixed across all possible scenarios, and in particular the relative performance might be extremely dependent on a large number of extrinsic factors such as sea state, background noise, presence of interferers causing false positives, etc. Therefore, understanding the factors influencing system performance is an important part of the process, ideally with some theoretical underpinning. Further, one must assume that the probability of an animal being detected by the system being tested is independent from that of the system generating the trials. In the previous visual survey example, the tracker detects animals far in advance, so that the probability of an animal being detected by the primary team can be safely assumed to be independent of the probability of it having been detected by the tracker. Consider a deep diving species that only echolocate at depth: if a similar method were to be used to set trials for PAM using visual observers, the visual observation would need to take place far ahead of the survey platform such that sufficient time passed before the acoustic system was in range. Otherwise the strong negative correlation in availability for detection between the two systems would be a confounding factor. This protocol would be impractical for animals whose dive cycles may last for over an hour, since it would be hard to visually survey far enough ahead and keep track of the animal to evaluate its successful, or missed detection.

For situations where it is possible to consider dual platform methods, trials would need to be conducted in a high density area for the target species to achieve adequate sample sizes in a realistic time frame. One has to be aware that the results may only hold for the specific system configurations tested, are only valid for the species investigated (or those species with comparable animal dependent extrinsic factors), and for the environmental conditions during the experiments. Many months of data collection may be required to obtain the sample sizes required to make reliable inferences covering the variety of extrinsic and intrinsic factors that might potentially influence the detection performance of the tested systems. In addition, an understanding of some aspects of the performance of individual systems in a range of environmental conditions could be gained by trials using artificial cues. Also, the effects of noise can be explored by mixing high signal-to-noise data with different levels of noise and determining how this affects the performance of software detectors.

6.2. Improving individual systems

All systems will perform better if they are optimised for a specific detection scenario based on the species of interest and local noise conditions. For PAM systems, their frequency range needs to cover the frequency spectrum of the vocalisation and detectors may be tuned for specific species of local importance. Most important for PAM is reducing acoustic noise, which is often determined by the way in which the hydrophone arrays are deployed from the monitoring platform. Furthermore, internal factors such as system noise or operational sound should be minimised. For AAM, the source level of the outgoing sonar pulses, their type and frequency should be adapted to the size of the target species. Receiver beam width, spatial coverage, steerability and stabilisation as well as the maximum operational depth influences the detection probability. The system resolution is also important for RADAR systems, as well as their power, scan rates and antenna type and height. These should all be adapted to the project specific purpose of the monitoring. With regard to internal factors influencing detection, IR systems should have a good thermal resolution, and low background noise level combined with high concurrent ocean coverage, while, for RADAR, polarimetric antenna and filtering raises the detection abilities in sub-optimal conditions.

Evaluations of the systems as described above will provide a framework for assessing the effectiveness of improvements. Work to improve system performance is primarily the responsibility of system developers. Collaboration between the oil and gas industry and the developers wishing to test/quantify the performance of improved systems should be encouraged.

6.3. Predicting the performance of combined systems

The use of multiple complimentary monitoring methods that complement each other's detection performance is a natural and necessary way of increasing the overall detection efficiency. As an example, methods for detecting animals at the surface will be well complemented by a method able to detect submerged animals, such as passive (for sound producing animals) or active acoustics.

It is clear that, given the magnitude of the problems of evaluating individual systems via field trials, attempting this for all the possible combinations of methods is a near impossible task. We suggest that a more effective way to understand the overall efficiency of combined systems would be to simulate combined detector performance, with parameters describing the performance of individual systems informed by real life trials in a model that includes realistic information on patterns of cue production and movements for the marine mammal species concerned. This would allow the prediction of the performance of different combinations of monitoring methods in a range of environmental conditions, for different species, and how well these would meet the requirements of particular regulations and guidelines. We recommend that a computer simulation tool should be developed, which would use the best available knowledge of animal movement, behaviour and cue production (surfacings, blows, vocalisations, etc.), and the most up to date information on the performance of individual detection systems, to simulate realistic scenarios in which combined performance would be assessed. Of course the utility of such a tool will be dependent on the data underpinning it: how reliably it represents the characteristics of a given scenario and the performance of individual systems under that scenario, so considerable work is required to obtain baseline data to inform the simulation before it can be used in practice. Similar simulation tools are used for example by navies to inform the likelihood of real life operations impacting marine mammals (e.g. Siderius and Porter, 2006), however, they only consider the performance of single systems.

Both literature reviews and dedicated field studies would be needed to provide the information needed for such a simulation, including:

• Studies that provide detailed information on temporal patterns and strength of relevant cues. Combined cue pattern data are required (e.g. combined time series for every blow and every click made by a sperm whale). This is fundamental because to evaluate the

performance of combined systems of complementary methods it is key to understand the correlation structure between the patterns of availability of the different cues that could be detected by the systems being tested. If one assumes that two systems detect animals independently, and that availability patterns are independent, then the probability of detecting an animal is the complement of missing the species in both systems, i.e. one minus the product of the two systems' probabilities of missing the animal. In the presence of a negative correlation in availability for the two systems, i.e. if at times when an animal is less available to be detected by one system it is more available to the other, the animals will actually be available for detection for a greater proportion of the time than the independence assumption would imply. Conversely, if the correlation between the availability processes is positive, the probability of detection will be lower than what would be the case for independent availability. Thus, the correlation structure of availability patterns must be accounted for to understand the overall performance of combined systems.

- Studies of the performance of individual systems at detecting individual cues under different conditions (e.g. different sea states, different fog conditions, different ambient noise, etc.).

Individual (or agent) based models (e.g. Grimm and Railsback, 2005) is one framework that might be used to develop the above simulation. This would require placing virtual animals, or groups of animals, in the area of interest, each moving and producing cues in a realistic manner. Other required parameters would include the size of the mitigation zone, the operational settings of a seismic survey, such as line change, survey scale and duration, the sound characteristics and movement of the sound source and the performance of different detection systems in different environmental conditions. Multiple detection systems could be used alone or in different combinations. By running a large number of simulations, probabilities for the four monitoring outcomes outlined in Section 3.2 could be estimated.

The simulation could be run with the species of interest in a particular area, the likely environmental conditions and the monitoring requirements (from existing regulations) using various combinations of sensors to evaluate likely cost performance parameters for different monitoring procedures. This would provide stakeholders with a realistic assessment of how effectively monitoring for mitigation efforts would achieve its goals. It would be very desirable to ground truth simulations by comparing predictions with real world observations made during seismic monitoring exercises.

A key requirement for developing useful models is to have realistic information on animal behaviour, in particular with regard to their cue production patterns and movements. Ideally these data should describe the behaviour that would occur before and during the activities that require mitigation measures. There are few examples of detailed datasets of this type and the problem is made considerably more complicated because the requirement is to know how animals move and behave in the presence of a seismic vessel both with and without an operating seismic sound source. Seismic surveys affect diving behaviour and movements of marine mammals (e.g. Gordon et al., 2003) and Blackwell et al. (2015) showed that bowhead whales change their vocalisation rates at different ranges from seismic surveys. These examples highlight the need for high quality data collection during seismic surveys. Some useful data could be collected with minimal additional effort as part of current observer duties by introducing appropriate monitoring protocols (for example, recording detailed cue production rates from sightings). Many detection systems are capable of recording raw data, which can be further processed offline not only to check for detections, but also to assess signal to noise levels and false positive rates. Other data, such as swimming behaviour and surfacing rate, would require dedicated field studies involving focal follows and recording appropriate cue parameters.

7. Conclusions

(1) PAM, AAM, thermal IR systems, and potentially high performance RADAR offer monitoring tools for the detection of marine mammals at sea which could usefully supplement visual observer effort. None of these methods on their own is likely to provide in-time detection of all animals or species in all conditions during real-time monitoring for mitigation purposes for seismic surveys. Combinations of two or more techniques may be necessary to provide the high level of detection required.

(2) Thermal IR and RADAR systems (like visual observers) detect cues made at and above the surface. Animals cannot be detected by these systems when they are submerged. Acoustic methods, such as PAM and AAM complement these efforts by enabling detection of animals when they are underwater.

(3) Passive acoustics is clearly a key modality for detecting many marine mammal species (mainly cetaceans) underwater. The extent to which PAM could be useful for detecting marine mammals for low visibility real-time monitoring for mitigation purposes varies considerably between species and with applications, being influenced in particular by the vocal behaviour of particular species (which may vary with time of year, location and gender), how these sounds propagate in the environment and the total noise field in which detections must be made. The method works well with most odontocete species, although the detection range for some species vocalising at very high frequencies, such as porpoise, may still remain too short to cover the monitoring zones required by many regulations.

(4) Vessel-mounted lower frequency (below 50 kHz) AAM systems have been demonstrated to be able to detect large marine mammals such as large odontocetes, pinnipeds and mysticetes at the ranges required by the industry for real-time monitoring for mitigation purposes. Detection ranges for small marine mammals may, however, not be sufficient to cover the required monitoring zones. Localisation and tracking is straightforward with AAM systems, but animal classification to either species or a species group is not possible. An animal must provide sufficient reflectivity to enable an adequate echo, and while the target strength has been measured and modelled for some species, for most species it is unknown. The potential for additional impact of the acoustic emission of an AAM system on marine mammals would need to be assessed.

(5) Thermal IR whale detection works best with short-diving, large animals, and worst with long-diving elusive and small animals. A 360° detection of animals is possible. Due to decreased noise from sunlight reflections, automatic detection of whale signatures in thermal IR works better during night than during day, rendering it ideal for the most common low visibility conditions (low light or darkness) and it is also quite robust to the effects of sea state. To date, thermal IR whale detection has mainly been performed in areas with cold to moderate water temperatures. In these waters performance measures (detection probability for different distances, true and false positive ratios) for large whales are adequate for real time monitoring. Detection ranges for tropical regions and small marine mammals are largely unknown.

(6) RADAR detection works best with short-diving animals with large and long above water expressions or surface activity and can also detect animals on ice. A 360° detection of animals is possible and large whales can, in ideal sea conditions, be detected at the ranges required by the industry for real-time monitoring for mitigation purposes (~ 1 km). However, a standard marine RADAR is unlikely to provide adequate monitoring performance for mitigation across a range of species and sea states. High performance vessel-mounted RADARs and polarimetric antennas (coupled with more sophisticated detection and clutter reducing software) are reported by system developers to perform better in high sea states, fog and rain than the standard marine RADAR. However, no empirical detection

reliability data is currently available, particularly to determine false positive rates and cue masking, which are of particular concern in higher sea states. The utility of proprietary target detection software is also unknown.

(7) Most systems use software to automatically detect cues. These all rely on human verification of potential targets to reduce false positive errors.

(8) Classification of marine mammals down to species level can currently only be reliably achieved to the level required for mitigation using PAM; though with AAM and IR, body size, body shape and movement/behavioural pattern of the animals should enable classification to species groups (such as seal, large whale or dolphin). Species or size classification using RADAR is unlikely.

(9) The performance (for a variety of species and over a range of conditions) of monitoring systems that utilise more than one detection methodology can best be explored in a modelling framework that incorporates real world data on the performance of the individual systems and the behaviour of target species. A model of this sort could be used to identify combinations of methods that would provide the best performance and to direct efforts to improve methods. Such a model could also help to develop better monitoring regulations and improved strategies for directing available monitoring effort in space and time. However, a model can only ever be as good as the data on which it is based and substantial new data will be required on cue production rates and movements of species of interest and on the performance of systems.

Acknowledgements

References

Abadi, S.H., Tolstoy, M., Wilcock, W.S., 2015. Baleen whale localization using a dual-line towed hydrophone array during seismic reflection surveys. J. Acoust. Soc. Am. 138, 1762.

ACCOBAMS Scientific Committee, 2004. Resolution 2.16 assessment and impact assessment of man-made noise. In: The Meeting of the Parties to the Agreement on the Conservation of Cetaceans of the Black Sea, Mediterranean Sea and Contiguous Atlantic Area.

Anderson, S.J., Morris, J.T., 2010. On the detection of marine mammals with ship-borne polarimetric microwave radar. IEEE 1–6.

Au, W.W., 2012. The Sonar of Dolphins. Springer Science & Business Media.

Au, W.W., Kastelein, R.A., Rippe, T., Schooneman, N.M., 1999. Transmission beam pattern and echolocation signals of a harbor porpoise (*Phocoena phocoena*). J. Acoust. Soc. Am. 106, 3699–3705.

Au, W.W., Ford, J.K., Horne, J.K., Allman, K.A.N., 2004. Echolocation signals of free-ranging killer whales (*Orcinus orca*) and modeling of foraging for chinook salmon (*Oncorhynchus tshawytscha*). J. Acoust. Soc. Am. 115, 901–909.

Blackwell, S.B., Nations, C.S., McDonald, T.L., Thode, A.M., Mathias, D., Kim, K.H., Greene Jr., C.R., Macrander, A.M., 2015. Effects of airgun sounds on bowhead whale calling rates: evidence for two behavioral thresholds. PLoS One 10, e0125720.

Borchers, D.L., Buckland, S.T., Zucchini, W., 2002. Estimating Animal Abundance: Closed Populations. Springer Science & Business Media.

Borowski, G.H., 1781. Gemeinnüzzige Naturschichte des Tierreichs. G. L. Lange, Berlin and Stralsund. 2 (1), 21.

Briggs, J.N., 2004. Target Detection by Marine Radar, 16th ed. IET.

Buckland, S.T., Turnock, B.J., 1992. A robust line transect method. Biometrics 48, 901–909.

Buckland, S.T., Anderson, D.R., Burnham, K.P., Laake, J.L., Borchers, D.L., Thomas, L., 2001. Introduction to Distance Sampling. Estimating Abundance of Biological Populations. Oxford University Press, Oxford.

Buckland, S.T., Anderson, D.R., Burnham, K.P., Laake, J.L., Borchers, D.L., Thomas, L., 2004. Advanced Distance Sampling. Estimating Abundance of Biological Populations. Oxford University Press, Oxford.

Buckland, S.T., Rexstad, E.A., Marques, T.A., Oedekoven, C.S., 2015. Distance Sampling: Methods and Applications. Springer.

Clarke, R., 1982. An index of sighting conditions for surveys of whales and dolphins. In: Report of the International Whaling Commission. 32. pp. 559–561.

Cucknell, A.-C., Boisseau, O., Leaper, R., Mclanaghan, R., Moscrop, A., 2016. Harbour porpoise (*Phocoena phocoena*) presence, abundance and distribution over the Dogger Bank, North Sea, in winter. J. Mar. Biol. Assoc. U. K. 1–11.

Current Corporation, 2011. Night Navigator. pp. 3.

DeProspo, D., Mobley, J., Hom, W., Carron, M., 2005. Radar-Based Detection, Tracking and Speciation of Marine Mammals from Ships. Arete Associates, Arlington VA.

DeRuiter, S.L., Southall, B.L., Calambokidis, J., Zimmer, W.M., Sadykova, D., Falcone, E.A., Friedlaender, A.S., Joseph, J.E., Moretti, D., Schorr, G.S., 2013. First direct measurements of behavioural responses by Cuvier's beaked whales to mid-frequency active sonar. Biol. Lett. 9, 20130223.

DEWHA, 2008. Australian Government EPBC Act Policy Statement 2.1 - interaction between offshore seismic exploration and whales. In: The Department of the Environment Water Heritage and the Arts.

Dunlop, R.A., 2016. Changes in vocal parameters with social context in humpback whales: considering the effect of bystanders. Behav. Ecol. Sociobiol. 70, 857–870.

Erbe, C., 2002. Underwater noise of whale-watching boats and potential effects on killer whales (*Orcinus orca*), based on an acoustic impact model. Mar. Mamm. Sci. 18, 394–418.

Forsyth, C.P., 2011. Radar Detection of Marine Mammals.

Gillespie, D., Mellinger, D., Gordon, J., McLaren, D., Redmond, P., McHugh, R., Trinder, P., Deng, X., Thode, A., 2008. PAMGUARD: semiautomated, open source software for real-time acoustic detection and localisation of cetaceans. J. Acoust. Soc. Am. 30, 54–62.

Gordon, J., Gillespie, D., Chappell, O., Lewis, T., Swift, R., Belford, R., 2000. The Role of Acoustic Monitoring in Minimising the Impact of Seismic Aquisition on Cetaceans, 62nd EAGE Conference & Exhibition.

Gordon, J., Gillespie, D., Potter, J., Frantzis, A., Simmonds, M.P., Swift, R., Thompson, D., 2003. A review of the effects of seismic surveys on marine mammals. Mar. Technol. Soc. J. 37, 16–34.

Grimm, V., Railsback, S.F., 2005. Individual-Based Modeling and Ecology. Princeton University Press. Princeton, New Jersey, USA.

Harwood, L.A., Joynt, A., 2009. Factors influencing the effectiveness of Marine Mammal Observers on seismic vessels, with examples from the Canadian Beaufort Sea. In: Canadian Science Advisory Secretariat = Secrétariat Canadien de Consultation Scientifique.

Hastie, G., 2013. Tracking Marine Mammals Around Marine Renewable Energy Devices Using Active Sonar. (SMRU Ltd. report to the Department of Energy and Climate Change).

Hastie, G.D., Donovan, C., Götz, T., Janik, V.M., 2014. Behavioral responses by grey seals (*Halichoerus grypus*) to high frequency sonar. Mar. Pollut. Bull. 79, 205–210.

IAGC, 2011. Recommended Mitigation Measures for Cetaceans During Geophysical Operations.

Janik, V.M., 2009. Acoustic communication in delphinids. In: Advances in the Study of Behavior. 40. pp. 123–157.

Janik, V.M., Sayigh, L.S., 2013. Communication in bottlenose dolphins: 50 years of signature whistle research. J. Comp. Physiol. A. 199, 479–489.

JNCC, 2010a. JNCC Guidelines for Minimising the Risk of Injury and Disturbance to Marine Mammals from Seismic Surveys.

JNCC, 2010b. JNCC Guidelines for Minimising the Risk of Injury to Marine Mammals from Using Explosives.

JNCC, 2010c. Statutory Nature Conservation Agency Protocol for Minimising the Risk of Injury to Marine Mammals from Piling Noise.

JNCC, 2017. JNCC Guidelines for Minimising the Risk of Injury to Marine Mammals from Geophysical Surveys.

Ketten, D.R., 1995. Estimates of blast injury and acoustic trauma zones for marine mammals from underwater explosions. In: Sensory Systems of Aquatic Mammals, pp. 391–407.

Leaper, R., Gordon, J., 2001. Application of photogrammetric methods for locating and tracking cetacean movements at sea. J. Cetacean Res. Manag. 3, 11.

Leong, H., Ponsford, A., 2008. The effects of sea clutter on the performance of HF surface wave radar in ship detection. IEEE 1–6.

Lewis, T., Gillespie, D., Lacey, C., Matthews, J., Danbolt, M., Leaper, R., McLanaghan, R., Moscrop, A., 2007. Sperm whale abundance estimates from acoustic surveys of the Ionian Sea and Straits of Sicily in 2003. J. Mar. Biol. Assoc. U. K. 87, 353–357.

Lilljeborg, 1861. Forh. Skand. Naturf. Ottende Møde, Kopenhagen (1860). 8, 602.

Linnaeus, C., 1758. Systema Naturae per regna tria naturae, secundum classes, ordines, genera, species, cum characteribus, differentiis, synonymis, locis. Editio decima, reformata. Laurentius Salvius, Holmiae (ii, 824 pp).

Madsen, P., Wahlberg, M., Mohl, B., 2002. Male sperm whale (*Physeter macrocephalus*) acoustics in a high-latitude habitat: implications for echolocation and communication. Behav. Ecol. Sociobiol. 53, 31–41.

Madsen, P., Johnson, M., De Soto, N.A., Zimmer, W., Tyack, P., 2005. Biosonar performance of foraging beaked whales (*Mesoplodon densirostris*). J. Exp. Biol. 208, 181–194.

Marsh, H., Sinclair, D.F., 1989. Correcting for visibility bias in strip transect aerial surveys of aquatic fauna. J. Wildl. Manag. 1017–1024.

Martin, N., St John, K., Gill, G., 2014. Review of Sound and Marine Life Guidelines for Marine Seismic Operations. (This paper was prepared for presentation at the 2014 SPE International Conference on Health, Safety, and Environment in Oil and Gas Exploration and Production held in Long Beach, California, 17-19 March 2014).

Matthews, J., Brown, S., Gillespie, D., Johnson, M., McLanaghan, R., Moscrop, A., Nowacek, D., Leaper, R., Lewis, T., Tyack, P., 2001. Vocalisation rates of the North Atlantic right whale (*Eubalaena glacialis*). J. Cetacean Res. Manag. 3, 271–282.

McDonald, M.A., Mesnick, S.L., Hildebrand, J.A., 2006. Biogeographic characterization of blue whale song worldwide: using song to identify populations. J. Cetacean Res. Manag. 8.

Mohl, B., Andersen, S., 1973. Echolocation: high frequency component in the click of the Harbour Porpoise (*Phocoena ph. L*). J. Acoust. Soc. Am. 54, 1368–1372.

Mohl, B., Wahlberg, M., Madsen, P.T., Heerfordt, A., Lund, A., 2003. The monopulsed nature of sperm whale clicks. J. Acoust. Soc. Am. 114, 1143–1154.

Norris, T.F., Mc Donald, M., Barlow, J., 1999. Acoustic detections of singing humpback whales (*Megaptera novaeangliae*) in the eastern North Pacific during their northbound migration. J. Acoust. Soc. Am. 106, 506–514.

Nowacek, D.P., Southall, B., 2016. In: IUCN (Ed.), Effective Planning Strategies for Managing Environmental Risk Associated with Geophysical and Other Imaging Surveys: A Resource Guide for Managers, pp. 42.

OGP, IAGC, 2011. An Overview of Marine Seismic Operations.

Palka, D., 1996. Effects of Beaufort sea state on the sightability of harbour porpoises in the Gulf of Maine. In: Report of the International Whaling Commission. 46. pp. 575–582.

Parente, C.L., de Araujo, M.E., 2011. Effectiveness of monitoring marine mammals during marine seismic surveys off Northeast Brazil. RGCI-Revista de Gestao Costeira Integrada 11.

Payne, R., Webb, D., 1971. Orientation by means of long range acoustic signaling in baleen whales. Ann. N. Y. Acad. Sci. 188, 110–141.

Pirotta, E., Brookes, K.L., Graham, I.M., Thompson, P.M., 2014. Variation in harbour porpoise activity in response to seismic survey noise. Biol. Lett. 10, 20131090.

Pyć, C.D., Geoffroy, M., Knudsen, F.R., 2016. An evaluation of active acoustic methods for detection of marine mammals in the Canadian Beaufort Sea. Mar. Mamm. Sci. 32, 202–219.

Quick, N.J., Janik, V.M., 2012. Bottlenose dolphins exchange signature whistles when meeting at sea. Proc. R. Soc. B Biol. Sci. 279, 2539–2545.

Schulz, T.M., Whitehead, H., Gero, S., Rendell, L., 2008. Overlapping and matching of codas in vocal interactions between sperm whales: insights into communication function. Anim. Behav. 76, 1977–1988.

Seber, G.A.F., 1986. A Review of estimating animal abundance. Biometrics 42, 267–292.

Sheldon-Robert, M., Beerens, S., Lam, F., 2008. The Delphinus array for passive marine mammal detection. In: 3rd Annual Maritime Systems and Technology Global Conference-MAST 2008, 12-14 November 2008. Cadiz, Spain.

Siderius, M., Porter, M.B., 2006. Modeling techniques for marine-mammal risk assessment. IEEE J. Ocean. Eng. 31, 49–60.

Silber, G., Betteridge, S., Cottingham, D., 2009. Report of a workshop to identify and assess technologies to reduce ship strikes of large whales, providence, Rhode Island 8-10 July 2008. In: NOAA Technical Memorandum NMFS-OPR-42.

Sirovic, A., Hildebrand, J.A., Wiggins, S.M., 2007. Blue and fin whale call source levels and propagation range in the Southern Ocean. J. Acoust. Soc. Am. 122, 1208–1215.

Smith, J.N., Goldizen, A.W., Dunlop, R.A., Noad, M.J., 2008. Songs of male humpback whales, *Megaptera novaeangliae*, are involved in intersexual interactions. Anim. Behav. 76, 467–477.

Southall, B.L., Bowles, A.E., Ellison, W.T., Finneran, J.J., Gentry, R.L., Greene, C.R.J., Kastak, D., Ketten, D.R., Miller, J.H., Nachtigall, P.E., Richardson, W.J., Thomas, J.A., Tyack, P.L., 2007. Marine mammal noise exposure criteria: initial scientific recommendations. Aquat. Mamm. 33, 411–414.

Stafford, K.M., Fox, C.G., Clark, D.S., 1998. Long-range acoustic detection and localization of blue whale calls in the northeast Pacific Ocean. J. Acoust. Soc. Am. 104, 3616–3625.

Stone, C., 2015. Marine mammal observations during seismic surveys from 1994 to 2010. In: JNCC Report, No. 463a.

Thomas, L., Buckland, S.T., Rexstad, E.A., Laake, J.L., Strindberg, S., Hedley, S.L., Bishop, J.R.B., Marques, T.A., Burnham, K.P., 2010. Distance software: design and analysis of distance sampling surveys for estimating population size. J. Appl. Ecol. 47, 5–14.

Urick, R.J., 1984. Ambient Noise in the Sea. Catholic University of America, Washington DC.

Verfuss, U.K., Miller, L.A., Schnitzler, H.-U., 2005. Spatial orientation in echolocating harbour porpoises (*Phocoena phocoena*). J. Exp. Biol. 208, 3385–3394.

Verfuss, U.K., Miller, L.A., Pilz, P.K., Schnitzler, H.-U., 2009. Echolocation by two foraging harbour porpoises (*Phocoena phocoena*). J. Exp. Biol. 212, 823–834.

Verfuss, U., Gillespie, D., Gordon, J., Marques, T., Miller, B., Plunkett, R., Theriault, J., Tollit, D., Zitterbart, D., Hubert, P., Thomas, L., 2016. Low Visibility Real-Time Monitoring Techniques Review.

von Benda-Beckmann, A.M., Lam, F.P.A., Moretti, D.J., Fulkerson, K., Ainslie, M.A., van IJsselmuide, S.P., Theriault, J., Beerens, S.P., 2010. Detection of Blainville's beaked whales with towed arrays. Appl. Acoust. 71, 1027–1035.

Weir, C.R., Dolman, S.J., 2007. Comparative review of the regional marine mammal mitigation guidelines implemented during industrial seismic surveys, and guidance towards a worldwide standard. J. Int. Wildl. Law and Policy 10, 1–27.

Weissenberger, J., Zitterbart, D.P., 2012. Surveillance of Marine Mammals in the Safety Zone around an Air Gun Array with the Help of a 360 Infrared Camera System. Society of Petroleum Engineers.

Zitterbart, D.P., Kindermann, L., Burkhardt, E., Boebel, O., 2013. Automatic round-the-clock detection of whales for mitigation from underwater noise impacts. PLoS 8, 1–6.