582
2008

Berichte
zur Polar- und Meeresforschung

Reports on Polar and Marine Research

Automated passive acoustic detection, localization and identification of leopard seals: from hydro-acoustic technology to leopard seal ecology

Automatisierte, passiv-akustische Detektion, Lokalisation und Identifikation von Seeleoparden: von hydro-akustischer Technologie zur Ökologie des Seeleoparden

Holger Klinck

HELMHOLTZ GEMEINSCHAFT

ALFRED-WEGENER-INSTITUT FÜR
POLAR- UND MEERESFORSCHUNG
In der Helmholtz-Gemeinschaft
D-27570 BREMERHAVEN
Bundesrepublik Deutschland

ISSN 1866-3192

Hinweis

Die Berichte zur Polar- und Meeresforschung werden vom Alfred-Wegener-Institut für Polar-und Meeresforschung in Bremerhaven* in unregelmäßiger Abfolge herausgegeben.

Sie enthalten Beschreibungen und Ergebnisse der vom Institut (AWI) oder mit seiner Unterstützung durchgeführten Forschungsarbeiten in den Polargebieten und in den Meeren.

Es werden veröffentlicht:

- Expeditionsberichte (inkl. Stationslisten und Routenkarten)
- Expeditionsergebnisse (inkl. Dissertationen)
- wissenschaftliche Ergebnisse der Antarktis-Stationen und anderer Forschungs-Stationen des AWI
- Berichte wissenschaftlicher Tagungen

Die Beiträge geben nicht notwendigerweise die Auffassung des Instituts wieder.

Notice

The Reports on Polar and Marine Research are issued by the Alfred Wegener Institute for Polar and Marine Research in Bremerhaven*, Federal Republic of Germany. They appear in irregular intervals.

They contain descriptions and results of investigations in polar regions and in the seas either conducted by the Institute (AWI) or with its support.

The following items are published:

- expedition reports (incl. station lists and route maps)
- expedition results (incl. Ph.D. theses)
- scientific results of the Antarctic stations and of other AWI research stations
- reports on scientific meetings

The papers contained in the Reports do not necessarily reflect the opinion of the Institute.

The „Berichte zur Polar- und Meeresforschung” continue the former „Berichte zur Polarforschung”

* Anschrift / Address
Alfred-Wegener-Institut
Für Polar- und Meeresforschung
D-27570 Bremerhaven
Germany
www.awi.de

Editor in charge:
Dr. Horst Bornemann

Assistant editor:
Birgit Chiaventone

Die "Berichte zur Polar- und Meeresforschung" (ISSN 1866-3192) werden ab 2008 ausschließlich als Open-Access-Publikation herausgegeben (URL: http://epic.awi.de).

Since 2008 the "Reports on Polar and Marine Research" (ISSN 1866-3192) are only available as web based open-access-publications (URL: http://epic.awi.de)

Automated passive acoustic detection, localization and identification of leopard seals: from hydro-acoustic technology to leopard seal ecology

Automatisierte, passiv-akustische Detektion, Lokalisation und Identifikation von Seeleoparden: von hydro-akustischer Technologie zur Ökologie des Seeleoparden

---

Holger Klinck

ISSN 1866-3192

„Erfolg ist die Fähigkeit, von einem Misserfolg zum anderen zu gehen, ohne seine Begeisterung zu verlieren.“

(Winston Churchill)

Diese Arbeit ist gewidmet

Hedi Simon	&	**Adam Simon**
* 21.03.1918		* 22.12.1914
† 11.01.2002		† 12.08.1998

Impressum

Holger Klinck
Alfred Wegener Institute for Polar and Marine Research
Am Handelshafen 12
27570 Bremerhaven, Germany

✉
Holger Klinck
Rudolf-Kinau-Strasse 10
24340 Eckernförde, Germany

E-mail: holger@klinck.net

Die vorliegende Arbeit ist die inhaltlich unveränderte Fassung einer Dissertation, die im Januar 2008 dem Fachbereich VI (Geographie/Geowissenschaften) der Universität Trier zur Erlangung des akademischen Grades Doktor der Naturwissenschaften (Dr. rer. nat.) vorgelegt wurde.

Table of contents

Zusammenfassung

Im Rahmen dieser Dissertation wurde ein hydro-akustisches Observatorium - das Perennial Acoustic Observatory in the Antarctic Ocean (PALAOA) - entwickelt, in der Antarktis aufgebaut und seit Januar 2006 quasi kontinuierlich betrieben. Seit Inbetriebnahme des Observatoriums wurden insgesamt 9400 Stunden (4.5 TByte) an akustischen Daten, die verschiedenste biotische und abiotische Signale beinhalten, aufgezeichnet. Das Observatorium verfügt über eine Echtzeitdatenanbindung an das Alfred-Wegener-Institut für Polar- und Meeresforschung (AWI) in Bremerhaven, wodurch es möglich ist, die akustischen Daten in Echtzeit zu analysieren.

Um die enormen Datenmengen nach definierten Signalen und in einem akzeptablen Zeitraum durchsuchen zu können, ist die Entwicklung von automatisierten Detektionsalgorithmen unabdingbar. In dieser Arbeit werden erstmalig automatische Algorithmen zur Erkennung von Robbenvokalisationen vorgestellt. Zwei komplementäre Algorithmen zur Detektion des low (LDT) und high (HDT) double trill Rufes des Seeleoparden (*Hydrurga leptonyx*) wurden auf Basis der Nulldurchgangsrate (zero-crossing rate - ZCR) bzw. eines Hidden Markov Models (HMM), beides Standardmethoden in der menschlichen Spracherkennung, entwickelt. Die Effizienz beider Detektoren wurde anhand eines einwöchigen Testdatensatzes ermittelt. Wie diese Validierung zeigte, erlauben die Algorithmen eine schnelle und robuste automatische Detektion der Signale. Die Detektion der LDTs war bei einem „Signal zu Rausch Abstand“ (SNR) > 13 dB zu 74 % erfolgreich. Der auf dem Hidden Markov Model basierende Algorithmus zur Detektion der HDTs erreichte bei Signalen mit einem SNR > 14 dB eine Detektionsrate von 85 %.

Beide Algorithmen wurden nach ausführlichen Tests auf die PALAOA Langzeitdatensätze appliziert. Die Resultate wurden hinsichtlich zeitlicher und saisonaler Muster im Vokalisationsverhalten des Seeleoparden untersucht. Die Analysen zeigten, dass Seeleoparden im Zeitraum Ende September bis Mitte Januar an der PALAOA Station akustisch präsent sind. Vor dieser Studie war lediglich bekannt, dass freilebende Seeleoparden im Zeitraum Anfang November bis Mitte Januar vokalisieren. Auch konnten zeitliche (24-stündige) Muster (sog. diurnal pattern) identifiziert werden, die mit hoher Wahrscheinlichkeit durch das „haul-out“ Verhalten (Ruhephasen auf dem Eis) der Tiere hervorgerufen werden. Die höchsten Vokalisationsraten treten demnach in der Nacht auf, die geringsten Raten am späten Vormittag, wenn die Sonneneinstrahlung am höchsten ist und die Tiere auf den Eisschollen liegen.

Darüber hinaus wurden mit Hilfe von verschiedenen Cluster- bzw. Klassifikationsverfahren (wie zum Beispiel Neuronale Netze) die Vokalisationen der Seeleoparden hinsichtlich geographischer Dialekte und individueller Unterschiede untersucht. Die durchgeführten Analysen zeigten beispielsweise, dass Individuen anhand ihrer Vokalisationen identifiziert werden können. Mit Hilfe eines hierarchischen Clusterverfahrens konnten 105 Vokalisationen eines Testdatensatzes, der von insgesamt drei Seeleoparden stammt, eindeutig den jeweiligen Individuen zugeordnet werden.

Zusammenfassend kann festgehalten werden, dass im Rahmen dieser Dissertation ein technisch neuartiges hydro-akustisches Observatorium entwickelt und erfolgreich in einer der abgeschiedensten Regionen unseres Planeten aufgebaut und betrieben wurde bzw. wird. Darüber hinaus wurden automatische Detektions- und Identifikationsalgorithmen entwickelt, die es erlauben, die ständig wachsende Datenmenge systematisch und automatisiert zu analysieren. Zukünftige biologisch orientierte Projekte werden unter Zuhilfenahme der entwickelten Technologien und Methoden vielfältige Fragestellungen auf dem Gebiet der Ökologie antarktischer mariner Warmblüter untersuchen können.

Summary

In the course of this study the Perennial Acoustic Observatory in the Antarctic Ocean (PALAOA) was developed, installed and operated (quasi-)continuously since January 2006 on the Antarctic ice shelf. The autonomous observatory provides real-time access to the Antarctic underwater soundscape and provides an acoustic data set of more than 9400 hours (4.5 TByte) of various biotic and abiotic sounds remotely from the Alfred Wegener Institute for Polar and Marine Research (AWI), Bremerhaven, Germany.

To screen the huge data set for signals of interest within a reasonable amount of time, automated pattern recognition algorithms are essential. For the first time, automated detectors were developed in this study for the detection of pinniped vocalizations. The two complementary algorithms focus on the detection of the predominant vocalizations of the leopard seal (*Hydrurga leptonyx*): the low (LDT) and high (HDT) double trill. Both algorithms are based on techniques commonly used in human speech recognition. The LDT detector is based on the signals' zero-crossing rate (ZCR) while the detector for the HDT utilizes the temporal changes in the pulse repetition rate (PRR) of the amplitude modulated signal in combination with a Hidden Markov Model (HMM) for the detection. For both detectors a performance test was conducted using a reference data set of one week duration. The performance test revealed that each detector allows a fast and reliable detection of its respective leopard seal call type. The detection rates of the detectors were estimated at 74% for LDTs (signal-to-noise ratio (SNR) > 13 dB) and 85% for HDTs (SNR > 14 dB). Once constructed and validated, the detectors were applied to large data sets to extract information on seasonal and temporal patterns of the vocal behavior of leopard seals. One result of this analysis revealed that leopard seal vocalizations are present in the vicinity of PALAOA between late September and mid-January. Before this study, free-living leopard seals were known to be vocal only between November and mid-January. Furthermore a diurnal pattern in the vocal behavior of leopard seals with high calling rates during night time and low calling rates during late forenoon was found, and most likely reflects the haul-out (resting on the ice) behavior of the seals.

State of the art, data driven methods such as hierarchical cluster analysis and neural networks were tested for their suitability to analyze geographical and individual variations of the HDT. The performance of one of these approaches, the hierarchical cluster analysis, was tested by analyzing data sets from three known individuals. The system correctly allocated all 105 seal vocalizations of a test set to the corresponding seal.

Overall, this study successfully built a reliable tool to obtain acoustic underwater recordings for one of the most remote regions on this planet and developed automated detection and identification tools for systematic analysis of the resulting, continuously growing, large acoustic data set. Future studies can now build on the ever growing data set and use the developed methods to investigate the many questions of pinnipeds and cetacean ecology.

1. Introduction

1.1 Motivation

The Southern Ocean[1.1] provides a nearly pristine habitat to a great variety of marine mammals, birds, fishes, invertebrates, micro-organisms and plants, many of which are endemic to this region. The investigation of these species, however, is seriously hindered by this region's remoteness and extreme environmental conditions, as manifested particularly in the formation of sea ice, low temperatures, and lack of daylight during polar winter.

The sea ice around Antarctica covers - at its maximum during polar winter - approximately 20 million km^2 (see Figure 1.1). During polar summer, when the solar radiation reaches maximum intensity, the ice covered area shrinks to 4 million km^2 (Kaiser *et al.* 2005). Antarctica's air temperature varies from a few degrees above freezing during austral mid-summer to temperatures below -50°C during austral winter. Light conditions alternate between continuous daylight during austral mid-summer and complete darkness for up to five months (depending on the latitude) during austral winter (El-Sayed 1971).

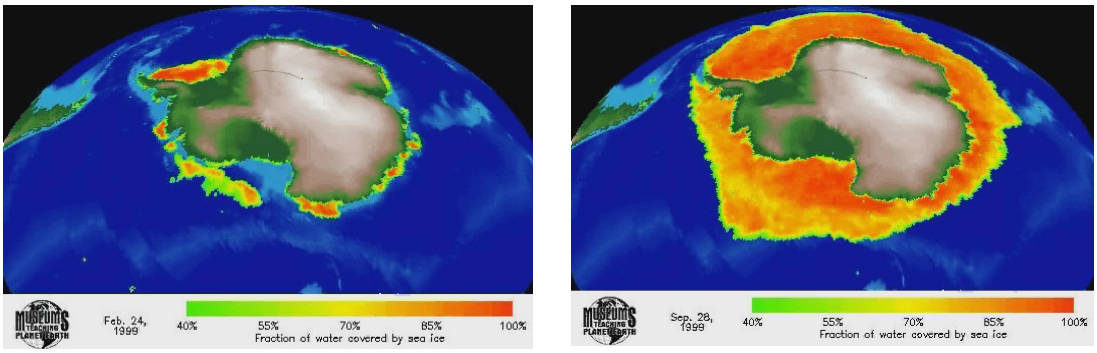


Figure 1.1: Sea ice concentration around Antarctica in austral summer (end of February – left side) and austral winter (end of September – right side) - derived by satellite imagery.
Source: http://earth.rice.edu/mtpe/cryo/cryosphere/topics/sea_ice/antarctic_sea_ice.html

However, increasing our knowledge of the Southern Ocean's biosphere is a pressing issue at times of climate change. Even baseline knowledge on many of the Antarctic species is still missing, as exemplified by our current knowledge of the distribution of Arnoux's beaked whales (*Berardius arnuxii*), which is based on sparse visual sightings (Figure 1.2). Thus (seasonal) distributions and abundance estimates of these and most other marine mammals are unknown. This makes quantifying the effects of climatic changes on these species difficult at best.

[1.1] Definition: The Southern Ocean, also known as the Antarctic Ocean or the South Polar Ocean, is, by definition of the International Hydrographic Organization, the oceanic division encircling Antarctica. It comprises the southernmost waters of the World Ocean south of 60° S latitude.

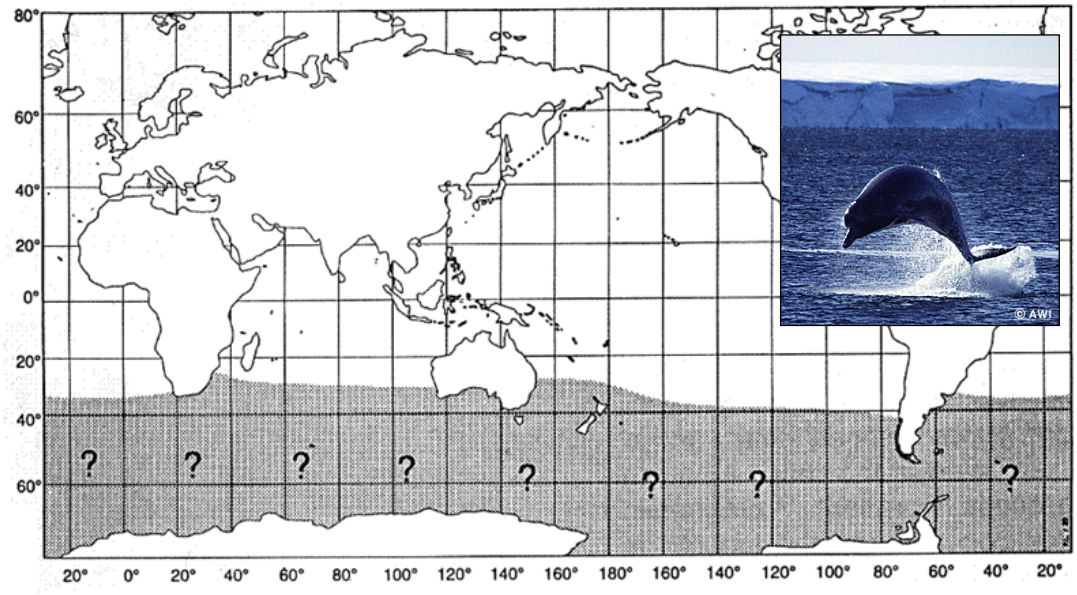


Figure 1.2: Supposed distribution of the Arnoux's beaked whale (*Berardius arnuxii*).
Source: Jefferson *et al.* (1993)

Particular concern focuses on the response of krill populations on a possible warming of the Southern Ocean. There, krill is the most important food source with highest levels of abundance in the summer months and shortages during winter (Hempel 1987). Marine mammals, such as baleen whales and seals (including the leopard seal), directly or indirectly (via penguins) depend on krill, and are hence immediately affected by changes in krill populations or distribution.

Extraordinary large gaps in our knowledge concern observations during austral winter, when the limited accessibility of the Southern Ocean does not allow studying the animals present during these harsh conditions. It is not known - except for Weddell seals (*Leptonychotes weddellii*) and minke whales (*Balaenoptera bonaerensis*) - which whale and seal species stay at high latitudes during polar winter. Many of these knowledge gaps can be overcome by the use of passive acoustic monitoring (PAM) (Tyack 1998), a method that recently become quite popular due to progress in digital audio technology (e.g. Richardson *et al.* 1995).

Marine mammals use underwater sound for communication, navigation and prey detection. Thus, PAM has the potential to provide year-round information on the presence/absence of animals, independent of weather conditions and without the need of direct (visual) contact with the animals. Detection of species' specific vocalizations unequivocally implies the presence of this species in the vicinity of the recording station. However, the absence of vocalizations does not necessarily imply the absence of animals and in many cases the direct correlation between vocalization rate and species presence is unknown. At this stage, it is simply not known to what degree species vocalize continuously, rendering this issue is an interesting research question by itself.

Depending on the frequency and intensity of the respective vocalization, the vocal behavior of marine mammals can be investigated within a range of several kilometers (for audible to ultrasound vocalizations) to about one hundred kilometers (for infrasound vocalizations) around a single hydrophone (underwater microphone) (Sirovic *et al.* 2007). Further refinement of PAM studies can be achieved by using multiple, synchronized hydrophones (a so called passive acoustic array), which allow calculation of the acoustic source's positions. Knowledge of the latter is a necessity to (a) determine the source level and (b) estimate the number of animals in the vicinity of the recording station. A passive acoustic array can also be used to determine migratory tracks of animals by calculating subsequent positions of an individually recognized vocalizing animal.

Furthermore, the use of PAM in the Southern Ocean in particular, allows investigation of the acoustic behavior of marine mammals in an environment almost undisturbed by humans. Recordings from this region will provide baseline information for future studies on (a) natural acoustic background levels and their response to a changing environment and (b) the influence of anthropogenic noise on the acoustic and locomotive behavior of marine mammals.

However, these latter research questions are beyond the scope of this thesis. To gain appropriate acoustic recordings from the Southern Ocean, an autonomous listening station named PALAOA[1.2] (Perennial Acoustic Observatory in the Antarctic Ocean) was developed. PALAOA was intended to continuously record the underwater soundscape over a period of several years. Consequently, PALAOA is producing enormous amounts of data. Extracting the signals of interest from this data set proved challenging. Obviously, human "observers" will not be able to efficiently manage such a task, but rather, numerical algorithms need to be developed to perform an automated, computer-based signal detection. The resulting time series of species specific vocalizations will then form the data base for ecological studies with focus on diurnal and seasonal patterns of the vocal behavior and their interrelation with the changing physical environment.

To obtain real-time acoustic broad band data from the Antarctic coastal ocean the following engineering issues had to be solved:

- How to reliably operate hydrophones in the Southern Ocean year-round?
- How to reliably power an autonomous station year-round?
- How to handle the large data volumes?
- How to establish a real-time data connection between an observatory in Antarctica and the Institute in Germany?

[1.2] PALAOA = Hawaiian; means (sperm) whale, whale tooth.

Once these engineering issues had been addressed and (at least partially) solved, PALAOA's acoustic data stream revealed a multitude of vocalizations, which – in the course of this thesis – could be attributed to various seal and whale species. However, analysis of all different vocalization types would by far exceed the scope of this study, which is why a single species was selected for the further development of automated pattern recognition algorithms.
For a number of reasons the species of choice was narrowed down to the leopard seal:

- The leopard seal is one of the key species for indicating changes in the Antarctic ecosystem. As a top predator, leopard seals play an important role in the Antarctic ecosystem. Consequently, leopard seals offer the possibility to study the effects of changes in the Antarctic environment as for example caused by global warming or commercial exploitation of krill resources.

- In 2003, acoustic recordings conducted in the Drescher Inlet (located 500 km S-SW of Neumayer Base) contained many leopard seal vocalizations but not a single individual was sighted by the experienced seal biologists who spent several hours each day on the sea ice. This result reiterated that (1) leopard seals are difficult to investigate visually at least in fast ice areas, and that (2) passive acoustics will likely provide new information without the need of direct access.

- Additional interest in a reliable detection of leopard seals developed in the context of diver safety. In 2003, a snorkeling scientist of the British Antarctic Survey (BAS) was attacked by a leopard seal and drowned by being pulled down to a depth of 70 m (Muir *et al.* 2006). The study by Muir *et al.* (2006) determined, that the number of injuries (not deaths) inflicted by leopard seals over the last 100 years is likely to be less than 50 in total. While the probability to be physically attacked by a leopard seal is low, the possible consequences of such an attack might be serious. When the Alfred Wegener Institute for Polar and Marine Research (AWI) conducted diving operations in the Southern Ocean during the expedition ANT XXIII-6 of R/V Polarstern (June to August 2006), safety regulations were implemented, which required acoustic observations prior to and during the diving operations. To this end, the underwater soundscape in the vicinity of the camp was monitored (Klinck & Burkhardt 2008). To do so most effectively, the acoustic characteristics of the leopard seal vocalizations needed to be studied in detail prior to the expedition and robust and fast detection algorithms had to be developed.

Given the sparse knowledge on leopard seals, it is not surprising that a substantial number of open scientific questions regarding leopard seals offered themselves to be addressed using the emerging new PALAOA data set:

- Do leopard seal vocalizations occur near PALAOA, and if yes, when?

- Which call types do they use, and how do they relate to recordings elsewhere?

- Which call types are suitable for automated detection (at different signal-to-noise ratios[1.3]), both in context of scientific research as well as diver safety, and how reliably do these calls occur when leopard seals are present?

- What are the typical patterns of frequency of occurrence of leopard seal vocalizations, and which factors influence these?

- Is it possible to identify individual leopard seals by the acoustic features of their vocalizations?

Before approaching the development of (hard- and software) research tools to answer these scientific questions, in the following paragraph the state of knowledge of passive acoustic monitoring techniques, automated detection methods and the biology of the leopard seal are described/summarized.

1.2 State of knowledge

1.2.1 Passive acoustic monitoring (PAM) devices

Passive acoustic monitoring (PAM) devices can be classified as (a) towed, (b) portable and (c) stationary. Towed systems (also known as passive acoustic streamers) are being operated from ships and are therefore beyond the scope of this thesis. Portable PAM devices are “on-site” systems. The acoustic sensor (hydrophone) and the recording electronics are merged in a waterproof housing and deployed on the ocean bottom for a certain time. The acoustic data are not available until retrieval of the device. The limitations of the devices are caused by the restricted amount of power (batteries) and storage capacity (hard discs) and by the high cost to deploy and retrieve these hydrophones.

Stationary PAM devises are “off-site” systems. Only the acoustic sensor (sometimes also the A/D converter) is deployed in the water. The acoustic signals are sent via optical or electrical cable to a station on land where the recording electronics are located. The main advantages of these systems are the “unlimited” amount of storage capacity and the real-time accessibility of the data. However, stationary PAM devices are more expensive and technically more complex than the portable systems. Table 1.1 gives an overview of the currently most prominent PAM devices.

[1.3] Signal-to-noise ratio (SNR) is defined as the ratio of a signal power to the noise power corrupting the signal.

Table 1.1: Prominent PAM devices.

	Name of Device	***Developed - Distributed by***	***Location***	***Technical specifications***
a	ARP Acoustic Recording Package	Scripps Institution of Oceanography, USA	Portable system	No. of hydrophones: 1 Sampling rate: 1 kHz/16bit Storage capacity: 72 GB Max. operation: 400 days (continuously recording)
b	HARP High-Frequency Acoustic Recording Package	Scripps Institution of Oceanography, USA	Portable system	No. of hydrophones: 1 Sampling rate: 200 kHz/16bit Storage capacity: 1920 GB Max. operation: 55 days (continuously recording)
c	AURAL M2	Multi-Électronique (MTE) Inc., Canada	Portable system	No. of hydrophones: 1 Sampling rate: 32.6 kHz/16bit Storage capacity: 60 GB Max. operation: 11.5 days (continuously recording)
d	T-POD	Chelonia Limited, UK	Portable system (stores only events)	No. of hydrophones: 1 Target frequencies: 9-170 kHz Storage capacity: 128 MB Max. operation: 90 days (continuously recording not possible)
e	Pop-up Recorder	Cornell University, USA	Portable system	No. of hydrophones: 1 Sampling rate: 2 kHz/16 bit Storage capacity: 80 GB Max. operation: 90 days (continuously recording)
f	HARUphone	NOAA-PMEL, USA	Portable system	No. of hydrophones: 1 Sampling rate: 2 kHz/16 bit Storage capacity: 400 GB Max. operation: >1 year (continuously recording)
g	EAR Ecological Acoustic Recorder	Hawaii Institute of Marine Biology, USA	Portable system	No. of hydrophones: 1 Sampling rate: 64 kHz/16 bit Storage capacity: 120 GB Max. operation: >1 year (scheduled recording)
h	PAL Passive Aquatic Listener	R.S.T. "Environmental Remote Sensing Technologies" Ltd., Greece	Portable system (stores only events)	No. of hydrophones: 1 Sampling rate: 100 kHz/16 bit Storage capacity: 2 GB Max. operation: > 1 year (continuously recording not possible)
i	ALOHA Cabled Observatory	University of Hawaii, USA	Hawaii, 022° 45' N 158° 00' W	No. of hydrophones: 1 Sampling rate: 96 kHz/24 bit Storage capacity: unlimited Max. operation: unlimited (cabled)
j	SOSUS Sound Surveillance System	U.S. Navy, USA	Mainly North Atlantic and North Pacific Position: classified	No. of hydrophones: classified Sampling rate: 256 Hz/16 bit Storage capacity: unlimited Max. operation: unlimited (cabled)

a, b) (H)ARP – (High-Frequency) Acoustic Recording Package

The ARP and also the HARP were developed by the Scripps Institution of Oceanography, University of California, USA. ARPs and HARPs are portable bottom moored recording systems (see Figure 1.3) and can be deployed anywhere in the world. The main components consist of a hydrophone, which is suspended 10 m above the mooring, batteries, ballast weights and a series of electronics handling acoustic release and data logging.

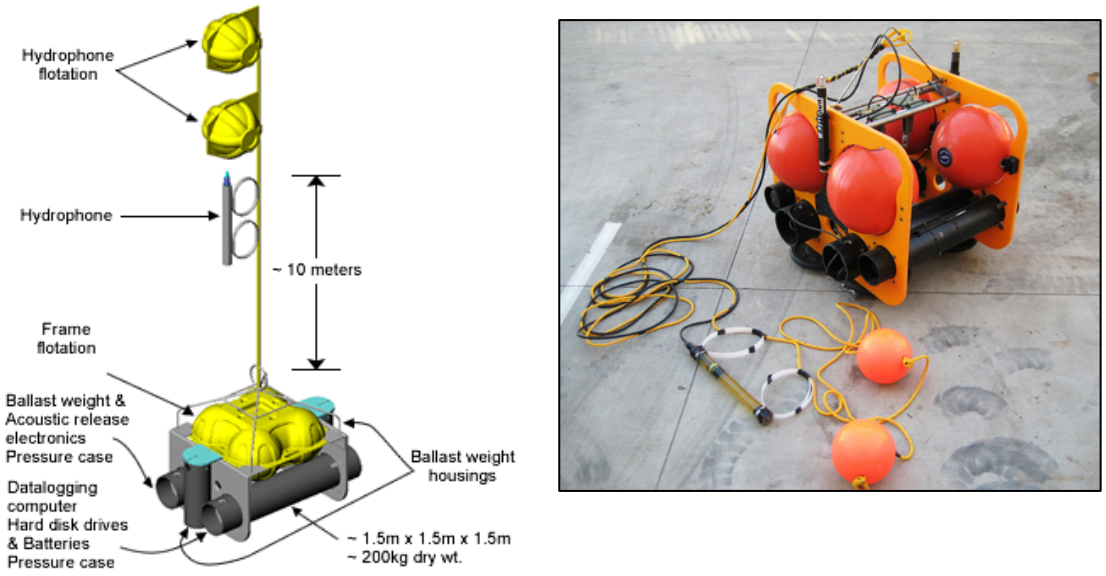


Figure 1.3: Design and image of an ARP.
Source: http://cetus.ucsd.edu/technologies_AutonomousRecorders.html

Both systems differ in the maximum sampling rate (ARP: 1 kHz; HARP: 200 kHz) and the storage capacity (ARP: 72 GB, HARP: 1920 GB). They are designed to either record continuously or record intermittently. The ARP can be operated for 400 days (HARP 55 days) in continuous recording mode. The system can be operated in a depth of at least up to 4000 m.

c) Aural M2

The Aural M2 recorder (see Figure 1.4) was developed by the Canadian company Multi-Électronique (MTE) Incorporation. The portable Aural M2 recorder can be operated continuously or intermittently with sampling rates up to 32.6 kHz.

Figure 1.4: Image of an Aural M2 recorder.
Source: http://www.multi-electronique.com/index_en.htm

Beside the acoustic system, the Aural M2 recorder is also equipped with a temperature and pressure sensor. Maximum operation depth is 300 m (in sea water).

d) T-POD

The T-POD (see Figure 1.5) was developed by the company Chelonia Limited, UK. The POD - 'POrpoise Detector' - is an automated detection system for porpoises, dolphins and other toothed whales.

Figure 1.5: Image of two T-PODs.
Source: http://www.chelonia.co.uk/index.html

The T-Pod works differently from all other mentioned PAM devices. The device is equipped with a built-in energy detector and records only date and time of detected events rather than the whole acoustic data. The device can be deployed for up to one year at a maximum operation depth of 3000 m. However, the system is only built for the passive acoustic detection of echolocation clicks of toothed whales (in a frequency range between 9 kHz and 170 kHz) and is therefore not suitable for detecting e.g. tonal sounds of pinnipeds.

e) Pop-up recorder

The pop-up recorder was developed by the Cornell University, USA. The device was designed as a portable bottom moored recording system (see Figure 1.6) and is in use in several oceans.

The device includes a microprocessor, several hard disks for data storage, acoustic communications circuitry, and batteries, all sealed in a single 17-inch glass sphere. An external hydrophone is connected to the internal electronics through a waterproof connector. At the conclusion of a mission, the positively buoyant sphere separates itself from its anchor and "pops up" to the surface for retrieval. The recorder can be operated continuously or intermittently with sampling rates up to 2 kHz. The storage capacity is 80 GB and allows to record continuously for 90 days. Maximum operation depth of the pop-up recorder is 6000 m.

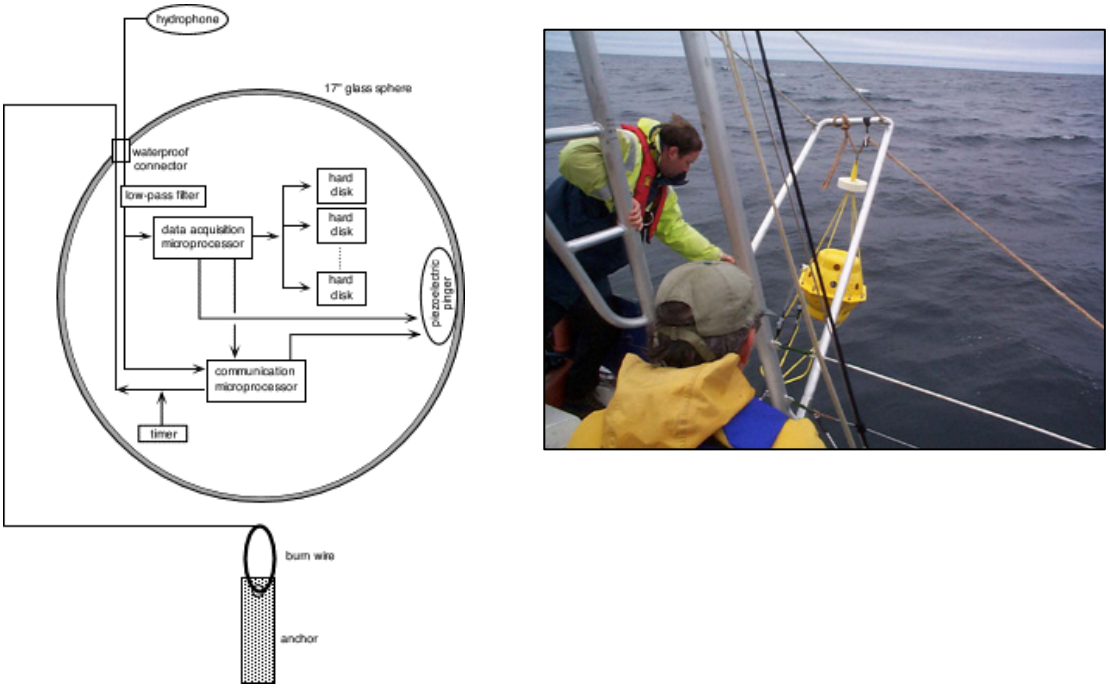


Figure 1.6: Design and image of a pop-up recorder.
Source: http://www.birds.cornell.edu/brp/hardware/pop-ups

f) HARUphone

The HARUphone was developed by the NOAA Pacific Marine Environmental Laboratory, USA. The device, which was primarily built to investigate sounds associated with seismic activities (e.g. earthquakes and underwater volcanoes), is deployed within a mooring and thus the operation depth can be varied. The sampling rate of the HARUphone is currently limited to 2 KHz. However, the device is suitable for long-term deployments > 1 year duration and is used to investigate baleen whales in various oceans of the world. A schematic of the HARUphone is given in Figure 1.7.

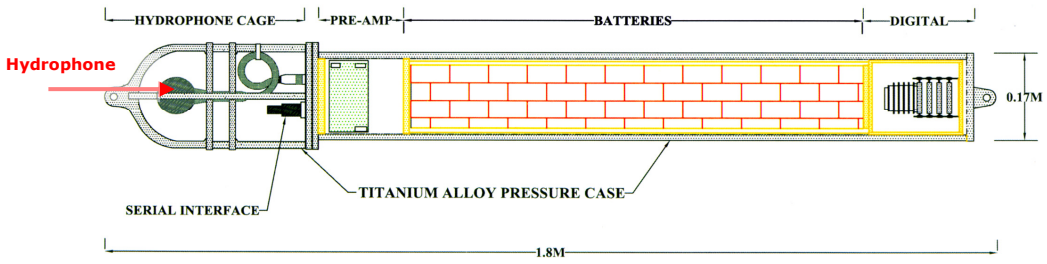


Figure 1.7: Design of the HARUphone
Source: Dr. Haru Matsumoto, NOAA-PMEL

g) EAR

The EAR recorder (see Figure 1.8) – developed by the Hawaii Institute of Marine Biology, USA – is particularly used to monitor the changing status of coral reef environments. The ear supports continuously/scheduled recording as

well as detection driven recording. The maximum sampling rate of the system is 64 kHz and the devices' maximum operation depth is 500 m.

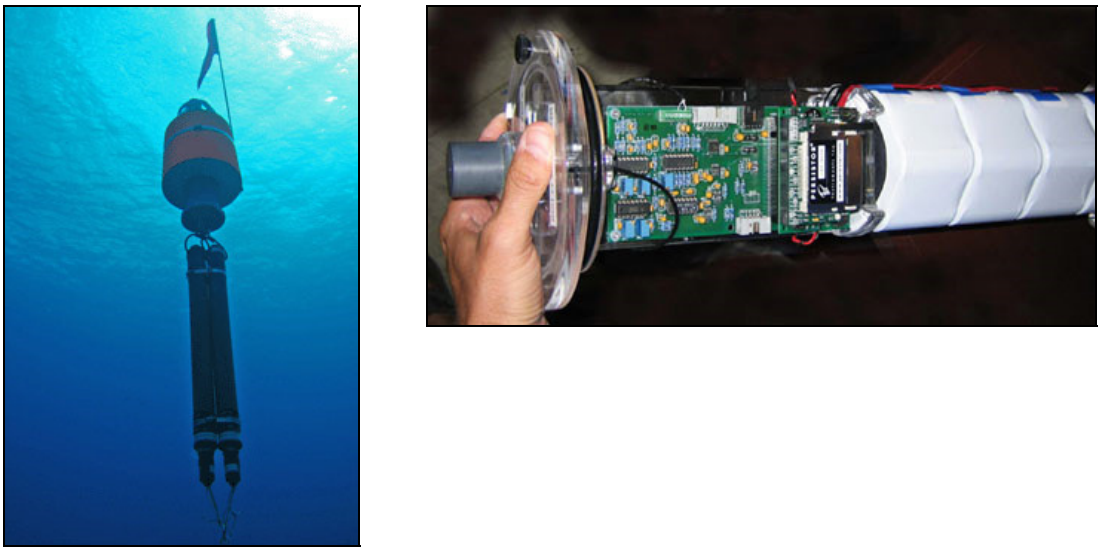


Figure 1.8: Images of the EAR recorder.
Source: http://www.pifsc.noaa.gov/cred/ear.php

h) PAL

The PAL recorder is distributed by R.S.T. "Environmental Remote Sensing Technologies" Ltd., Greece. The recorder was built for geophysical measurements especially wind speed and rainfall at sea. However, recently the devices' software was modified to enable marine mammal research. The PAL is laid out to detect and store events and cannot be operated continuously or intermittently. However, the high sampling rate of 100 kHz allows to detect ultrasounds produced by toothed whale species.

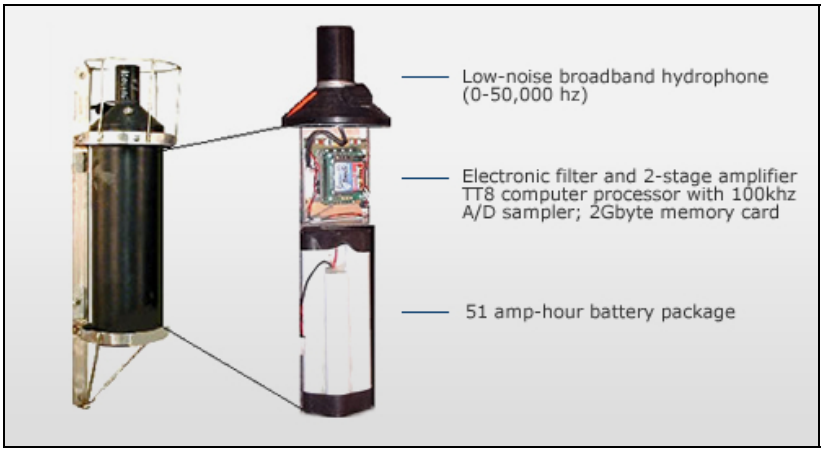


Figure 1.9: Image and design of the PAL recorder.
Source: http://www.rstech.com.gr/products_01.html

i) ALOHA Cabled Observatory

The ALOHA Cabled Observatory is operated by the School of Ocean and Earth Science and Technology, University of Hawaii (USA). The ALOHA Cabled Observatory is a multi sensor platform deployed at a depth of 5200 m in the Pacific ocean approx. 150 miles from Pearl Harbor (see Figure 1.10).

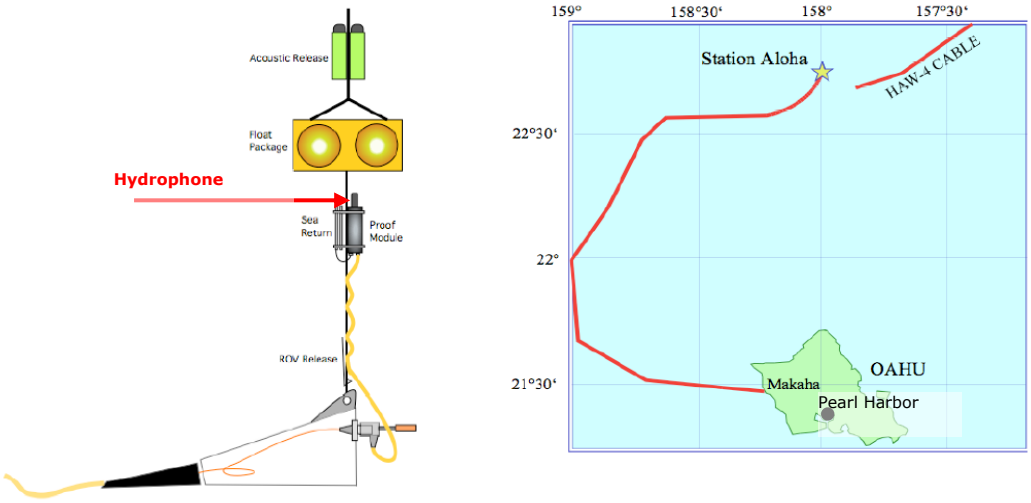


Figure 1.10: Design and location of the ALOHA Cabled Observatory.
Source: http://www.soest.hawaii.edu/GG/DeepoceanOBS/aco_home_page.htm

ALOHA has been in operation since end of September 2006 and was equipped with a single hydrophone. The data are sampled with up to 96 kHz/24 bit and transmitted via cable to Makaha, Oahu Island. A live stream of the ALOHA acoustic data is presented to the public in the World Wide Web at http://station-aloha.soest.hawaii.edu:8000.

j) SOSUS - Sound Surveillance System

The Sound Surveillance System is a multibillion-dollar project of the US Navy started during the Cold War in the mid 50's. The goal of this project was to install an underwater surveillance system to detect and track submarines. The so-called Sound Surveillance System (SOSUS) is a network of hydrophone arrays mounted on the seafloor throughout Atlantic and Pacific Ocean (see Figure 1.11). The hydrophones were placed into a "sound channel" that exists in the ocean and allows low-frequency sounds to travel over long distances. This sound channel is called the SOund Fixing And Ranging (SOFAR) channel. Submarines produce low-frequency sounds which can be detected at long ranges by hydrophone arrays located on continental slopes and seamounts.

The hydrophone arrays monitor the oceans, record sounds, and transmit the data back via undersea cables to onshore stations for further analysis. At the end of the Cold War, the Navy decided to allow this system to be used by scientific researchers with suitable security clearances, in what was called "dual-use" (Nishimura *et al.* 1994). SOSUS is now also used to study the vocal behavior of whales. Scientists can study and track whales in the Atlantic and

Pacific oceans using the SOSUS hydrophone arrays (http://www.pmel.noaa.gov/vents/acoustics/sosus.html).

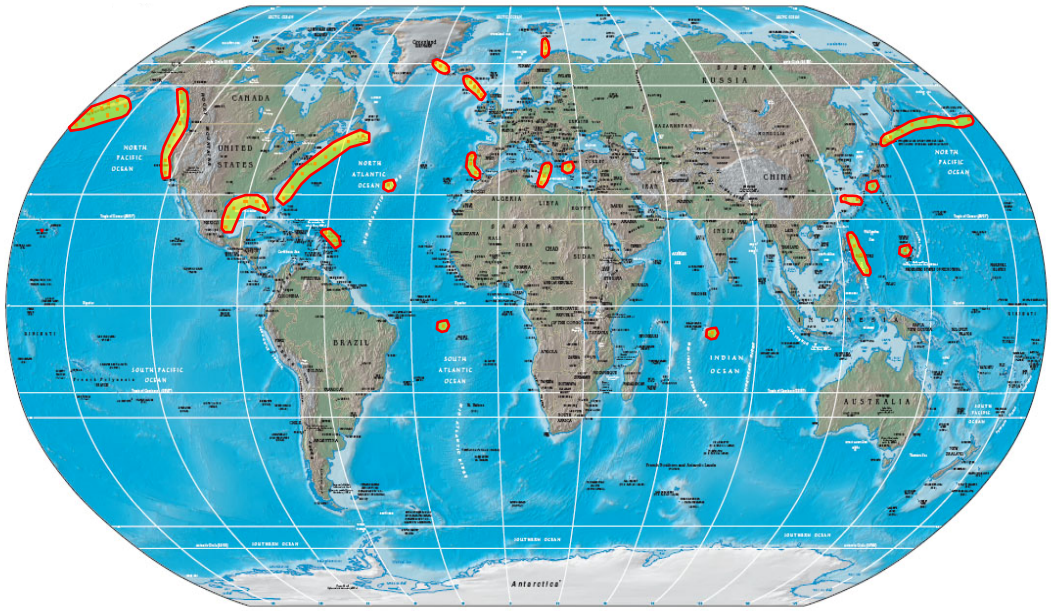


Figure 1.11: Assumed (unconfirmed!) positions of the SOSUS arrays (yellowish highlighted areas). Source: http://de.wikipedia.org/wiki/Sound_Surveillance_System

The large-aperture arrays of high sensitivity located within the SOFAR channel allow to detect radiated acoustic power of less than a Watt at ranges of several hundred kilometers. However, the sampling rate of 256 Hz/16 bit only allows to study the low-frequency vocalizations of baleen whales.

1.2.2 Automated detection algorithms

There are four automated detection methods commonly used in marine mammal research. The presented methods are already implemented in freely available bioacoustic research software packages (for example Ishmael[1.4] or XBAT[1.5]).

a) Energy summation

The energy summation method is a simple but fast method often used to detect any signal in a data set which could be of interest. The output of the energy summation method (any signal of interest) can be fed into a more sophisticated classification system to filter the detections for the interesting ones. The energy summation method sums up the energy in all frequency bands in a time frame. The resulting time series of sums is then used in combination with a threshold to detect signals which stand out against the normal background noise level. If the frequency band of the target signals is known, a band-pass filter can be applied to the acoustic data in advance to reduce the number of false positive

[1.4] http://www.pmel.noaa.gov/vents/acoustics/whales/ishmael/
[1.5] http://www.xbat.org

detections. A more advanced version of the energy summation method involves working with a target frequency band and one or two reference frequency bands. In this case the detector compares the energy per time frame in these frequency bands by calculating a ratio. The ratio is high when the energy in the target frequency band is relatively high compared to the reference frequency bands. The resulting time series of ratios is then again used in combination with a threshold to detect signals of interest. This method allows discarding unwanted detections of broad frequency sounds.

b) Matched filtering

Matched filtering is used to detect signals with low variability within samples of a species in the time domain. In a first step a reference kernel has to be calculated. The kernel is calculated by selecting several high-quality (good signal-to-noise ratio) samples of the target signal. For each sample the waveform is normalized to zero mean and unit energy. The normalized waveforms are placed together in a matrix. A sample covariance matrix is then calculated. Based on the covariance matrix the *eigenvector* with the highest *eigenvalue* is computed (Mellinger & Clark 2000).
In a second step the detector is used to calculate the cross-correlation of the input signal and the kernel. If the cross-correlation reaches a certain threshold, a call is detected. Matched filtering works best for detecting signals in white Gaussian noise (van Trees 1968).

c) Spectrogram correlation

The spectrogram correlation detector is used in the time-frequency domain and allows detecting signals with a higher degree of variation compared to the matched filtering detector. First, a synthetic time-frequency kernel is computed. This computation is a rather complex process as many parameters of the kernel can be adjusted (Mellinger & Clark 2000). This allows defining the degree of variability in the calls to be detected. The resulting kernel (reference spectrogram) is then cross-correlated with the spectrogram of the input signal. Again, a certain threshold is used as a final decision step to detect the target signals.

d) Machine learning algorithms

Machine learning algorithms such as artificial neural networks and support vector machines have been only sparsely used for the detection of marine mammals in the past. However, recent publications (Mellinger & Clark 2000) make the efficiency of machine learning algorithms clear. It is assumed that these algorithms will play an important role within the research field of passive acoustic detection of marine mammals in the next years.

e) Human speech recognition methods

State of the art human speech recognition systems are based on Hidden Markov Models (HMM). These statistical models represent a language unit, for instance a word or a phoneme. A HMM has a finite number of states and the transitions between these are probabilistic and take place once every time unit

(a model may also remain in the same state). Each state has a probabilistic output function which represents a random variable or a stochastic process (Rabiner & Juang 1993). Gaussian distributions are a common choice for representing these functions, and in reality, mixtures of Gaussians with individual means and variances and mixture weights are often used, as these allow any arbitrary function to be approximated. When presented with an observation sequence, a model can determine the probability of having generated the observations, but since the observations do not uniquely define a particular state sequence, it is not possible to know which states were active, and in what order. This is why the process is said to be hidden. The transition probabilities and the probability distributions along with their weights are the parameters of a HMM. During training, these are optimized with respect to the training data to increase the likelihood of the models having generated the data. Hidden Markov Models are used in the ornithology community for the detection and classification of bird vocalizations. However, the Hidden Markov approach has not been widely used by the marine mammal acoustic research community yet.

1.2.3 Leopard seal (*Hydrurga leptonyx*)

1.2.3.1 Distribution

Leopard seals - solitary living pack ice seals[1.6] – have a circumpolar distribution around the Antarctic continent between 50°S and 80°S. The main population (around 200,000 animals) lives within the circumpolar pack ice with higher densities near the pack ice edge (e.g. King 1983; Siniff 1991; Bester *et al.* 2002). The seasonal abundance of leopard seals on Subantarctic Islands peaks between August and September with absence of animals from late austral spring to mid-autumn (Walker *et al.* 1998). During austral spring the adults move towards the Antarctic continent into the inner pack ice to breed (Brown 1957; Erickson *et al.* 1971; Siniff & Stone 1985). Broad-scale movement of leopard seals are related to ice conditions (Erickson *et al.* 1971; King 1983). Animals move northwards with the expanding pack ice edge in the austral winter when the ice reaches a maximum cover and thickness (Brown 1957; Erickson *et al.* 1971; Siniff & Stone 1985). The periodic appearance of leopard seals on Bird Island (54°00'S, 38°03'W) is due to frequency and timing of the Antarctic Circumpolar Wave (Jessopp *et al.* 2004). Years with low temperature and extensive pack ice cause more seals to migrate further north (Jessopp *et al.* 2004). Other reasons for such movement may be inter-specific aggression or food shortages (Øritsland 1977). The latter could be the reason why on some islands the presence of leopard seals is correlated with penguin (Borsa 1990; Lesinski 1993; Hiruki *et al.* 1999) and fur seal (Jessop *et al.* 2004) breeding seasons. Most leopard seal sightings at these islands were nearby penguin rookeries (Rogers *et al.* 2005). There is still much to learn about leopard seal distribution and habitat use because of their solitary nature (Borsa 1990) and because of the logistic difficulties of working within the Antarctic pack ice

[1.6] Sea ice is classified into two categories: pack ice (floating ice) and fast ice (ice attached to a shore). Thus, the appellation pack ice seals (like the leopard seal) is resulting from the fact that pack ice seal species give birth to their pubs in the pack ice zone. By contrast Weddell seals give birth to their pubs on fast ice.

(Rogers *et al.* 2005). However, passive acoustic investigations are becoming a powerful means of increasing our knowledge of the leopard seal's population structure and migratory behavior.

1.2.3.2 Physiology and Morphology

The leopard seal has a long and slender body with a disproportionately large head (Riedmann 1990) emphasizing the often described "reptilian appearance" (Riedmann 1990; Laws 1993) of this seal (see Figure 1.12). Leopard seals can reach a length of more than 4 m and a weight of up to 500 kg.

Figure 1.12: The leopard seal (*Hydrurga leptonyx*).
Source: http://channel.nationalgeographic.com/channel/crittercam/show_description_10.html

The long snout has a wide gape, including powerful jaws with large re-curved canines and incisors as well as upper and lower tricuspid molars that interlock (Riedmann 1990; Rogers & Cato 2002). This dentition is well adapted for gripping and tearing prey while the three interlocking cusp molars provide an effective krill sieve (Rogers & Cato 2002). As described in Reeves *et al.* (2002), adults have a grey pelage that is dark dorsally and light ventrally, with dark blotches scattered over the chest, belly and flanks. They are silvery just after moulting in January and February but fade to grey over the next several months. The pelage of newborns is similar to that of adults, although their hair is longer and softer. The leopard seal travels with agility in a manner similar to the crabeater seal (*Lobodon carcinophaga*) but with less speed and with little use of the fore flippers (Ray 1970). Leopard seals exhibit sexual dimorphism. Females are generally larger than males, being 3.8 m in length and weighing up to 500 kg, while males reach 3.3 m with 300 kg in mass (Reeves *et al.* 2002; Rogers & Cato 2002). Leopard seals can reach an age of 25 or 30 years (Siniff 1991).

1.2.3.3 Diet

Leopard Seals, along with killer whales, are the top predators in the Antarctic marine ecosystem (Reeves *et al.* 2002). Leopard seals are known to feed on diverse prey including krill, fish, penguins, seals, cephalopods and crustaceans (e.g. Kooyman 1981; Laws 1984; Hiruki *et al.* 1999; Hall-Aspland *et al.* 2005). From available data on body weights and an assumed feeding rate

corresponding to 7 percent of the body weight, the amount of food consumed per day is estimated to be about 20 kg for an average (~ 260 kg) adult leopard seal (Øritsland 1977). The maximum volume found by Murphy in 1948 was 64 kg of penguins, reported from a leopard seal stomach, which refers to periods of intense feeding or situations of abundant food supply (Øritsland 1977). The lack of bones or intestines found in scats suggests that leopard seals do not consume their prey (seals and penguins) whole (Hall-Aspland & Rogers 2004). Accordingly, observations confirm that leopard seals only eat the blubber or skin of seals, leaving the remainder of the carcass intact (Bonner 2004). Peaks in the incidence of various prey species in the diet of the leopard seal apparently coincide with periods of abundance of different prey species (Siniff & Stone 1985). The leopard seal is responsible for more predation on warm-blooded animals than any other pinniped (Muir *et al.* 2006). Apart from crabeater seals (*Lobodon carcinophaga*), leopard seals also prey on southern elephant seals (*Mirounga leonina*), Antarctic fur seals (*Arctocephalus gazella*) (Rogers & Cato 2002), Ross seals (*Ommatophoca rossi*) and Weddell seals (*Leptonychotes weddellii*) (Perrin & Bronwell 2002), although predation on the two last named species is likely to be insignificant (Siniff & Stone 1985). During the austral mid-summer period of December and January, newly-weaned crabeater seals become important prey. Then in late January and February, young penguins, mostly Gentoo (*Pygoscelis papua*), Macaroni (*Eudyptes chrysolophus*) and Adélie penguins (*Pygoscelis adeliae*) become available and are extensively taken when they enter the ocean for the first time (Lowry *et al.* 1988; Siniff 1991). Krill (mainly *Euphausia superba*), makes up the largest proportion of the leopard seal's diet (Perrin & Bronwell 2002), and is an important food item in late September until the beginning of the austral summer when other prey items are available (Siniff & Stone 1985). It becomes important again in March, and probably remains so throughout the austral winter (Siniff & Stone 1985). Cephalopods are an important food item in January, but some are taken throughout the year (Siniff & Stone 1985).

The annual combined biomass of prey taken by the leopard seal has been estimated due to stomach content analysis and mathematical calculation to consist of 45% krill, 35% seal, 10% penguin and 10% fish and cephalopods (Siniff & Stone 1985). However there might be site-specific variation in the composition of the leopard seals' diet. The opportunistic nature of the foraging pattern does not principally differ from the feeding behavior of other pinniped species (Geraci 1975), but it does seem extreme with respect to the size and taxonomic range of prey items (Siniff & Stone 1985). It seems that winter could be a period of food stress in certain areas. During winter, leopard seals probably compete with krill-feeding specialists such as crabeater seals and Adélie penguins. In comparison to these specialists, the leopard seal is almost certainly not as efficient in foraging for krill and could, therefore, be one of the first species adversely affected by the fluctuation of krill stock (possibly caused by krill fisheries, climate variability and/or climate change) (Siniff & Stone 1985). Other evidence for food stress during winter might be the dispersal of immature and some adult leopard seals to Northern islands outside the pack ice area (Gwynn 1953; Rounsevell & Eberhardt 1980).

1.2.3.4 Social behavior

Leopard seals are generally solitary and most observations have been of single animals (Gilbert & Erickson 1977; Hiruki *et al.* 1999), but there are also observations of leopard seals interacting in captivity as well as in the wild. Rogers *et al.* (1996) described twelve different behavior patterns in a three-year study on two captive leopard seals, most of which consisted of different swimming patterns of the seals. These behaviors were observed in association with vocalizations. Nine of the patterns were considered to be agonistic in function, including mainly aggressive or threat behavior. The remaining three were categorized as breeding or non-interactive behavior (Rogers *et al.* 1996). Hiruki *et al.* (1999) described observations of leopard seals in the wild which were seen in pairs on five occasions at Seal Island, mostly in situations assumed to be mating or playing. Another observation describes the situation where a smaller leopard seal followed or waited nearby a larger seal that was hunting, either to imitate the hunting technique or to scavenge from prey carcasses that the larger seal left behind (Hiruki *et al.* 1999). It seems unlikely that co-operative hunting exists as a common hunting strategy, because the probability of a single hunter capturing prey is high enough. Nevertheless, it is possible that juveniles follow adults to learn hunting techniques (Hiruki *et al.* 1999). Acoustic behavior is likely to be an important key to improve our understanding of leopard seal social interactions.

1.2.3.5 Reproductive behavior and life cycle

According to a study by Kuhn *et al.* (2006), who observed a juvenile male leopard seal fitted with a satellite tag, the majority of time (58.9%), is spent hauled out for sleeping, resting and for adults pupping on the ice floes of the pack ice (Rogers 2007). However this percentage may differ according to age and season, as the mentioned study from Kuhn only observed one juvenile over a period of 23 days (8 August - 31 August 2002). In summer months the peak haul-out periods are around 0930 and 1300 local time, probably due to a lack of krill in the surface waters (Gilbert & Erickson 1977). According to the study of Kuhn *et al.* (2006), haul-out behavior is followed by cruising (27.4%) and the rest of the time is passed underwater while diving (13.7%). Most of the dives observed were shallow with only 13% exceeding 100 m, and 2% deeper than 150 m. The maximum depth of dive recorded for this animal was 424.5 m and the mean dive duration was quite short (2 min, max = 9.35 min; Kuhn *et al.* 2006).

Female leopard seals are believed to be sexually mature at 4 years and males at 4.5 years of age (Laws 1984). Leopard seals are solitary breeders and are sparsely distributed on the outer fringes of the pack ice (Laws 1984). In this habitat ice floes provide numerous platforms to whelp and nurse pups in areas of abundant food supply with unrestricted access to the sea (Riedmann 1990; Rogers *et al.* 1996). The constant break-up of pack ice during summer forces the leopard seal, like other pack ice-breeding seals, to have short breeding seasons (Riedman 1990). Pupping is believed to occur from October to mid-November. At birth pups are about 120 cm, with rapid growth through the first 6 months (Rogers & Cato 2002). Lactation is believed to last about 4 weeks

(Rogers & Cato 2002). It is assumed that the female is not accompanied by the male during this period (Siniff 1991) as only mother-pup groups were sighted (Muir *et al.* 2006). Mating probably takes place from December to early January after pups have been weaned, although it has never been observed in the wild (Muir *et al.* 2006). Leopard seal males and females are widely dispersed at the start of the mating season. In this situation of finding a sexually receptive mate during the brief breeding season, underwater acoustic communication could play an important role (Rogers *et al.* 1996). A detailed discussion on this topic will follow later on in this thesis. Communication must occur over long distances and both females and males have been found to produce loud broadcast calls during the breeding season (Rogers *et al.* 1996; Rogers & Bryden 1997). Lone males are known to vocalize underwater for many hours each day, which may serve as an indicator of male fitness as these displays require the male to be in good body condition (Rogers 2007). Females are thought to produce broadcast calls to advertise their sexual receptivity to distant males (Rogers *et al.* 1996). However, further acoustic studies are necessary to understand the vocal behavior of the leopard seal.

1.2.3.6 Underwater acoustic repertoire and behavior

The first (partial) spectrogram of a leopard seal vocalization was published by Ray (1970), while Stirling & Siniff (1979) described four different leopard seal call types quantitatively. A comprehensive description of the vocal repertoire of leopard seals was published by Rogers *et al.* (1995). Rogers identified twelve different call types by analyzing recordings of captive and free living animals (Prydz Bay, Antarctica). The frequency span of the analyzed call types ranges between 65 Hz and 4800 Hz. Thomas *et al.* (1983) recorded ultrasonic vocalizations with frequencies up to 164 kHz of captive leopard seals during hunting activity. However, ultrasonic vocalizations have so far not been reported during any of the field studies.

Free living leopard seals are known to be vocal underwater before and during breeding, from November to mid-January (Thomas & DeMaster 1982; Stirling & Siniff 1979). The function of the respective vocalizations is often difficult to assess, since they are rarely correlated to visual observations or to behavioral data sampled by autonomous bio-logging devices. Stirling & Siniff (1979) suggested territoriality as an associated function, but Rogers *et al.* (1996) were the first to link leopard seal vocalizations to the behavioral context in which a call is used by working with captive animals. Rogers *et al.* (1996) proposed that underwater acoustic behavior is important during the mating period, and that broadcast calls are used by mature female leopard seals to advertise their sexual receptivity, and most likely also by mature males in search of mates.

In contrast to otariids[1.7], females of many phocid[1.8] species remain with their pups for the duration of lactation so there is less need for a female to distinguish her own pup from others (Insley 2001). However, mother-pup communication in leopard seals and vocal learning in leopard seal pups has not been investigated

[1.7] (Antarctic) eared seals: Antarctic fur seal.
[1.8] (Antarctic) true seals: Southern elephant, Weddell, Ross, crabeater and leopard seal.

at all. Rogers (2007) studied age-related differences in the acoustic characteristics of male leopard seal sounds and revealed that juveniles produce additional variants of the commonly known calls. This phenomenon is probably related to an overproduction stage (Nelson & Marler 1993) where young animals produce a wide range of vocalizations memorized at an early age or imitated from nearby adults (Nottebohm 1972; Ralls *et al.* 1985; Catchpole & Slater 1995; Podos *et al.* 1999).

1.3 Approach

1.3.1 The Perennial Acoustic Observatory in the Antarctic Ocean

To obtain hydro-acoustic data from the Southern Ocean, the first goal of this thesis was to develop a tool to obtain multi-channel (for localizing sound sources), broad-band (for recording low and high frequency sounds) long-term (for investigating seasonality in the occurrence of sounds) recordings of the Antarctic underwater soundscape. To avoid multi-year delays between deployment and retrieval of data typical for moored underwater-recorders, and to overcome local data storage capacities, the recordings needed to be forwarded in real-time to the AWI in Bremerhaven. For these reasons the "**P**erenni**AL** **A**coustic **O**bservatory in the **A**ntarctic Ocean" (PALAOA) was developed. Adjacent to the Atlantic sector of the Southern Ocean, PALAOA was/is designed to autonomously record the Antarctic underwater soundscape year-round with an array of four synchronized hydrophones. The observatory was set up on the Ekström Ice Shelf, which permits, by using a WLAN connection to the German Neumayer Base, a real-time connection to Germany.
The implementation of PALAOA was challenging for many reasons. The installation of the observatory was hindered by the limited accessibility and the extreme climatic conditions. Determining the optimal position of the observatory required a detailed in-situ analysis of the thickness and shear of the ice shelf. To access to the Southern Ocean, a penetration of the 100 m thick ice shelf was necessary. To ensure year-round operation of PALAOA a sophisticated energy module had to be developed. Streaming the PALAOA data in real-time to Germany required an elaborate combination of hardware and software. In summary, the PALAOA project was a highly multidisciplinary project, covering aspects from the fields of glaciology/seismology, computer engineering, communication engineering, acoustic engineering and electronics/electrical engineering.

The observatory is providing the first, broad-band, long-term data record of the Southern Ocean, containing information on the seasonal and diurnal acoustic behavior of marine mammals as well as on the natural background noise. The acoustic module of PALAOA is in quasi-continuous operation since austral spring 2006, though some substantial data gaps still exist for polar winter when supplying power to PALAOA was difficult. To date (December 2007) more than 8727 hours of acoustic data have been collected and are available for further analyses.

1.3.2 Data analyses

The second part of the thesis, the analysis of the acoustic data stream for leopard seal vocalizations, was accomplished in three major steps.

1.3.2.1 Manual data analysis

Firstly, a "manual" data analysis of the PALAOA data set was conducted by listening to and visually detecting displayed spectrograms of the recorded data. The goals of the analysis were (a) to derive the different types of leopard seal vocalizations occurring in the vicinity of PALAOA and (b) to get detailed information on the acoustic features (e.g. length, frequency range) of the most prominent call types (see chapter 3.1.2.3), the low double trill (LDT) and the high double trill (HDT). For this reason a graphical user interface (GUI) was programmed in Matlab™, to determine the most important characteristics of a call in a reasonable amount of time. Furthermore a reference data set was created by manually counting the LDTs and HDTs over a one week period in certain time intervals. The resulting reference data set was of vital importance to test the performance of the automated detection algorithms.

1.3.2.2 Automated detection of leopard seal vocalizations

Based on the results of the manual analysis, automated detection algorithms for the LDT and HDT based on human speech recognition techniques were developed. For LDTs a detection method based on the zero-crossing rate of the waveform was developed. To detect the HDTs, a more sophisticated Hidden Markov Model was applied. The performance of each detection system was tested using the manually created reference data set. The goal was to automatically detect the HDTs and LDTs reliably and in real-time. The algorithms were used to screen the whole PALAOA data set for leopard seal vocalizations.

1.3.2.3 Identification of individual leopard seals

For the identification of individual leopard seals, the pulse repetition rates of the HDTs were analyzed in detail. The developed analysis tool tEST (the Envelope-Spectrogram Technique) - which was also used in the HMM based detection system - calculated the envelope of the amplitude modulated waveform of the HDT. The envelope was transformed into the frequency domain by means of a Fast Fourier Transformation (FFT). The resulting spectrogram provided information of the temporal course of the pulse repetition rate throughout the call. It is assumed that the call frequency and the associated pulse repetition rate is a function of the animals' vocal tract. Thus, these parameters should provide - similar to humans - individual information. To prove this hypothesis, data of three male leopard seals recorded in Prydz Bay, Antarctica, by Tracey Rogers[1.9], were analyzed. In addition, a previously unanalyzed sub-data set recorded at PALAOA was investigated to determine if the number of vocalizing animals can be derived using the developed technique.

[1.9] Director of the Australian Marine Mammal Research Centre, Sydney, Australia.

2. The Perennial Acoustic Observatory in the Antarctic Ocean

The Perennial Acoustic Observatory in the Antarctic Ocean (PALAOA) was developed to investigate the underwater vocal behavior of the leopard seal and other marine mammals. The general project requirements were (1) to build an autonomously operating observatory which (2) records the Antarctic underwater soundscape year-round and continuously, (3) covers a frequency range between 10 Hz and 96 kHz, (4) allows real-time access to the acoustic data and (5) permits the determination of the sound sources' positions.

2.1 Environmental conditions

2.1.1 Atmospheric conditions

Meteorological observations have been carried out at the Neumayer Base for more than 14 years since 1993. The observations are contributing to the World Meteorological Organization (WMO) network Global Atmospheric Watch and provide basic weather forecast information for the DROMLAN[2.1] network.

2.1.1.1 Radiation

Figure 2.1 depicts monthly averages of radiation data.

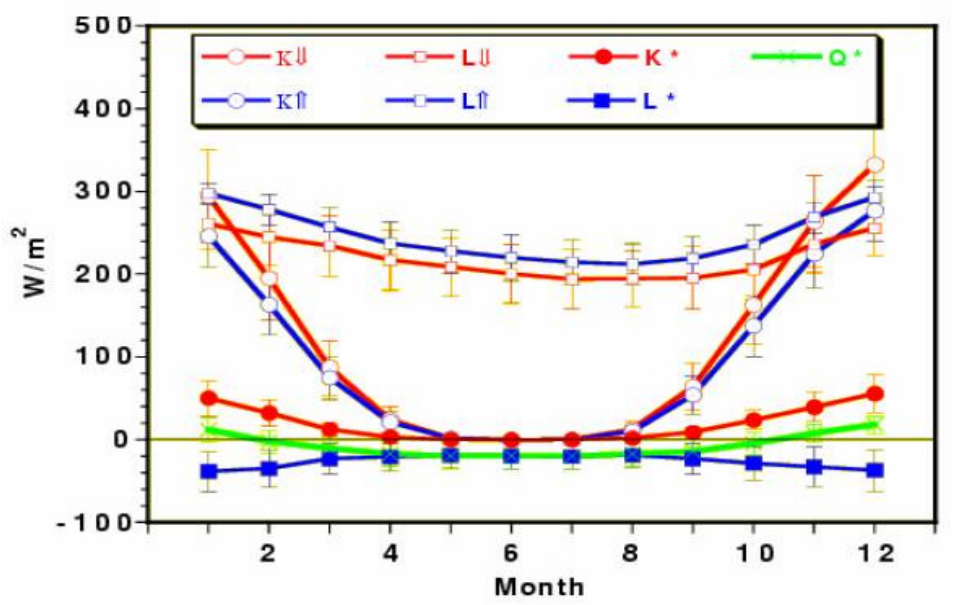


Figure 2.1: Radiation measured at Neumayer Base, Antarctica. K↓ = incoming solar radiation; K↑ = outgoing (reflected) shortwave radiation; L↓ = incoming longwave radiation; L↑ = outgoing longwave radiation; K* = net shortwave radiation; L* = net longwave radiation; Q* = net radiation. (Source: Dr. Gert König Langlo, Department of Meteorology, Alfred Wegener Institute)

The short wave components K are particularly relevant to the energy production via solar panels. During austral summer (October to March), K reaches a clear maximum because of the 24 hours of daylight (see Figure 2.1: K↓ = incoming solar radiation), while in austral winter (April to September) K is at a minimum because of the prevailing darkness.

[2.1] Dronning Maud Land Air Network: http://www.scar.org/information/dromlan/dromlanflyer.pdf

2.1.1.2 Wind

To get information on the monthly frequency of occurrence of the main wind speeds at Neumayer Base, the 3-hourly routine synoptic observations (measured at 10 m height) from the years 2005 and 2006 were analyzed (Figure 2.2.)

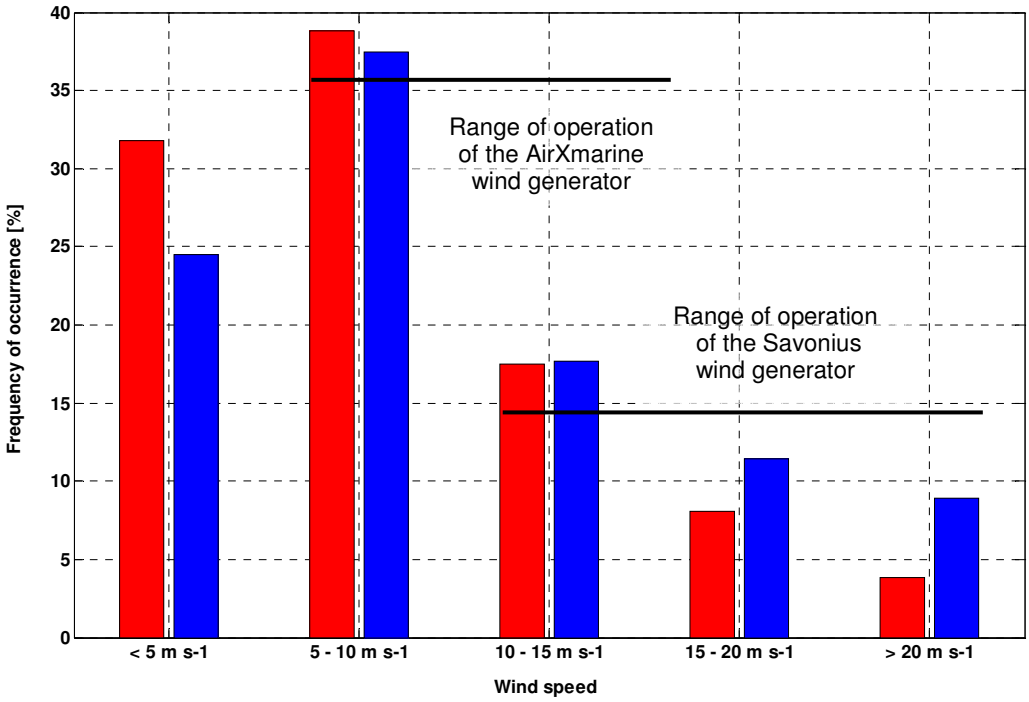


Figure 2.2: Frequency of occurrence of main wind speeds at Neumayer Base in austral summer (October to March, red) and austral winter (April to September, blue) and the range of operation of the wind generators implemented in PALAOA.

The frequency of occurrence of the main wind speeds was calculated for austral summer (October to March) and austral winter (April to September). During the latter period, only negligible amounts of sunlight are available to produce solar power (see Figure 2.1) and wind must hence serve as the primary energy source. Wind speeds in the range of operation of the AirXmarine wind generator occurred 56% of the time in austral summer and 55% of the time in austral winter (Figure 2.2) In contrast the frequency of occurrence of wind speeds in the range of operation of the Savonius wind generator increased from 29% (austral summer) to 38% (austral winter).

Monthly mean wind speeds at 10 m height range between 6.7 m s^{-1} in January and 10.1 m s^{-1} in August (see Table 2.1). In general, wind speeds are significantly higher during the winter months. However, occasional wind speeds exceeding 40 m s^{-1} (~ 150 km h^{-1}) can occur year-round.

Table 2.1: Monthly averaged wind speeds [m s^{-1}] measured at Neumayer Base in 10 m height. All wind speeds are given on a monthly time basis as a result of a 14 year observation period. Mean = mean wind speed, MeanMax = maximum mean wind speed and MaxMax = maximum wind speed. (Source: Dr. Gert König Langlo, Department of Meteorology, Alfred Wegener Institute)

Month of year	*Mean wind speed [m s^{-1}]*	*MeanMax wind speed [m s^{-1}]*	*MaxMax wind speed [m s^{-1}]*
1	6.68	20.16	28.32
2	7.71	21.54	28.81
3	9.92	24.30	28.81
4	9.90	26.76	32.96
5	9.61	26.69	32.41
6	9.78	27.69	35.49
7	9.62	28.14	36.52
8	10.11	29.89	36.52
9	9.49	26.83	32.92
10	9.24	27.39	33.95
11	9.71	25.26	33.95
12	7.09	21.34	30.35

For the installation of wind generators the most frequent wind directions have to be considered to avoid placing the generators in the lee of the pylon. Figure 2.3 displays mean wind speeds, wind directions and their frequency of occurrence at Neumayer Base. Southerly (185°) and easterly (90°) wind directions predominate with southerly winds being of lesser intensity.

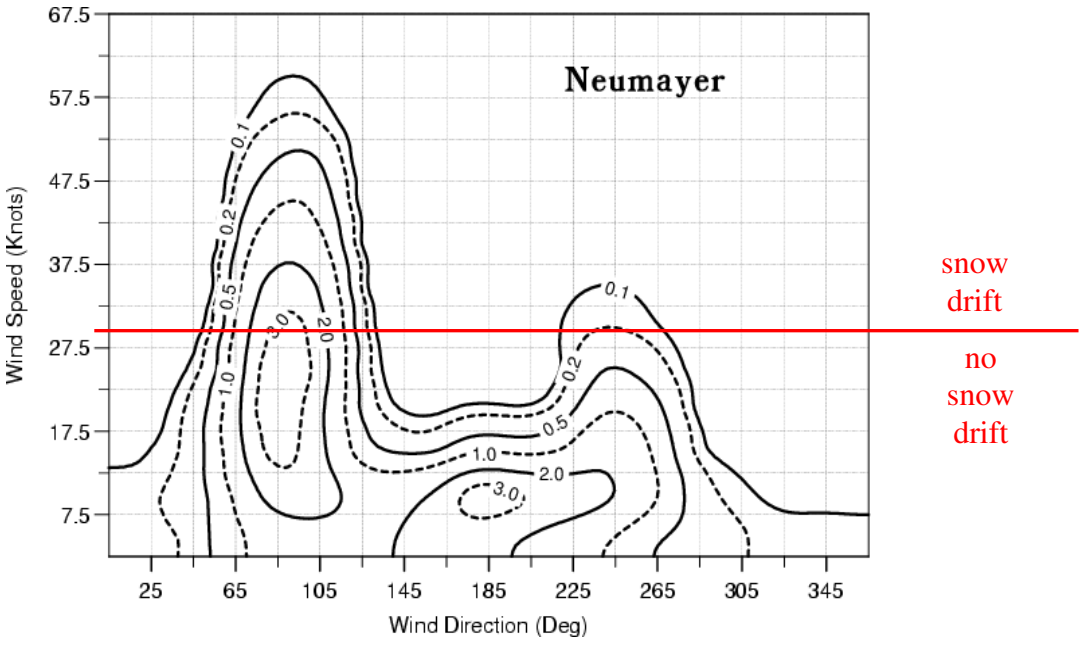


Figure 2.3: Direction (abscissa in degrees) of mean wind speeds (ordinate in knots[2.2]) and frequency of occurrence (contour lines in percent) of winds at Neumayer Base, measured at 10 m height. (Source: Dr. Gert König Langlo, Department of Meteorology, Alfred Wegener Institute)

[2.2] 1 knot equals 0.51 m s^{-1}.

2.1.1.3 Temperature

The mean temperatures measured at Neumayer Base are given in Table 2.2.

Table 2.2: Monthly averaged temperatures [°C] measured at Neumayer Base. All temperatures are given on a monthly time basis as a result of a 14 year observation period. Mean = mean temperature, MeanMax = maximum mean temperature, MaxMax = maximum temperature, MeanMin = minimum mean temperature and MinMin = minimum temperature. (Source: Dr. Gert König Langlo, Department of Meteorology, Alfred Wegener Institute)

Month of year	*Mean temp [°C]*	*MeanMax temp [°C]*	*MaxMax temp [°C]*	*MeanMin temp [°C]*	*MinMin temp [°C]*
1	-4.1	1.2	4.3	-16.4	-23.8
2	-8.1	-0.1	3.6	-21.6	-26.5
3	-12.7	-2.4	1.1	-29.2	-33.0
4	-17.6	-4.2	0.8	-35.1	-39.1
5	-20.5	-6.1	-0.2	-38.3	-44.3
6	-22.1	-7.5	-3.1	-39.6	-44.7
7	-24.0	-9.7	-3.9	-41.4	-45.6
8	-24.9	-10.6	-3.9	-41.8	-47.3
9	-23.1	-9.5	-2.6	-40.8	-45.9
10	-18.1	-6.3	0.8	-34.5	-42.6
11	-10.1	-1.6	1.5	-25.2	-32.6
12	-4.8	0.6	2.8	-17.0	-24.2

Temperature range is important when selecting the proper electronic components for the acoustic observatory. Table 2.2 shows that especially during the winter months (June to September) the mean temperatures drop below -22°C. In addition mean minimum temperatures range between -39.6°C and -41.8°C.

2.1.2 Ice and oceanographic conditions

2.1.2.1 Glaciological and oceanographic processes at the Antarctic coast

Figure 2.4 depicts the predominant glaciological and oceanographic processes of the Antarctic coastal ocean and ice shelf.

The Antarctic continent is covered by an ice sheet (i.e. glacier) which reaches a thickness of more than 4000 m in central Antarctica. The entire glacier moves (caused by gravity) towards the coast and into the ocean, predominantly along so called ice streams. At the grounding line, the ice sheet detaches from the sea-floor and starts to float. This floating part of the glacier - which features a thickness between 500 and 1000 meters at the grounding line (in the Dronning Maud Land Area) and several tenths of meters at its oceanic edge (Steinhage *et al.* 1999) - is called the ice shelf. Parts of this floating ice shelf regularly break off, a process termed calving. Both small chunks of ice and icebergs measuring many hundreds of square kilometers are formed. The icebergs continue drifting with the predominant current into warmer regions where they melt.

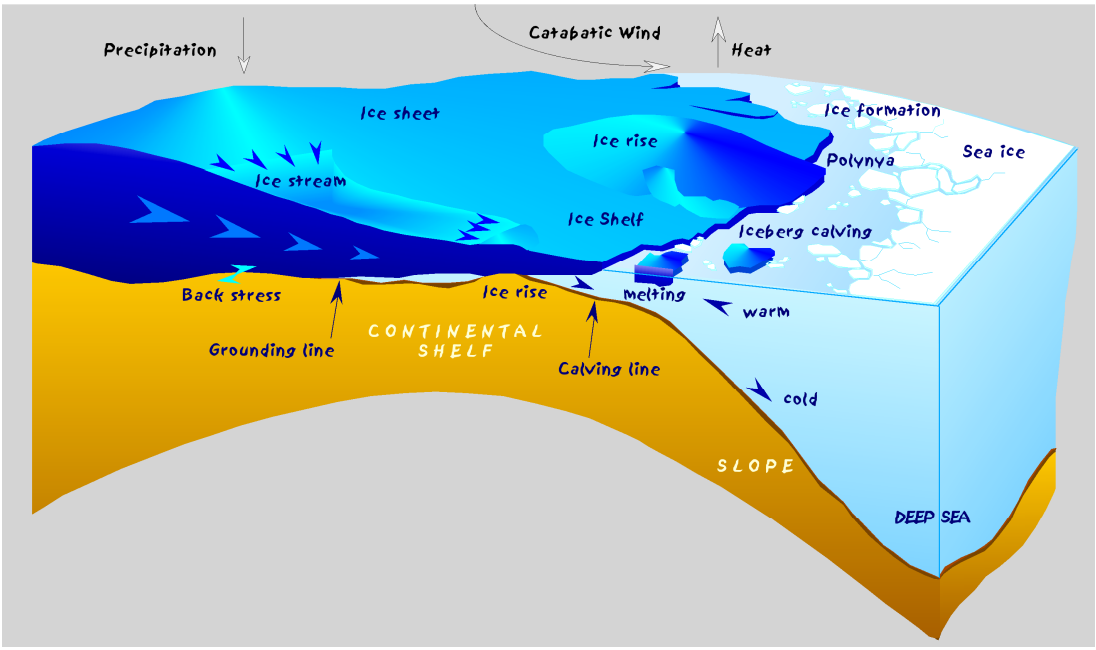


Figure 2.4: Sketch of the Antarctic coast with glaciological and oceanographic processes. Terms are described in the main text. (Source: Dr. Hannes Grobe, Department of Geoscience, Alfred Wegener Institute)

The ocean area directly adjacent to the ice shelf edge is most important for the formation of sea ice. Cold offshore winds from the high continental plateaus (so called catabatic winds) push the sea ice offshore. Within the resulting area of open water, known as the coastal polynia, intense air-sea exchange of heat leads to cooling of surface waters and formation of new sea ice. As a result, water of high density (low temperature and high salinity) is formed which contributes to the formation of Antarctic bottom water (Fahrbach & Rohardt 2008).

2.1.2.2 Currents, ocean temperature and sound velocities

The tidal current in the Atka bay region is characterized by two peaks per day superimposed by the Antarctic coastal current (Buxhoevenden *et al.* 1992). The water temperature near the sea surface varies throughout the year between -1 ° and -2 °C, which implies that the near-surface in-water sound velocity remains nearly constant throughout the year at around 1420 m s^{-1}.

2.2 PALAOA: design drivers and approach

To get an overview of the seasonal variation of the vocal activity of a target species, continuous long-term recordings covering all seasons are essential. This requires both year-round energy supply and access to the water body. For the power supply of PALAOA, only renewable energy sources were considered for environmental and logistic reasons, namely solar panels, wind energy and fuel cells. The previous sections showed that the necessary amount of power cannot be produced year-round by solar panels solely. In polar summer, solar radiation suffices to produce the necessary amount of power exclusively with solar panels. Wind generators are particularly suited to austral winter, as high wind speeds can be expected. On the other hand, during austral winter all

electronic equipment and installations have to endure the low temperatures of up to -45°C. The most severe impact of these low temperatures is their influence on the battery capacity: Below -30°C the capacity of a standard battery type will be less than 10%. Multiyear, continuous access to the water body is just as challenging. Long term hydrophone deployments over the ice shelf edge or through the sea-ice would be threatened by calving of the ice shelf or passing icebergs, while autonomous recording units moored at the sea-floor would not provide real-time access and are subject to damage by passing icebergs. Thus, the safest location for PALAOA appeared to be on the ice shelf proper. A position close to the ice shelf edge would thereby maximize the reception probability of vocalizations of marine mammals migrating within the coastal polynia. However, as the ice shelf steadily advances and breaks off by an average 100-150 m each year, a location too close to the edge would increase the risk of loosing the station due to break-off.

a) Glaciology

To determine the optimal station location, results from previous airborne radio-echo sounding surveys were consulted. These data suggested an ice thickness between 80 m and 200 m on the ice shelf north of Neumayer Base would be ideal. Satellite interferometric imagery (see Figure 2.11) indicated that the northernmost protrusion of the ice shelf exhibits little shear, a favorable condition for long term stability without putting stress on the cables inside the ice.

b) Logistic constraints

The observatory was also expected to need regular maintenance, particularly during its first few years of operation. As a consequence PALAOA was located within a radius of 2 hours drive with a ”Pistenbully” snow crawler from the Neumayer Base. In addition, to the observatory could not be too far from Neumayer Base because of the limited range of the wireless local area network (WLAN) connections. To transmit the data from the ice shelf edge to Neumayer Base (~15 km) a radio link has to be established. The topography of this area is flat and the stations are within sight. This allows for a p2p (point to point) connection without the need of a relay station. Real-time access is particularly important during the development phase, as this region is accessible for external visitors only during austral summer (November – February). Real-time access allows the analysis of the data (quasi) instantaneous year-round without the need of retrieving physical data storage units. Furthermore the real-time connection permits continuous monitoring of the station and to send out a repair team in case of malfunctioning. Development of PALAOA was an iterative process spanning the period between 2004 and 2007, within which three Antarctic expeditions were conducted (austral summer 2004/2005, 2005/2006 and 2006/2007). The detailed stepwise development process of PALAOA follows. For a complete understanding, this description is based on the work conducted in 2005, 2006 and 2007 as the results of each development stage provided input to the design goals and layout of the next stage (see Table 2.3).

Table 2.3: Overview over the PALAOA project[2.3]. Table columns indicate the three development stages. Table rows refer to design goals, design layout, field work and results and are further subdivided according to the various modules comprising PALAOA: AM = Acoustic Module; EN = Energy Module; COM = Communication Module; AS=Additional Sensors.

Timeline	**PALAOA 05**	**PALAOA 06**	**PALAOA 07**
2004	**2005**	**2006**	**2007**

Stages:		***PALAOA 05***	***PALAOA 06***	***PALAOA 07***
Design goals:	**AM:**	• Record internal ice shelf sounds.	• Record underwater soundscape.	• Multi-channel recordings. • Possibility to synchronize with external recording units. • Calibration of the system.
	EN:	• Produce energy with wind and solar. • Implement control system.	• Provide PALAOA 50% of time with power. • Reduce internal electronic noise. • Improvement of the control system.	• Provide PALAOA 75% of time with power. • Further reduce internal and environmental electronic noise. • Energy balance.
	COM:	• Establish data connection to Neumayer Base.	• Increase bandwidth of WLAN. • Establish real-time connection to AWI.	• Establish backup/low-power WLAN connection between PALAOA and Neumayer Base. • High quality stream.
	AS:	• Video observation of Atka Bay.	• Measure tidal currents. • Measure movement of PALAOA.	• Repair CTD cable. • Incorporate weather data into control system.
Design layout:	**AM:**	• Single hydrophone buried in the firn. • A/D conversion using Barix Instreamer.	• Signals of four hydrophones with amplifier / filter routed to PC with external high-quality soundcard. • Develop AsioRecorder	• Set up an additional station on sea ice to increase range of the acoustic module. • Substitution for failed hydrophones.
	EN:	• Savonius wind generator combined with solar panels. • Control system based on a Barix Barionet.	• Extend energy module with methanol fuel cell. • Integrate a control system based on a Barix Barionet. • Galvanic isolation between main electronics and audio module.	• Galvanic isolation of power supply and audio path of each hydrophone. • Installation of second wind generator. • Installation of more capable batteries. • Current hall sensors.
	COM:	• WLAN P2P connection between PALAOA and Neumayer Base.	• Extend communication module with high efficient antenna and booster. • Configure existing Neumayer satellite connection for real-time access.	• Additional WLAN dish antennas. • Further development of the AsioRecorder. • New WLAN hardware.
	AS:	• Webcam with sensitive sensor for poor light conditions.	• CTD sensor. • GPS sensor.	• Establish interface to meteorological data base.

[2.3] ”PALAOA project” is the synonym for the development process of the observatory while ”PALAOA” means the observatory itself.

Field work: AM = Acoustic Module EN = Energy Module COM = Communication Module AS= Additional Sensors	AM:	• Ice shelf hydrophone and box with electronics borrowed into ice.	• Seismic survey for evaluating the water depth and thickness of the ice shelf. • Hot water drilling operation for deployment of the hydrophones.	• Established PALAOA-S(atellite) on the sea ice.
	EN:	• Solar panels and wind generator fixed at a 6-m mast.	• Merged 6-m mast and 10 ft container. • Installation of methanol fuel cell. • Installation and configuration of control system.	• Maintenance: cleared “sastrugies” around container. • Changed wiring. • Installed additional wind generator and new methanol fuel cell.
	COM:	• WLAN technique installed at PALAOA and Neumayer Base.	• Installation of a WLAN booster at PALAOA.	• Installation of additional dish antennas at PALAOA and Neumayer.
	AS:	• Webcam assembled on top of the mast.	• Deployment of CTD through ice shelf. • Installing GPS antenna on top of container.	• Fixed cable breakage of the CTD.
Results: AM = Acoustic Module EN = Energy Module COM = Communication Module AS= Additional Sensors	AM:	• Electronic noise dominated acoustic recordings.	• Deployment of hydrophones successful. • Measuring with single hydrophone because of “sferics” noise.	• Streaming of real-time stereo signals to AWI. • PALAOA-S provided additional data and increased range of the acoustic module.
	EN:	• Energy module worked smoothly. • Energy shortage in winter. • Barix Barionet worked reliably.	• Energy control system allowed controlling the station from AWI. • Methanol fuel cell didn’t work at low temperatures. • Energy shortage in winter.	• “Sferics” noise reduced to a minimum. • Uptime significantly increased. • Control system optimized. • Methanol fuel cell didn’t work at low temperatures. • Energy shortage in winter.
	COM:	• WLAN connection stable (bandwidth: 1.3 Mbit/sec).	• Bandwidth of system increased to 3 Mbit/sec. • Data connection between AWI and PALAOA established.	• Backup/low-power connection established. • Stereo stream established. • High Quality multi stage buffering system.
	AS:	• Webcam pics/movies provided useful information of ice coverage of Atka Bay and weather conditions.	• CTD provided information on tidal current. • GPS measured movement of the station container.	• Weather data from Neumayer meteorological observatory fed into control system.

Table 2.3 summarizes for each development stage the corresponding design goal and layout, field work and results. A detailed discussion of these aspects is presented in the subsequent sections.

2.3 PALAOA 05

2.3.1 PALAOA 05 – design goals and layout

The electronic components of the acoustic, energy and communication modules was assembled into a Zargesbox (80 x 60 x 60 cm) which was placed in a trench about 1 m deep and covered with a plywood panel and snow.

2.3.1.1 PALAOA 05 – design goals and layout: Acoustic Module (AM)

The design goal for the acoustic module was to record shelf ice sounds which can be used to test the radio link between PALAOA 05 and the Neumayer Base. For this reason a Reson TC4033 hydrophone was selected because of its robustness and low cost. The hydrophone signals were amplified (50 dB) and filtered (1 Hz high-pass filter) with an Etec A1101 amplifier. A Barix Instreamer was used as an A/D converter. This is a small embedded Linux device with two analogue audio inputs and a network-compatible Ethernet interface. The Instreamer allows sampling of the acoustic signals at 48 kHz/16bit. The digitized data are mp3 encoded and streamed with 128 kBit via the Ethernet interface. A small self programmed C-routine, running on a desktop PC located at Neumayer Base, stored the streamed data in files of 1 minute length on a local hard disc.

2.3.1.2 PALAOA 05 – design goals and layout: Energy Module (EM)

The design goals for the energy module were (1) to supply PALAOA 05 with power generated by solar and wind and (2) to implement a rudimentary control system to observe the battery voltages. The energy module of PALAOA 05 had a solar module consisting of four solar panels (charging controller Votronic SR 300) and a Savonius wind generator type WS 0,30A (charging controller WGU-22), produced by the Finnish company Windside. A schematic wiring diagram is given in Figure 2.5. The energy module generated a main voltage of 24 V (operating voltage of main electronic components). The energy generated by the module was used to charge four Dryfit gel batteries, each with a capacity of 130 Ah. The consumer load was connected to the batteries via a battery guard to avoid deep discharge of the batteries.

The station was controlled with a Barix Barionet, a small embedded Linux device with several digital and analogue in- and outputs and one Ethernet and RS232 interface. The analogue inputs were used to monitor the voltage of the battery pack and the temperature inside the Zargesbox. With the analogue inputs of the Barionet covering a range of 0 - 5 V, measuring transducers had to be used to transform the observed voltages (0 - 40 V) and temperatures (-40° - +40°C) to a 0 - 5 V output signal. The digital inputs remained unused. All data were digitized by the Barionet and streamed via the Ethernet interface to Neumayer Base.

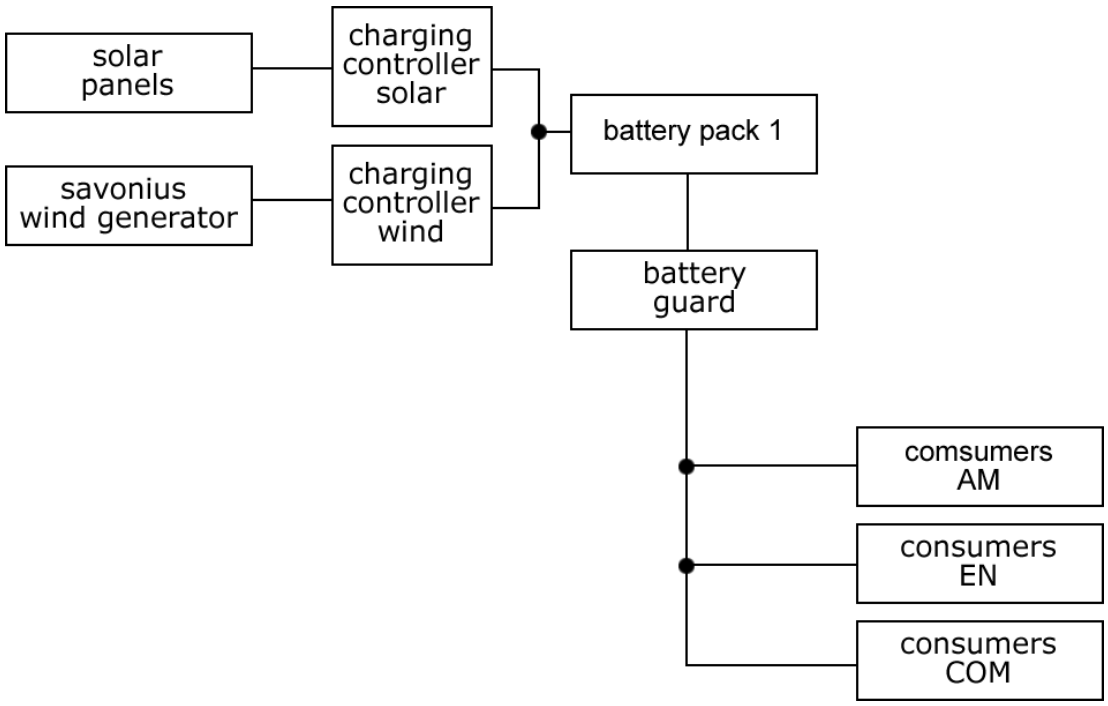


Figure 2.5: Schematic wiring diagram of PALAOA 05.

2.3.1.3 PALAOA 05 – design goals and layout:

Communication Module (COM)

The design goal for the communication module was to set up a radio link between PALAOA 05 and Neumayer Base. This radio link was established with WLAN technology. Two highly directional YAGI antennas with a gain of +18 dBi and two Cisco 350er series bridges were used. To increase the bandwidth of the WLAN link, an additional WLAN booster was installed at Neumayer Base. However, because of the booster's high power consumption (~ 25 Watt), no booster was installed at PALAOA 05. In contrast to later stages of PALAOA, no continuous real time connection to Germany was attempted at this stage. However, the Instreamer could be reached from the internet and live sound could be transmitted for a limited time until the satellite link was saturated because of data transmission by other users with higher priority.

2.3.1.4 PALAOA 05 – design goals and layout: Additional Sensors (AS)

The design goal for the additional sensors was to gain visual observations of the ice conditions in the Atka Bay (see Fig. 2.12). For this task a Mobotix webcam type M10 day/night was selected. This webcam is equipped with two built in CCD sensors (color and b/w). The camera switched automatically to the more sensitive b/w CCD sensor during limited light conditions. The webcam was set up to take a picture every 10 minutes with a resolution of 1.3 Megapixels.

2.3.2 PALAOA 05 – field work (January 2005 – February 2005)

The Zargesbox which contained the electronic components of the acoustic and energy module was prepared in Bremerhaven in advance to the expedition. The

assembling of PALAOA 05 on the Antarctic ice shelf was conducted by the overwintering team of Neumayer Base. During the Polarstern cruise ANT XXII/3, a three day visit to Neumayer Base in early February 2005 was used to put PALAOA 05 into operation.

2.3.2.1 PALAOA 05 – field work: Acoustic Module (AM)

The hydrophone was buried in the firn, approximately 4 meters below the surface. The Etec A1101 and the Barix Instreamer were placed into the abovementioned Zargesbox.

2.3.2.2 PALAOA 05 – field work: Energy Module (EM)

The Savonius wind generator and the four solar panels (facing to the North) were mounted on a pylon of 6 m height (see Figure 2.6). The charging controllers, the batteries and the Barix Barionet were placed together with the acoustic equipment into the Zargesbox.

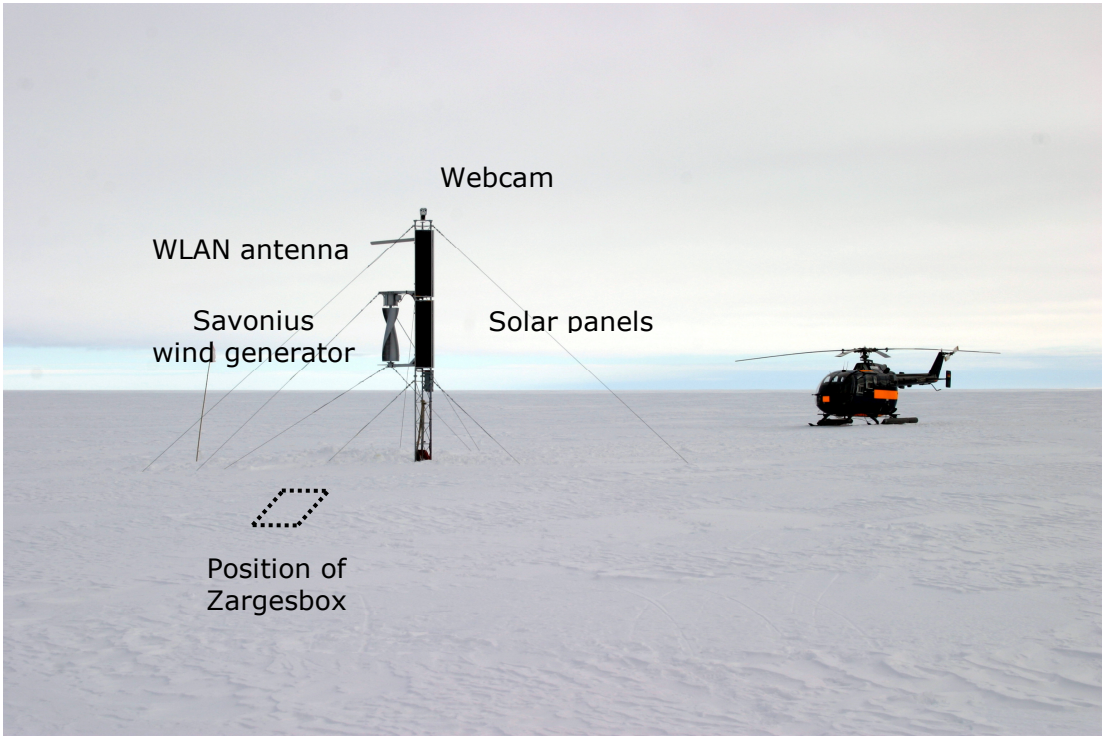


Figure 2.6: PALAOA 05 with components for energy production and communication. The location of the snow covered Zargesbox that contained electronics and batteries is indicated by the dotted line.

2.3.2.3 PALAOA 05 – field work: Communication Module (COM)

The WLAN antenna was fixed near the top of the pylon. The WLAN bridge was assembled with the other electronic equipment in the Zargesbox. A second WLAN antenna, a WLAN bridge and a booster were installed at Neumayer Base. An Ethernet cable was run to the electronic lab and connected to a Desktop PC.

2.3.2.4 PALAOA 05 – field work: Additional Sensors (AS)

The webcam was installed on the very top of the pylon and oriented in N-NE direction. A splitter was used to merge the power and the data cable of the webcam (power over Ethernet), which simplified the installation, as only one cable had to be installed to the top of the pylon.

2.3.3 PALAOA 05 – results and discussion

2.3.3.1 PALAOA 05 – results and discussion: Acoustic Module (AM)

Shelf ice recordings were obtained and transmitted to Neumayer Base for further analysis. However, the acoustic recordings were severely compromised by electronic noise. After a test in the lab the two charging controllers of the wind generator and the solar panels could be identified as the noise source. The problem was undetectable during prior tests in the lab, as the ambient electromagnetic noise in the lab prohibited the high amplification settings used in Antarctica.

2.3.3.2 PALAOA 05 – results and discussion: Energy Module (EM)

The wiring diagram of PALAOA 05 is given in Figure 2.5. The energy production worked quite effectively during austral summer. Because of the high amount of sunlight, the solar panels generated enough energy to power the station. During austral winter, when the station was powered by wind, the station dropped out during calm periods from time to time. Uptime of PALAOA 05 during austral winter (March to September) was around 65%. The total power consumption of the station averaged 25 Watt.

The temperature inside the Zargesbox varied between -15°C and +35°C. A comparison between the temperature in the Zargesbox and atmospheric temperatures are given in Figure 2.7. The difference between the temperature in the Zargesbox and the ambient temperature outside (measured at Neumayer Base) was often more than +10°C. Recalling that the Zargesbox (all-aluminum) was buried 1 m below the surface indicates that the observed temperature changes are not a direct consequence of ambient temperatures. Rather, the highest temperatures occurred during periods of an energy surplus, when significant amounts of energy were transformed into heat by the charging controllers. Because of the good isolation in the ice, the Zargesbox did not cool down to the lowest atmospheric temperatures of -35°C. Thus the operating time of PALAOA 05 was increased because of a higher battery capacity, which is strongly influenced by temperature. The Barix Barionet worked most reliably. At Neumayer Base all operating data were stored using a Matlab™ routine in a log file.

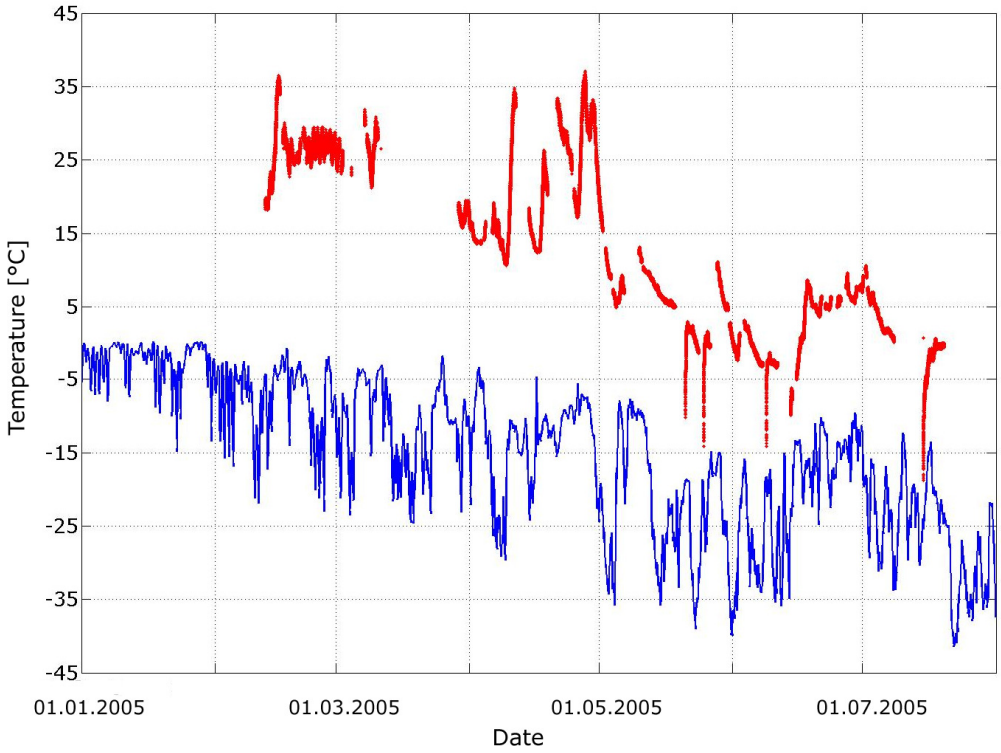


Figure 2.7: Temperature inside the Zargesbox (red) and atmospheric temperatures (blue).

2.3.3.3 PALAOA 05 – results and discussion: Communication Module (COM)

The test of the WLAN link was successful. All components withstood the harsh climatic conditions and the bandwidth of the WLAN link was determined at 1.2 – 1.3 Mbit/s. The link was rather stable and the connection was interrupted only when insufficient power was available to operate PALAOA 05. Quite surprisingly, no decrease in the WLAN bandwidth was noticed during storms (with or without snow drifts). However, the achieved bandwidth of the WLAN link is too small to permit the transmission of several high-quality uncompressed hydrophone signals to Neumayer Base.

2.3.3.4 PALAOA 05 – results and discussion: Additional Sensors (AS)

The Webcam survived the harsh climatic conditions. The pictures and movies taken provided useful information on the sea ice conditions in Atka Bay. The images captured the swaying of the floating ice shelf, providing the first hints of the tidal current. Furthermore, the presence/absence of the coastal polynia was visible in the webcam pictures.

2.4 PALAOA 06

2.4.1 PALAOA 06 – design goals and layout

Inside the Zargesbox, the space to house the electronic modules of PALAOA 05 was severely limited. Furthermore, whenever direct access to the electronics was necessary for testing or modification purposes, the Zargesbox had to be

"unearthed", which proved to be a tedious and time consuming affair. Hence a 10 ft isolated container was selected to house all electronics of PALAOA 06 and to serve as a mounting platform for the energy producing components and antennas. The container was fixed on a 20 ft sledge. If the station moved too close to the ice shelf edge, all of the equipment - except the hydrophone cables in the ice, the hydrophones and the CTD sensor - could be hauled to a safer position or even Neumayer Base with a "Pistenbully" snow crawler.

2.4.1.1 PALAOA 06 – design goals and layout: Acoustic Module (AM)

The design goals for the acoustic module were to gain continuous and broadband (frequency range up to 192 kHz) recordings of the Antarctic underwater soundscape – at least from a single hydrophone. To meet this goal, deployment of a single hydrophone would have sufficed. However, recognizing the considerable mobilization and demobilization efforts and availability of technical staff for the necessary hot water drilling (see below) it was decided that the array of 4 hydrophones planned for the final stage should be installed during this season. The array consisted of 3 active[2.4] hydrophones (Type Reson TC4032) and one passive[2.5] hydrophone (Type Reson TC4033). The advantage of an active hydrophone is the ability to transmit signals over long distances (up to some hundreds of meters) with reduced sensitivity for electromagnetic interference and lower capacitive attenuation of higher frequencies. In contrast the passive hydrophone also allows the emission of sounds and can therefore be used for playback experiments. Each hydrophone was connected to a Reson VP2000 amplifier/filter. Signals of each active TC 4032 hydrophone were passed through a Reson input module to a Reson VP2000 amplifier/filter, while the passive hydrophone was directly connected to its VP2000 amplifier/filter. The EC 6073 input modules serve to provide power to the active hydrophones' built-in pre-amplifiers. The output of each Reson VP2000 was routed into two different digitizing systems[2.6] (compare Figure 2.8).

One digitizing system consisted of a local, industrial PC (operating system: Windows XP) with an external FireWire soundcard, MOTU Traveller. This enabled sampling of all four hydrophone signals up to 192kHz/24bit. The MOTU data were recorded on the PC with the AsioRecorder software (programmed in C) developed in cooperation with the University of Kaiserslautern, Germany specifically for the PALAOA project. The AsioRecorder is a small and stable recording software supporting Asio multi channel driver. The recorded high-quality data (mono) were stored locally at PALAOA and retrieved daily via FTP (file transfer protocol) from PALAOA 06 to Neumayer Base. This system consumed about 53 Watt in total, which was likely to exceed the power limitations of the energy module during wintertime.

[2.4] Equipped with a built-in preamplifier.
[2.5] Not equipped with a built-in preamplifier.
[2.6] Low-quality system: provides compressed data with reduced frequency range (20 Hz – 15 kHz @ 16 bit); High-quality system: provides uncompressed or loss-less compressed data with frequency range up to 96 kHz @ 24 bit.

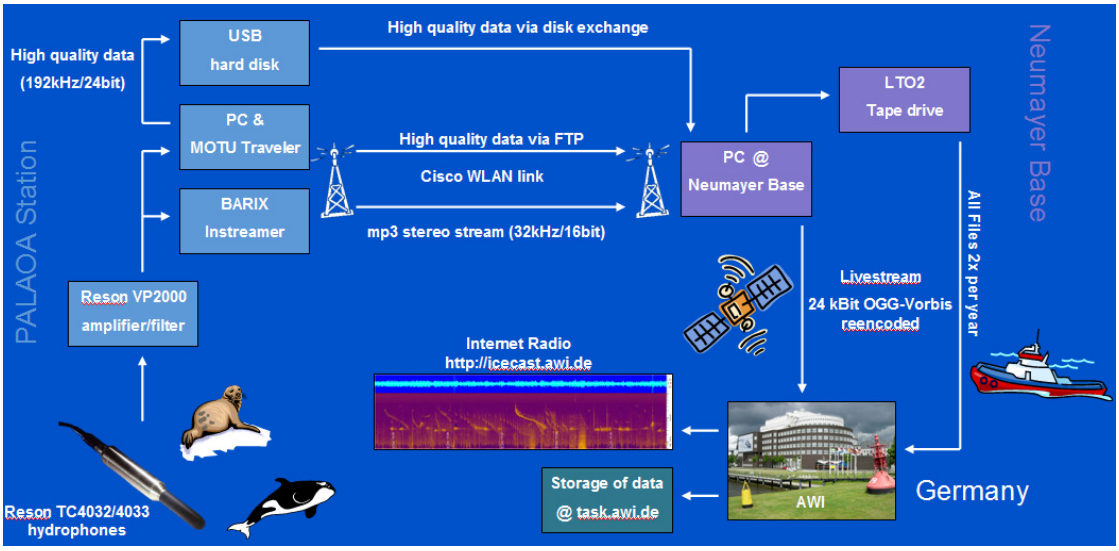


Figure 2.8: Schematic of the PALAOA 06 audio and communication module.

For this reason the audio module of PALAOA 05 was maintained as a second, alternative digitizing system. Signals of one active hydrophone were routed into the Barix Instreamer, which provided an mp3 stream of reduced frequency range (20 Hz - 15 kHz @ 16 bit). This approach minimized the power consumption of the entire system to about 30 Watt. When the low quality system was active, the PC and the external soundcard were remotely disconnected from the power supply.

2.4.1.2 PALAOA 06 – design goals and layout: Energy Module (EM)

The design goals for the energy module were (1) to reduce the electronic noise in the acoustic recordings to minimum and (2) to provide PALAOA 06 at least 50% of time with power. In the course of the upgrade of the acoustic module, the energy demand increased from 25 W (PALAOA 05) to 53 W (PALAOA 06). To provide the station with the necessary amount of power, the number of solar panels was doubled. Furthermore the number of batteries was increased from 4 (PALAOA 05) to 24 (PALAOA 06) to buffer periods without input from the solar panels or the wind generator. However, during austral winter, system drop-outs[2.7] were expected because of the increased energy demands in the course of the extension of PALAOA. A third energy source - independent of the weather conditions - appeared mandatory to minimize the number of drop-outs. A methanol fuel cell, type SFC A50R produced by the German company SFC Smart Fuel Cell AG, Munich, Germany and designed for temperatures as low as -35°C was integrated into the energy module and tested as an addition to solar and wind energy sources. It should – at least theoretically – be capable of providing sufficient energy during periods of both darkness and calm winds. This fuel cell, actually the first commercially available methanol fuel cell, unfortunately produced a voltage of 12 V, which necessitated using a DC/DC changer to transform the output voltage up to the stations operational voltage of

[2.7] A drop-out means: The PALAOA observatory cannot be accessed via remote control and no acoustic data are transferred to Neumayer Base. Drop-outs can be caused by various reasons (e.g. lack of energy or defective electronic components).

24 V. This caused a power loss of about 15% (efficiency of the DC/DC changer: 85%).

Based on its flawless performance in PALAOA 05, the Barix Barionet device was again selected as the main control unit. The Barionet was extended with several relay-cards to realize a galvanic isolation between the main electronic and the audio module and to be able to remotely switch on and off several devices. A galvanic isolation between the power supply of the EN/COM module and the AM module was necessary to reduce the electronic noise in the audio recordings. For this reasons, the 24 Dryfit gel batteries (12 V) with a capacity of 130 Ah were grouped into 6 battery packs. The freely programmable (BASIC dialect) Barix Barionet and four additional relay cards (six relays on each card) enabled the galvanic separation of the modules. A schematic wiring diagram of PALAOA 06 is given in Figure 2.9.

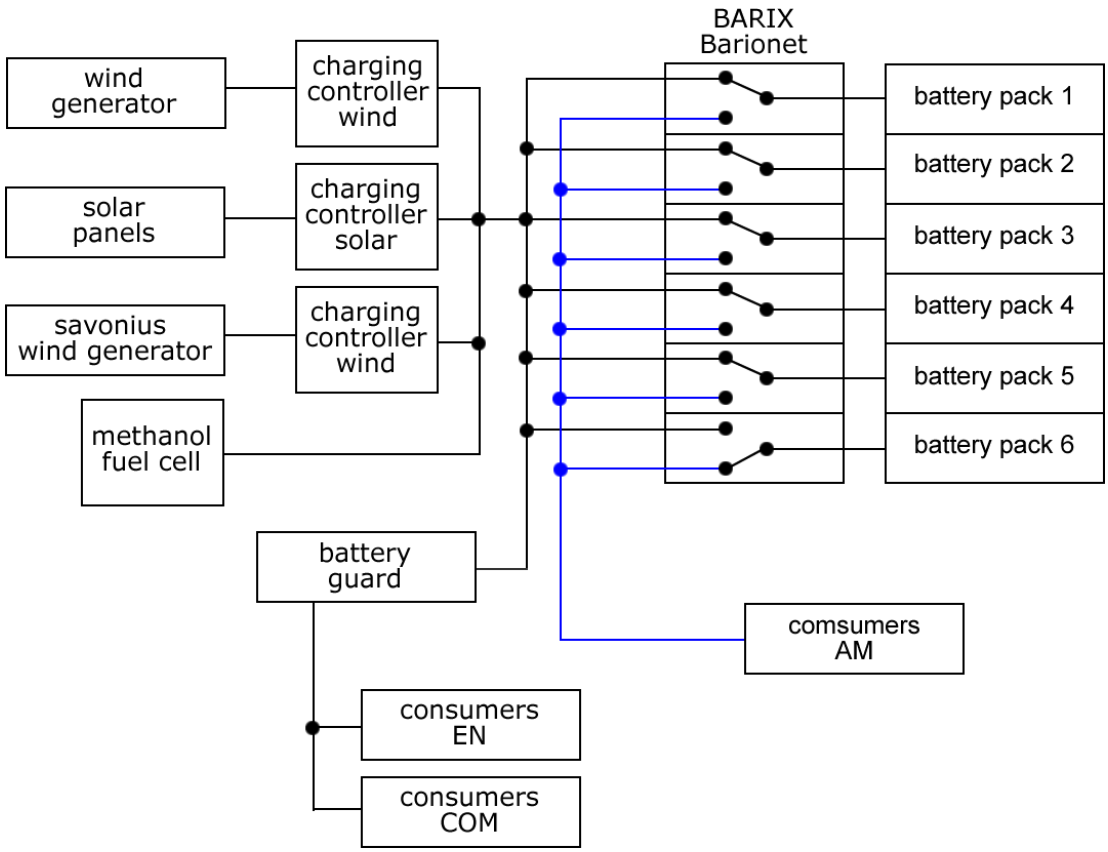


Figure 2.9: Simplified schematic wiring diagram of PALAOA 06.

The positive and negative poles of each battery block were connected to a relay input. Each relay has two outputs (toggle switchable) available. One output was connected to the audio module and the other to the energy module. This configuration allowed connecting every battery block to either one of the two modules. During operation, only one battery pack was used to provide the audio system with power. The other five battery packs were used to power main electronic system (EN and COM module) while simultaneously being charged. After a time interval of 10 minutes, one of these five packs was disconnected from the charging system and connected to the audio system. Then the

previous battery pack of the audio system was switched off and connected to the charging circuit again.

The battery voltage of each electric circuit was monitored using the Barix Barionet analogue inputs. With the analogue inputs of the Barionet covering a range of 0 - 5 V measuring transducers had to be used to transform the observed voltages (0 - 40 V) to the 0 - 5 V range. The power supply of the measuring transducers were designed to be galvanically isolated from their in- and outputs. This allows monitoring the voltage of the battery pack without producing a bypass of the galvanically isolated power lines.

2.4.1.3 PALAOA 06 – design goals and layout:

Communication Module (COM)

An acoustic livestream was expected to greatly facilitate the development of automated real-time analyses of the PALAOA 06 recordings in Germany within the time frame of this thesis. Also the underwater sounds could be presented to the public via the World Wide Web. Hence the design goals for the communication module were to increase the bandwidth of the WLAN connection between PALAOA 06 and Neumayer Base to allow the transmission of at least one hydrophone signal sampled with 192 kHz @ 24 bit (FLAC coded) and to establish the real-time connection to Bremerhaven. The WLAN link was boosted with a WLAN booster at PALAOA 06, in addition to the one installed at Neumayer Base. Furthermore the observatory was equipped with an IRIDUM modem/phone. The IRIDIUM device features an internal GPS receiver, which allows broadcasting the stations' position at predetermined time intervals, regardless of the WLAN link's availability. In case of a break off of the station's hosting ice shelf segment, PALAOA 06 would send its position to the Alfred Wegener Institute, facilitating the stations' recovery by ship. The IRIDIUM phone was intended to serve as a communication and emergency link with Germany, primarily for the overwintering team. The desired acoustic livestream to Germany was realized with Icecast[2.8] (released under a GPL license), which is frequently used for broadcasting radio transmissions on the internet. Figure 2.10 gives a schematic of the livestream processing chain.

[2.8] http://www.icecast.org

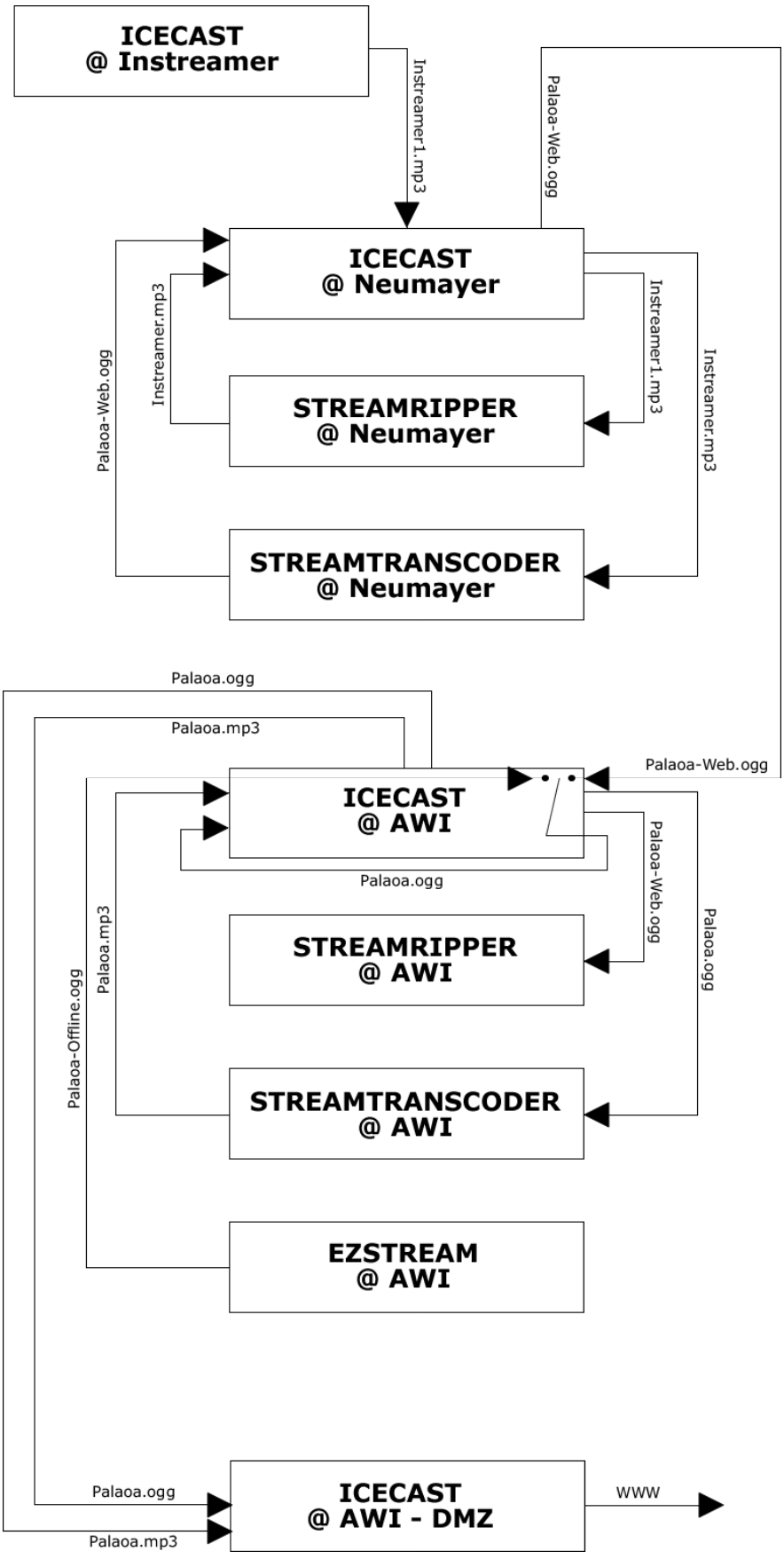


Figure 2.10: Schematic diagram of the livestream process.

The Barix Instreamer is equipped with an Icecast client streaming module. The audio signals of the hydrophones are digitized, mp3-coded and then sent as an Icecast stream (named “Instreamer.mp3”) via the WLAN link to the mount point “Icecast@Neumayer” at Neumayer Base. At Neumayer Base the software Streamripper[2.9] marks the stream with a time stamp and stores the data locally as files of one minute duration and named “yyyymmdd-hhmmss PALAOA Instreamer.mp3”. The time-stamped stream is again sent to the mount point “Icecast@Neumayer”. Another program – the so-called streamtranscoder[2.10] – is fed with the time-stamped stream. This software module resamples the stream to a 24 kBit ogg-stream called “PALAOA-Web.ogg” which is sent to Bremerhaven. The resampling (data reduction) is necessary because of the limited bandwidth of the satellite connection between Neumayer Base and Germany. The central mount point at the AWI is called “Icecast@AWI”. A streamripper module saves the incoming “PALAOA-Web.ogg” stream to a local hard disc in files of one minute duration. At the mount point “Icecast@AWI” a switch was established. As long as the incoming stream is running, the “PALAOA-Web.ogg” is mounted as “PALAOA.ogg”. When no live stream is available (e.g. during breakdown of the WLAN link between PALAOA and Neumayer Base) the Ezstream software[2.11] streams the last recent received files as “PALAOA-Offline.ogg” to the Icecast@AWI mount point.

The “PALAOA.ogg” or “PALAOA-Offline.ogg” stream is sent to a “streamtranscoder” module to create an additional mp3 stream – called “PALAOA.mp3”. The reason for this quality reducing re-transcoding is the fact that MS-Windows does not natively support OGG-vorbis playback, but most internet users can play back the mp3 stream without the need to install additional software. The “PALAOA.ogg” and the “PALAOA.mp3” streams are finally sent to the “Icecast@AWI-DMZ” mount point. This mount point is installed on a computer which is located in the “demilitarised zone” of the AWI providing an open port where the streams are accessible by the public (www.awi.de/PALAOA).

2.4.1.4 PALAOA 06 – design goals and layout: Additional Sensors (AS)

The design goals for the additional sensors were to gain information on the tidal and coastal currents and to determine the drift of PALAOA 06. A Trimble Lassen IQ GPS receiver with a high accuracy and low power consumption was connected to the RS232 interface of the Barix Barionet. The GPS was used to observe the stations’ drift, which is caused by the lateral movement of the ice shelf. In total, five temperature sensors were placed inside and outside the container and the control cabinets. The temperature sensors were directly connected to the Barix Barionet. The data were streamed to Neumayer and there stored in a log file using a Matlab™ routine.

The CTD was configured as a standalone device. The sensor was automatically switched on every half an hour and the data were stored on an internal memory.

[2.9] http://streamripper.sourceforge.net
[2.10] http://www.oddsock.org
[2.11] http://www.icecast.org/ezstream.php

From time to time, a serial connection was initiated to readout the data. For the readout procedure the CTD was connected to the RS 232 Interface of the industrial PC. This RS 232 interface was accessible via the Neumayer Desktop PC using a Windows HyperTerminal connection. The lithium batteries inside the CTD should provide enough power to run the device for a period of 2 years in this sample mode. However, a possibility to provide the sensor with external power exists and was planned to be established during the field season 2006/2007.

2.4.2 PALAOA 06 – field work (December 2005 – February 2006)

The field work was conducted between early December 2005 and late February 2006. The Ocean acoustics team was supported by the AWI's glaciology and logistics department for the seismic study and the hot water drilling operation.

The 05/06 field work commenced with a seismic survey to determine the final position of the hydrophone array which provided detailed information on the ice shelf thickness and sea floor bathymetry under the ice. A small seismic source, SISSY (Seismic Impulse Source SYstem), uses electrically triggered ammunition to generate by explosion a pressure pulse of approximately 1800 hPa (Buness *et al.* 2000) in a small pipe of 5 cm in diameter. With the shot being fired in a closed system, no packing material from the ammunition is left in the firn. Seismic recordings were obtained by a geophone chain of 24 geophones at a spacing of 10 m. Best results were obtained by placing the shot point one line length away from the middle of the geophone chain (see Figure 2.11). Due to the duration of the pulse, the ice proved too thin for a "normal" spread arrangement, with a shot point close by, while the source has insufficient power to allow for larger shot point separations. Assuming an in-ice p-wave sound speed of 3770 m s^{-1} (Kohnen 1974) and 1440 m s^{-1} for sound propagation in sea water, a first estimate of the thicknesses of the shelf ice and underlying water column was obtained for the proposed drill sites. Ice thickness varied in the area of investigation between 105 m and 120 m and the water column thickness under the ice shelf between 150 m to 170 m (see Table 2.4).

Altogether, 45 shots were fired as the location of the geophone chain was changed 10-times. Retrospective comparison of these results with those from the hot water drilling (depth of the central hole of the hydrophone array: 97.3 m), showed the ice to be approximately 10 m thinner than predicted (ice thickness estimate from seismic study: 108 m). This mismatch might have been caused by neglecting the lower velocity of the seismic wave in firn. To match the actual ice thickness with the seismic wave travel times, an average sound speed of 3550 m s^{-1} within the ice shelf has to be assumed. Under this assumption, the water column below the ice is about 160 m deep at all four drilling sites.

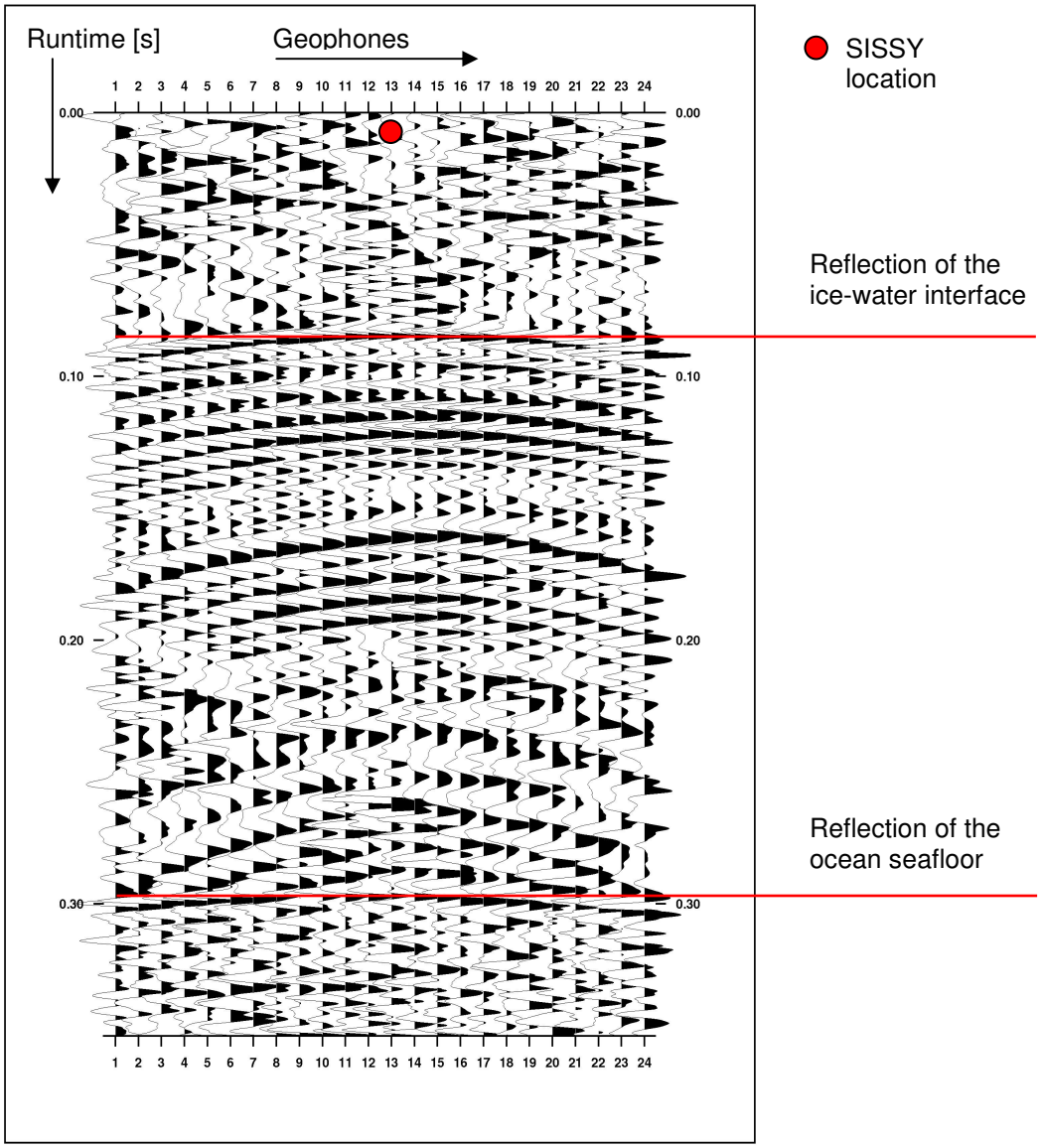


Figure 2.11: Seismogram measured in the vicinity of PALAOA using the SISSY. Graph courtesy of Daniel Steinhage, AWI.

After this seismic survey, PALAOA 06 was finally located at 70°31'S and 8°13'W, close to the shelf edge of the Ekström ice shelf, about 15 km north of the Neumayer Base (see Figure 2.12).

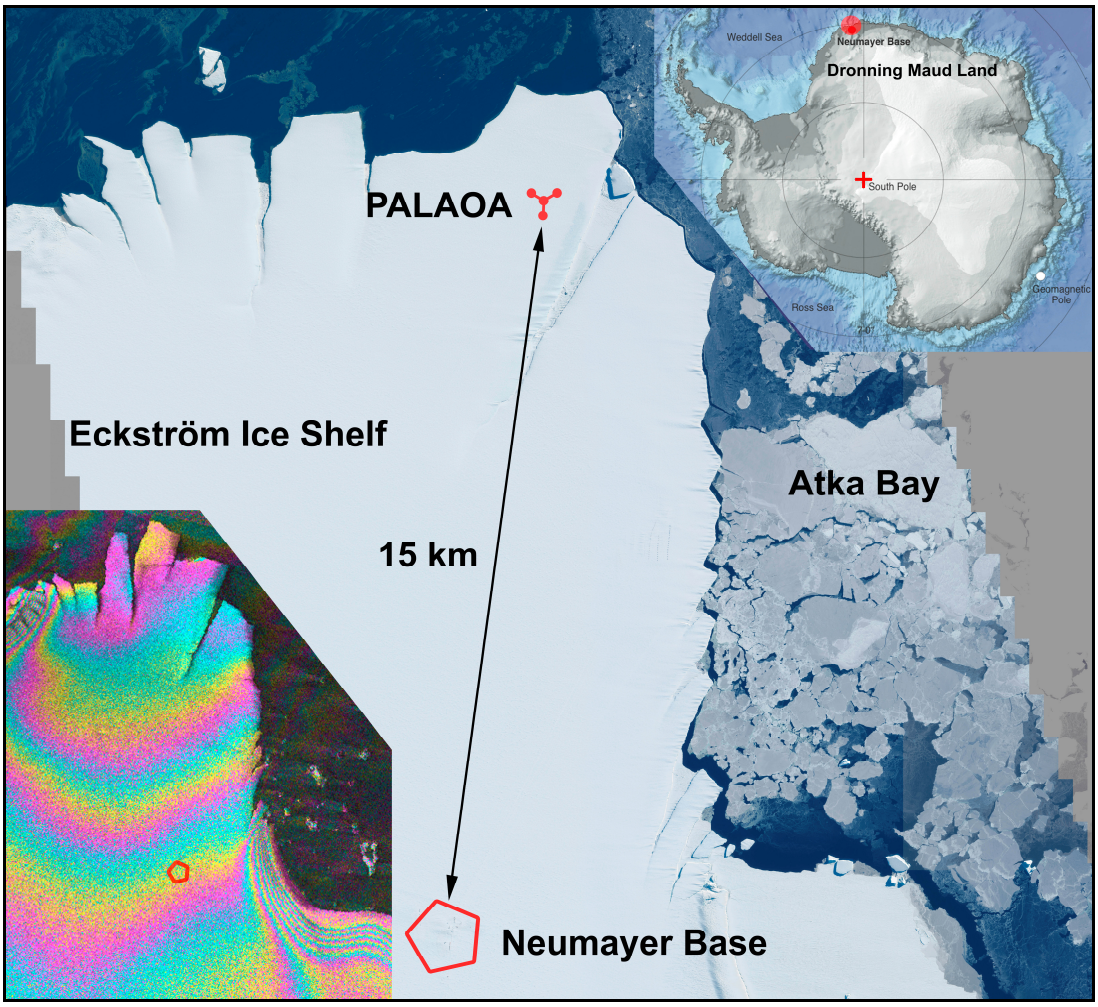


Figure. 2.12: Visual satellite image of the Eckström Ice Shelf with the position of PALAOA Station indicated. The lower left insert gives a satellite interferometric imagery of the same area. Images courtesy of Wolfgang Rack, University of Canterbury, Christchurch, NZ.

2.4.2.1 PALAOA 06 – field work: Acoustic Module (AM)

A control cabinet for the acoustic module was constructed in advance of the field campaign in Bremerhaven (see Figure 2.13). The system had been prepared such that – in theory – only the hydrophones had to be connected.

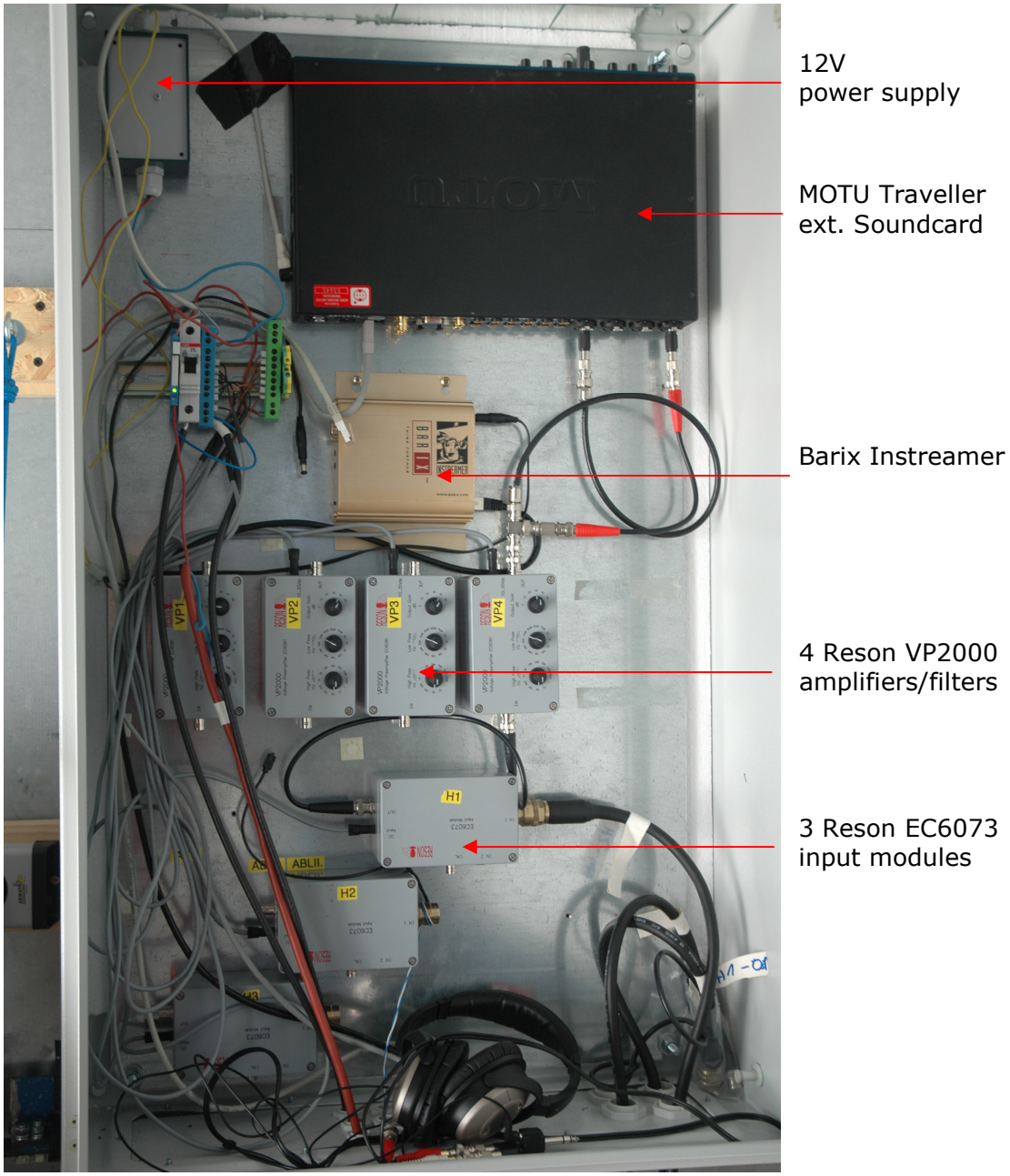


Figure 2.13: Control cabinet audio module.

To deploy the four hydrophones in the water body below the ice shelf, a hot water drilling operation was conducted to penetrate the 100 m thick ice shelf. The hot water drilling system was designed by the Alfred Wegener Institute and first used in January 1993 (Nixdorf *et al.* 1994a; Nixdorf *et al.* 1994b). The drillings for PALAOA 06 were conducted between 5 December 2005 and 10 January 2006 by the AWI's logistic department.

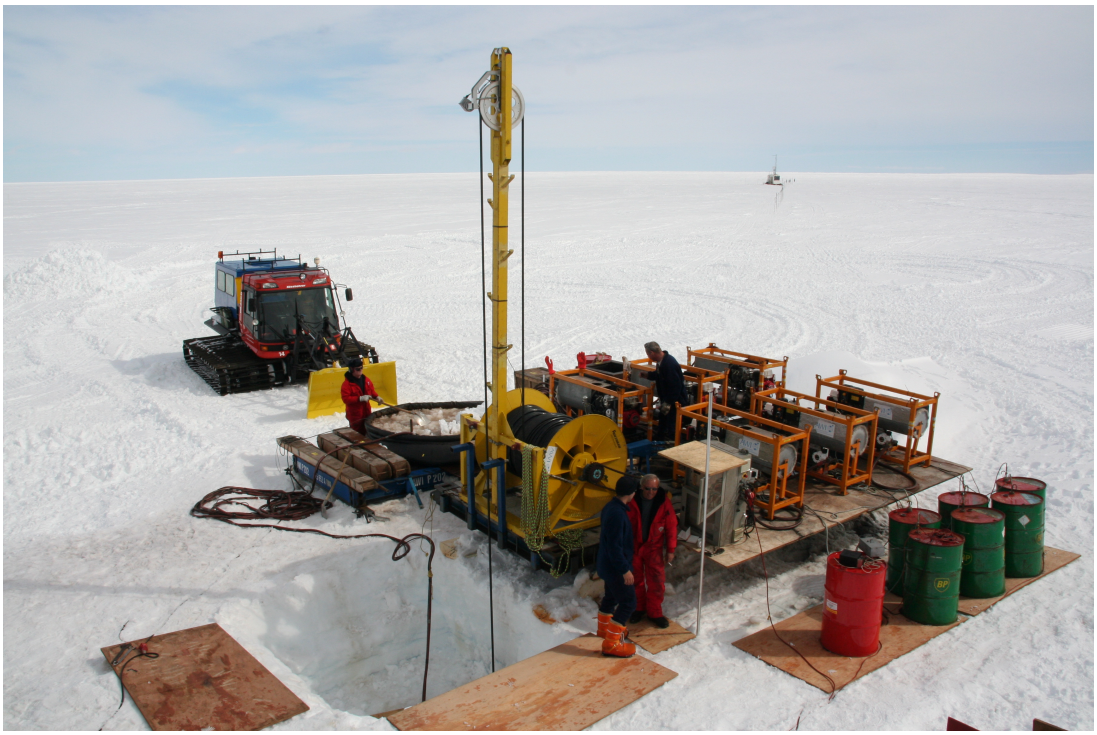


Figure 2.14: Drill site (from left): Pistenbully, water basin, inclinable mast with pressure hose on winch-drum, and six Waps. Right in front: One petrol barrel (red) to fuel the Honda generator modules of the Waps, and six diesel barrels (green) to fuel the heat exchangers of the Waps. Left in front: a 1.5 m deep manhole for safe attachment of splash lances. The hose between manhole and water basin communicates with an immersion pump inside a cavity 30 m below the ice surface (see text).

Figure 2.14 shows the major components of the drilling equipment which was mounted on a 20 ft container-sledge next to the fuel supplies. The sledge based equipment consisted of:

a) Six modified hot water high-pressure washers (Wap, 125 kW),
b) a winch-drum to lower and heave,
c) a pressure hose with the attached splash lance over,
d) a 5 m inclinable mast, and
e) a winch-monitor to control the cumulative hose/lance weight, the drilling speed and the nozzle depth.

A flexible 5 m^3 basin on a separate sledge was used for melting snow. From this basin, six immersion pumps supplied the WAPs with water. The WAPs outputs were connected in parallel to a high pressure valve, which distributed the hot water between the drill lance and the melt basin.

Boreholes were melted through the ice by injecting hot water (90°C) at flow rates of up to 80 l min^{-1} and up to 100 bar through the nozzles of the splash lances. Depending on the melting rate, the lances were lowered into the ice at speeds of between 10 and 40 m h^{-1}. Lances of 60, 120, and 300 mm diameter were used successively to melt and widen the duct of the holes. Holes usually are somewhat wider than the lance diameters, as warm water returns from the

tip of lance along its sides back to the cavern. At each of the four drilling sites, two parallel holes (approx. 1 m apart) were melted into the ice. The first hole was drilled to a depth of 30 m where a cavity of about 11 m^3 was melted into the firn by operating the 120 mm splash lance for 3 hours at 60 l min^{-1} flow rate and 40 bar pressure. The additional water reservoir generated in this cavity was necessary to establish a water recirculation system (see Figure 2.15) which allows continuous drilling.

Once the cavern was prepared, its water was recycled to the melt basin by a seventh immersion pump. This reservoir helped to minimize energy demands for melting snow and hence allows use of the available hot water for the ”drilling” process proper. Thereafter, the second (main) hole was drilled leading through the cavity into the ocean. As soon as the nozzle penetrated the ice shelf at a depth of about 100 m (Table 2.4), sea water rapidly rose into the cavity, providing an unlimited amount of water for further widening the borehole. The lowering of the instrumentation was achieved by a separate winch driving the cable via the inclinable mast exactly into the borehole. For lowering the relatively lightweight hydrophones and the CTD for the central borehole, 20 kg drop-weights were attached to the hydrophones and CTD by means of corrosion links.

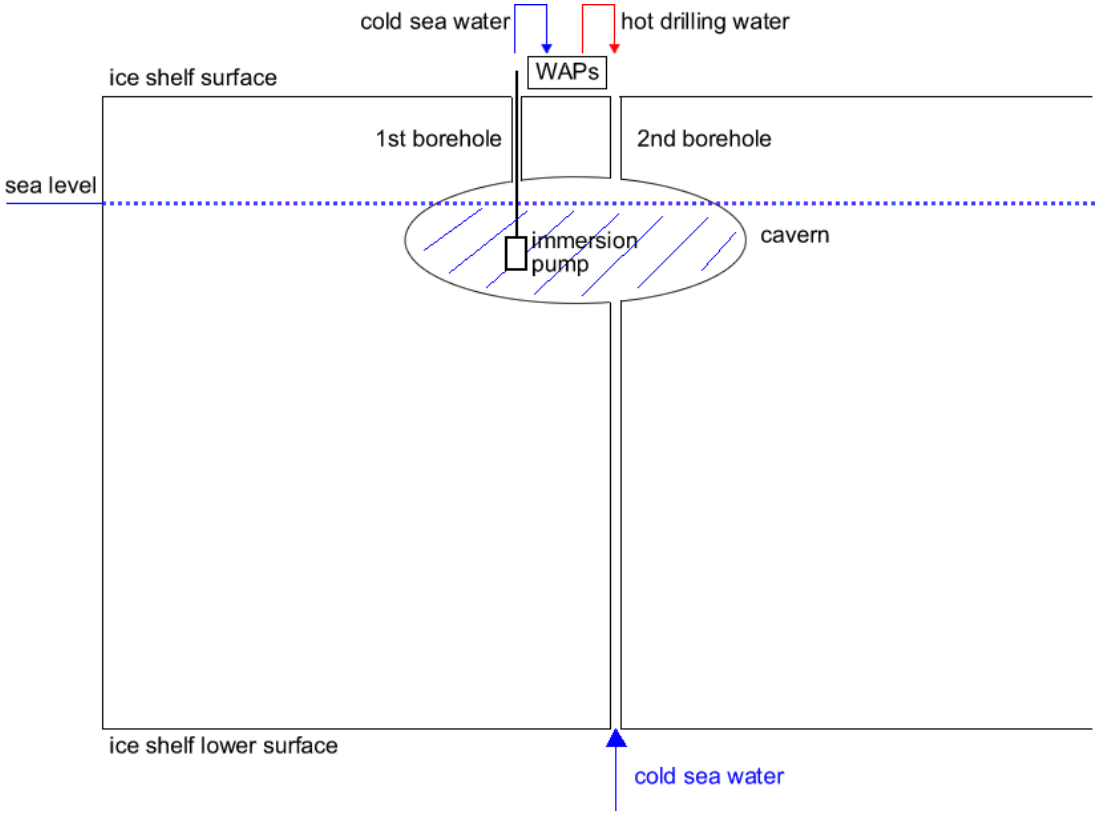


Figure 2.15: Schematic of the water recirculation system.

The set-up of the drilling equipment required ten days after it was unloaded from the R/V Polarstern. Three to four days were needed for preparing each drilling site, filling the cavity with water, penetrating the ice with splash lances, and deploying the hydrophones and CTD. The equipment was dismantled and

stored in a container within seven days. The four penetrations were achieved on 21, 27 and 30 December and 3 January.

The drilling sites were arranged in a triangle around the central borehole (70.52°S, 8.23°W) at distances of about 520 m as given in Figure 2.16. A 3D sketch of the array is given in Figure 2.17.

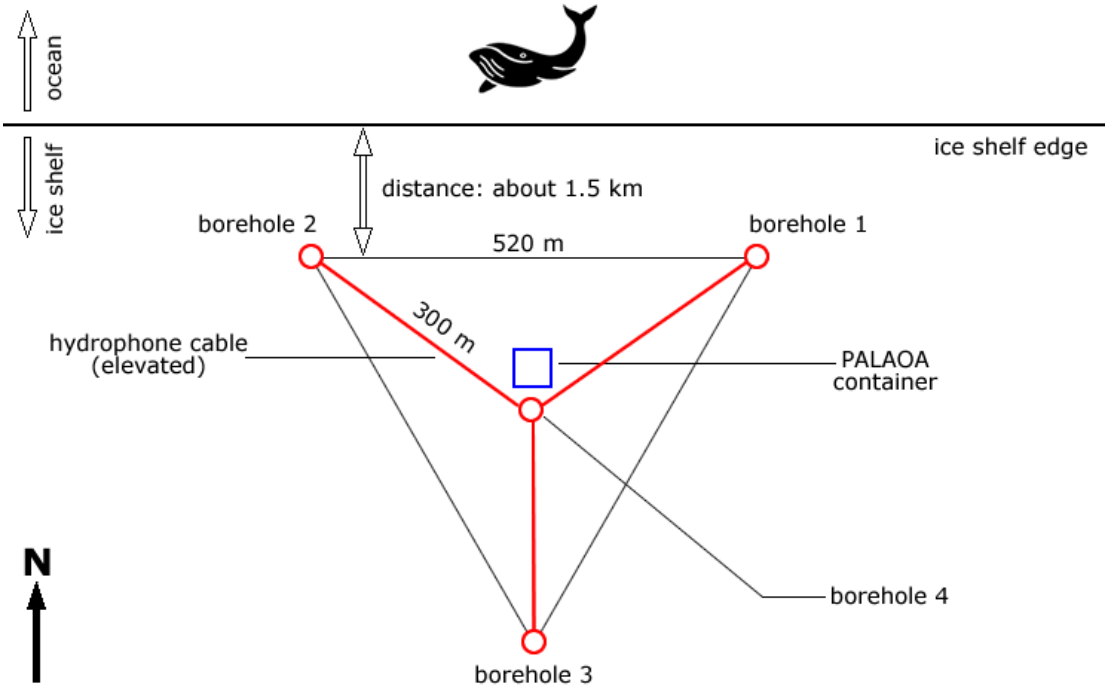


Figure 2.16: Top view to the PALAOA hydrophone array.

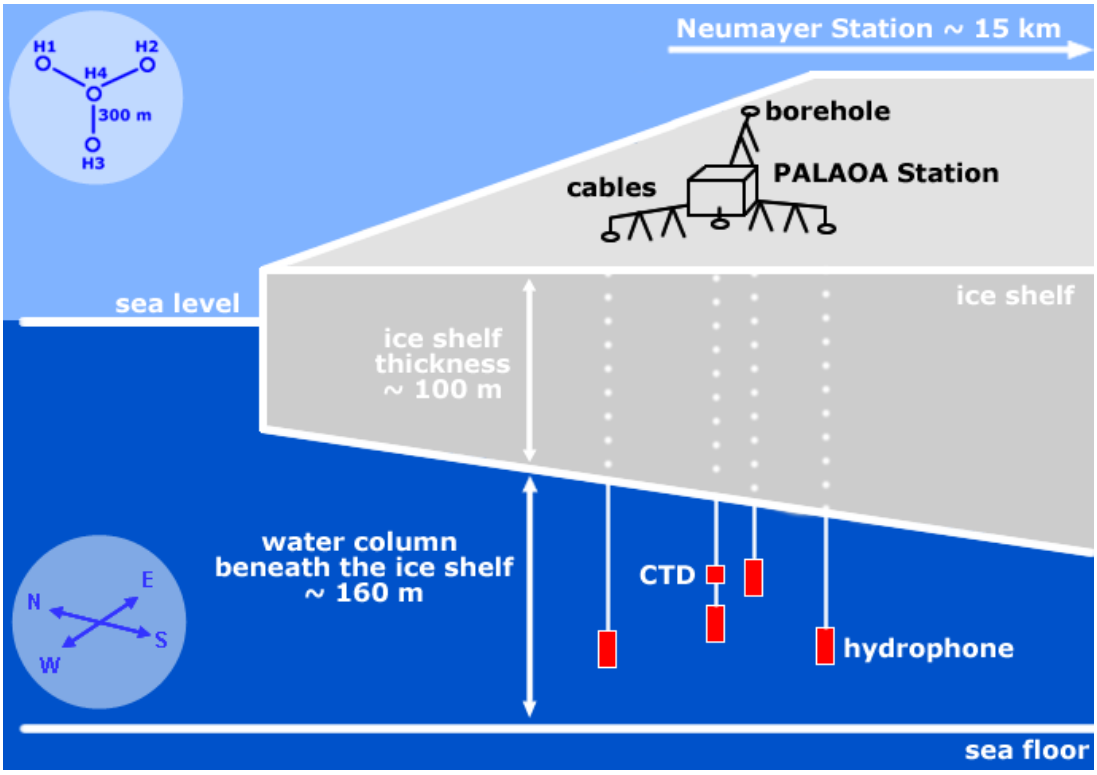


Figure 2.17: 3D sketch of the PALAOA hydrophone array.

Details about the hydrographic parameters and deployment depths at the four drilling sites are given in Table 2.4.

Table 2.4: Hydrographic parameters and deployment depths at the four drilling sites.

Site	*Central (4)*	*East (1)*	*West (2)*	*South (3)*
Position lat[1)] (°S)	70.5229	70.5216	70.5216	70.5256
Position lon[1)] (°W)	8.2296	8.2226	8.2366	8.2296
Ice thickness [m]	98	95	100	99
Water column thickness [m]	161	161	159	159
Sea floor depth[2)] [m]	243	240	245	244
Hydrophone cable length[3)] [m]	190	170	170	170
Hydrophone under ice[4)] [m]	92	75	70	71
Hydrophone depth[5)] [m]	175	155	155	155
CTD under ice[4)] [m]	72	-	-	-
CTD cable length [m]	170	-	-	-
CTD depth[5)] [m]	155	-	-	-

1) Positions measured with 4 differential GPS devices at January 14th 2006. (Remark: positions drift with time).
2) Sea floor depth = water column thickness + ice thickness – 15 m sea above water (estimated freeboard of ice shelf).
3) From surface of ice
4) Under ice = cable length – ice thickness
5) Depth = cable length – 15 m

During the drilling procedure, the lowering rate of the lance had to be controlled carefully. To guarantee the holes' exact vertical orientation, the splash lance had to hang freely inside the borehole, straightened by gravitational forces. Thus the eventual veering of the lance when touching ice at the bottom of the borehole could be prevented. To check whether the lance hung freely, the cumulative weight of hose and lance was monitored continuously, requiring that the weight would always increase (due to the increasing hose length). A decreasing weight would indicate that the lance was touching the (current) bottom of the melt hole. In this case, the lowering rate was instantly decreased until the weight increased again.

Figure 2.18 provides weight profiles derived from the westerly borehole as an example.

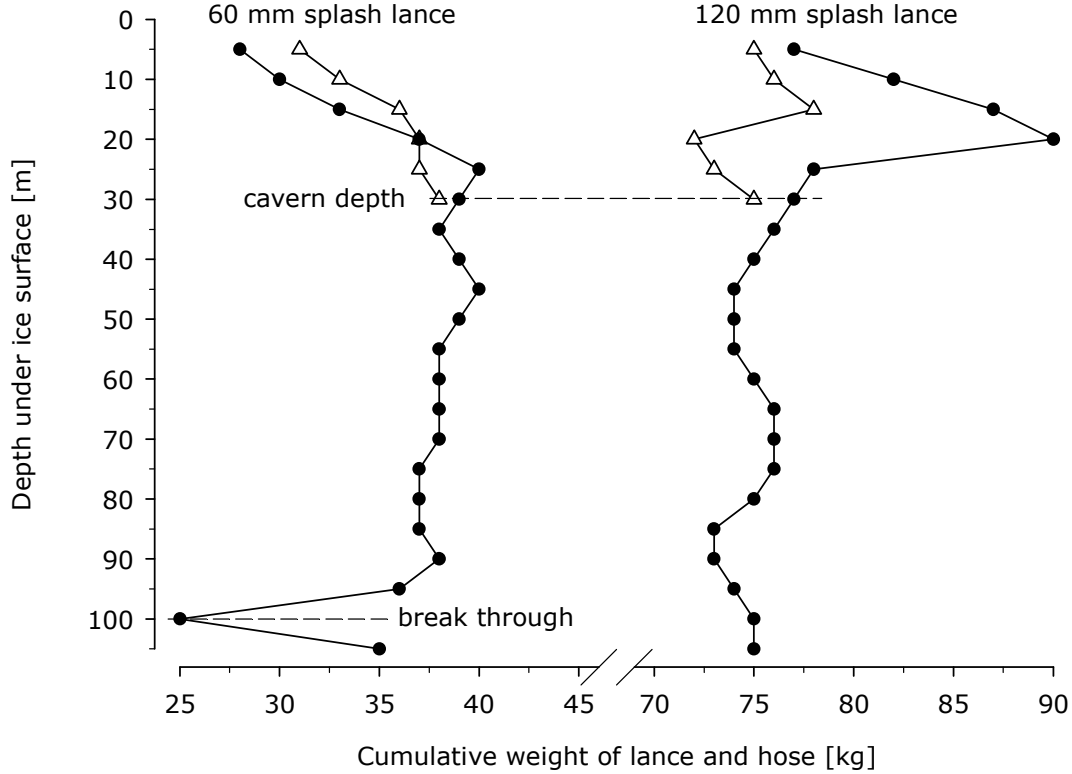


Figure 2.18: Weight profiles during lowering of splash lance of two diameters and hose for melting the cavern (hollow triangles) and up to the break through at 95 m depth of the westerly drilling site (solid circles). Boreholes were lowered through the solid shelf ice by injecting hot water at flow rates of 60 l min^{-1} for the cavern at 50 (60 mm lance) and 40 bar (120 mm lance). Flow rates to penetrate until the break through occurred were 70 l min^{-1} (60 mm lance) and 80 l min^{-1} (120 mm lance) at 60 and 50 bar respectively. Lowering rates were constantly either 35 m h^{-1} (cavern 120 mm lance) or 40 m h^{-1} (all other drilling until break through). Inflection points of the profiles represent alternating water levels inside the duct of the hole during the drilling progress.

The 60 mm splash lance (29 kg plus hose weight) was first lowered at 40 m h-1, 60 l min-1 flow rate at 50 bar until it reached the cavity depth 30 m below the ice-surface. The rate of weight increase was slightly – but not significantly - reduced between 15 and 30 m when water partially filled the duct of the hole leading to a weight reduction of the immerged lance. The cumulative weights of

both lances and hose remained almost stable during the melting process (36.19 ± 3.98 kg and 76.76 ± 4.39 kg; mean ± SD) (the hose apparently has little weight in water). The break-through into ocean water at the bottom of the glacier was clearly detectable at almost 100 m depth by a precipitous drop in weight caused by salt water intruding from the ocean into the borehole.

The total fuel consumption of about 9.000 l used during the five weeks drilling campaign is given in detail in Table 2.5. The technique and set-up of the equipment has been described in detail by Nixdorf *et al.* (1994a), Nixdorf *et al.* (1994b) and Nixdorf *et al.* (1997).

Four other 20 ft sledges carried a 90 kVA generator and fuel barrels, a container for equipment, a platform for supplies and materials, and a 15.000 l tank container. Two Kässbohrer Pistenbullies with hydraulic crane and bucket, respectively, assisted for transportation and machine support. A team of four operated the drilling. Drill operations lasted for about eight to ten hours per day, including time for transportation between Neumayer Base and the drill site.

Table 2.5: Fuel consumption during the five week hot water drilling campaign Dec. 05 – Jan. 06.

Date	***90 kVA Generator***	***Two Bullies***	***Six Waps***	
	Arctic Diesel [l]	***Petrol [l]***	***Arctic Diesel [l]***	***Petrol [l]***
05.-19.12.2005		900	1.200	200
20.12.2005	180	293	712	
22.12.2005	104	295	864	200
24.12.2005	110	265		
27.12.2005	99	155	579	200
30.12.2005	73	251	632	200
03.01.2006	89	155	654	
05.01.2006		269		
07.01.2006		258		
09.01.2006		120		
Sum:	655	2.961	4.641	800

2.4.2.2 PALAOA 06 – field work: Energy Module (EM)

The control cabinet of the energy module within the container was equipped in Bremerhaven in advance of the expedition (see Figure 2.19).

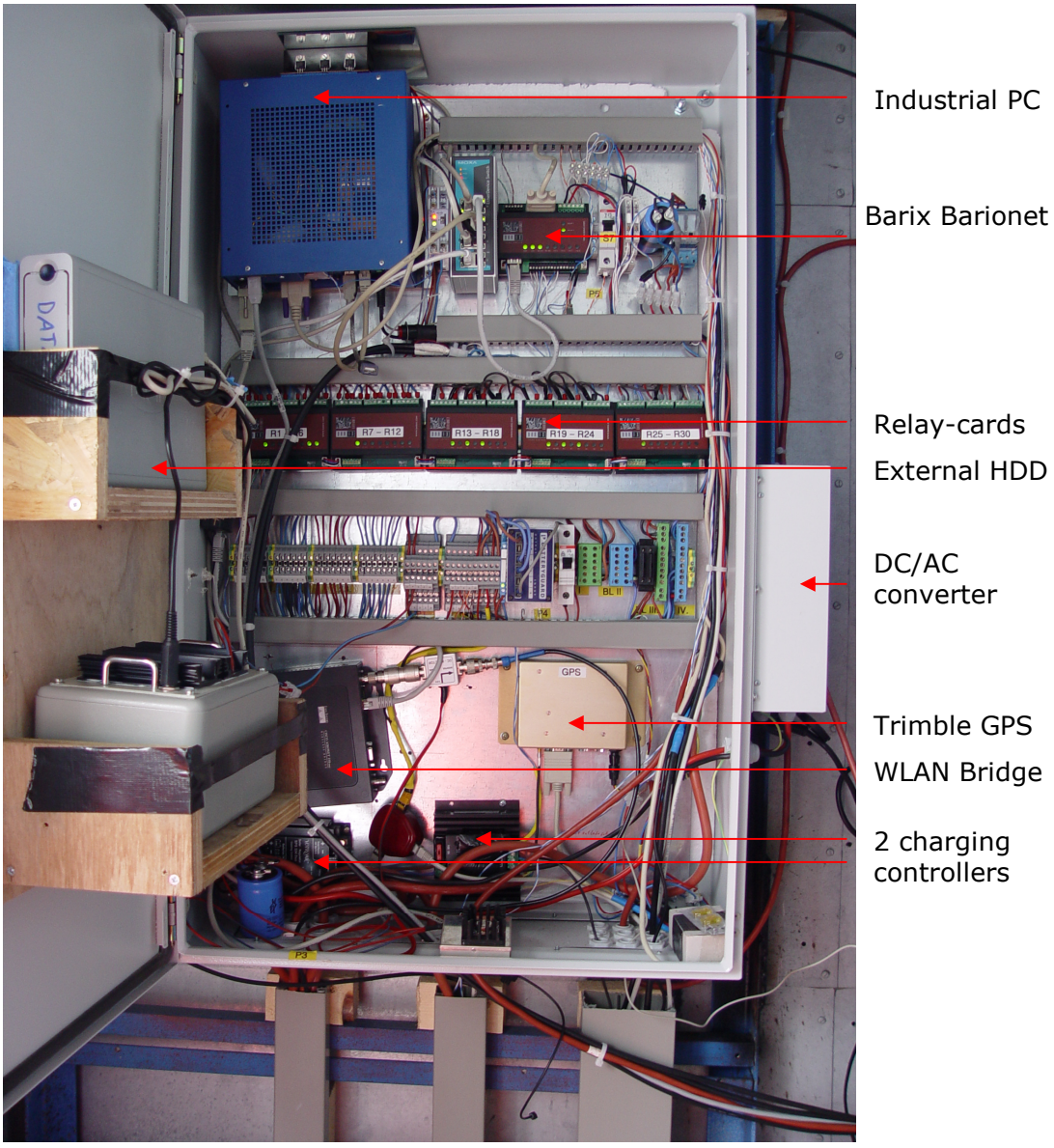


Figure 2.19: Control cabinet of the energy module.

The central container of PALAOA 06 was rigged concurrent with the drilling operations. The 6 m pylon of PALAOA 05 with WLAN antenna, wind generator and webcam was mounted on top of the container (see Figure 2.20). The fuel cell and the methanol canisters were placed inside the container. The tube for the CO_2 and water by-products was fed into a barrel.

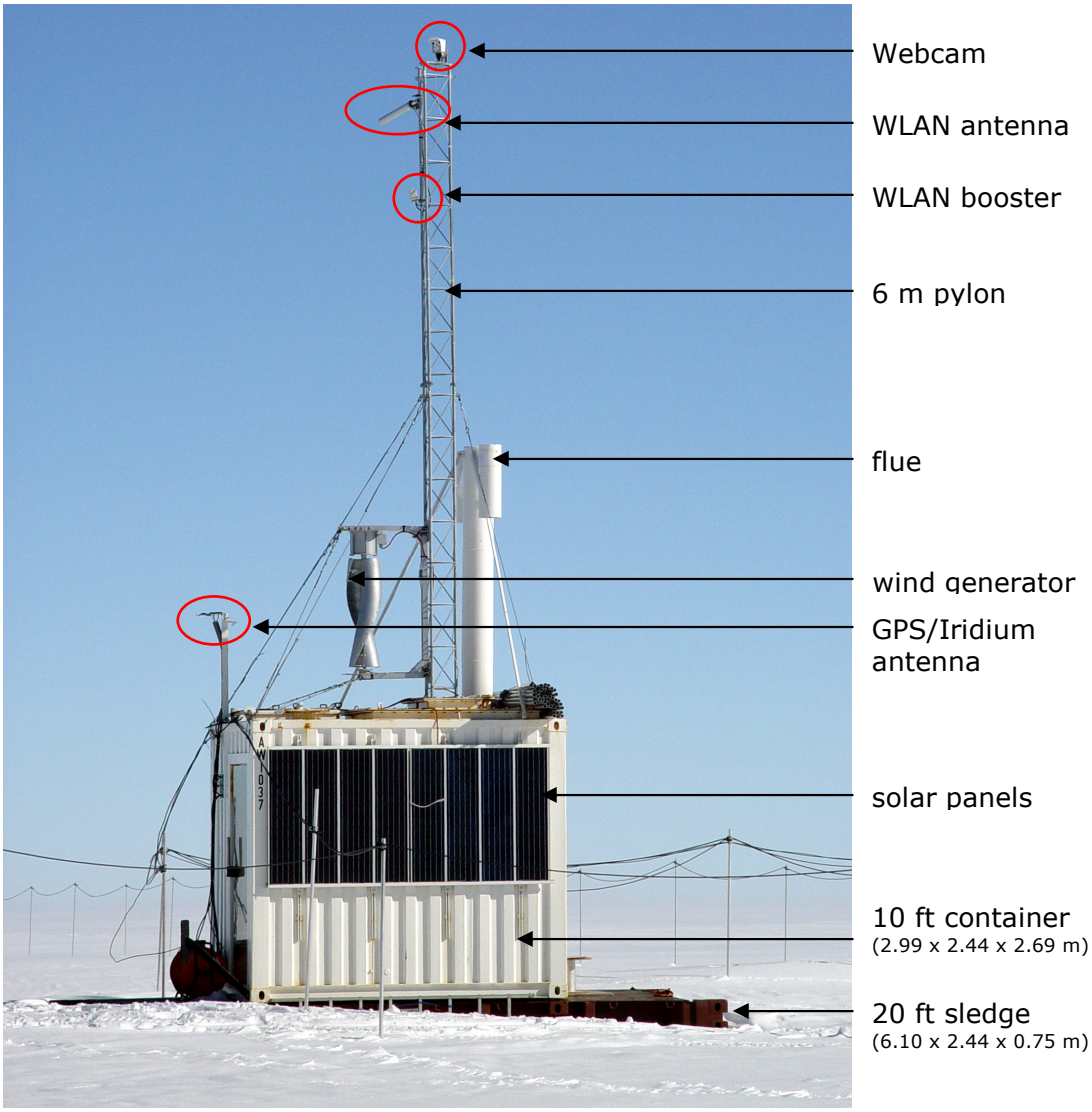


Figure 2.20: PALAOA 06 container.

2.4.2.3 PALAOA 06 – field work: Communication Module (COM)

A WLAN booster was installed at PALAOA 06. To reduce the attenuation of the WLAN link, the booster was installed close to the WLAN antenna (see Figure 2.20).

2.4.2.4 PALAOA 06 – field work: Additional Sensors (AS)

The CTD sensor was deployed together with the central hydrophone through the central borehole of PALAOA (see Figure 2.17). Furthermore a GPS was successfully integrated in PALAOA 06 (antenna installed on top of the container – see Figure 2.20).

2.4.3 PALAOA 06 – results and discussion

2.4.3.1 PALAOA 06 – results and discussion: Acoustic Module (AM)

All hydrophones were successfully lowered through the boreholes and tested successfully one by one after deployment. The hydrophone array allows the determination of the position of sound sources. An example is given in Figure 2.21.

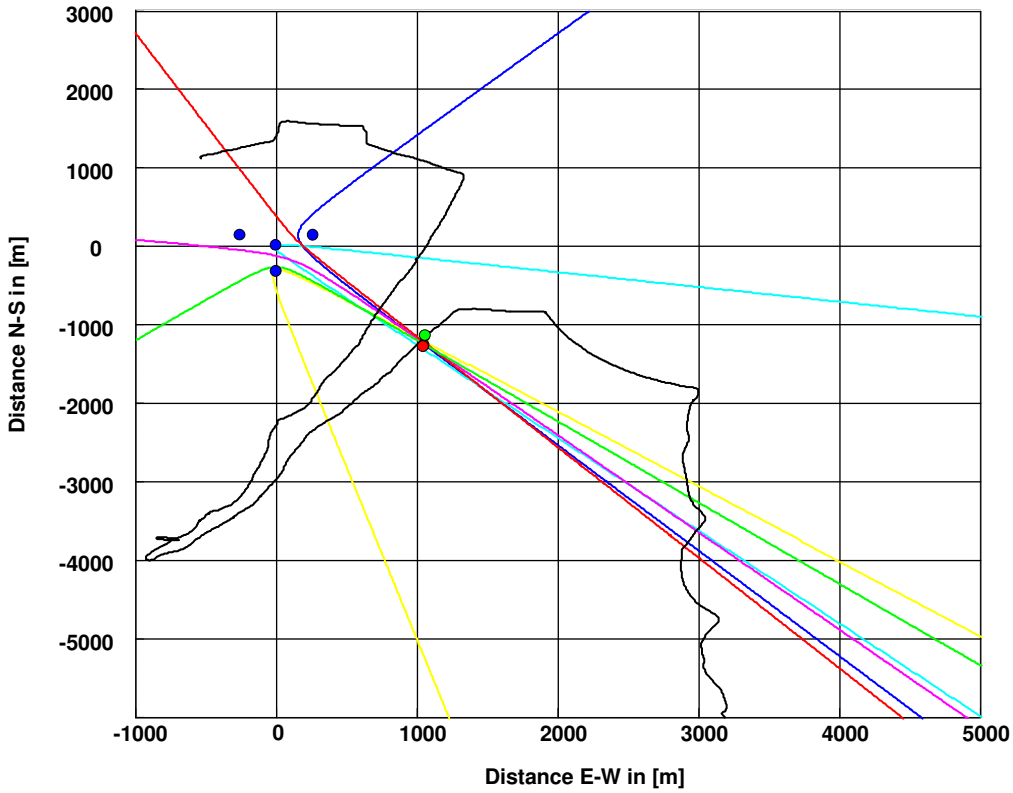


Figure 2.21: Example for a calculated position of a sound source (red dot) and calculated position (centre of head of intersect; green dot) using the PALAOA array (blue dots). The colored lines are representing the hyperbolas for the six hydrophone pairs – see description in test.

The blue stars indicate the hydrophones of the PALAOA 06 array. The acoustic source's position is determined as follows: A sound wave arrives at different times at the four hydrophones. The difference between the time of arrival between two hydrophones is called time of arrival difference (TOAD). For each pair of hydrophones (H1-H2, H1-H3, H1-H4, H2-H3, H2-H4 and H3-H4) a TOAD can be measured. The TOAD is then used to calculate a hyperbola of possible positions of the source for each hydrophone pair (colored lines in Figure 2.21). Any point on the hyperbola represents the same TOAD.

The six hyperbolas (one for each hydrophone pair) should intersect at one point which represents the possible location of the sound source. The advantage of an array with 4+ hydrophones is that the calculation results in a so-called "head of intersect" for a calculated 2D position. The head of intersect is a region where

the source is located. For the example shown in Figure 2.21 the centre of the head of intersects and the true position of the source agree quite nicely. For an array of only three hydrophones, the hyperbolas will always intersect in one point, with no empirical information on the possible error of the calculated position. Two hydrophones will provide information only on the bearing of a sound source.

When simultaneously connecting multiple hydrophones to the audio system, frequent spikes impaired the recordings severely. The problem is most likely caused by the creation of a large antenna (see Figure 2.22) when connecting multiple hydrophones in the audio control cabinet to a common ground.

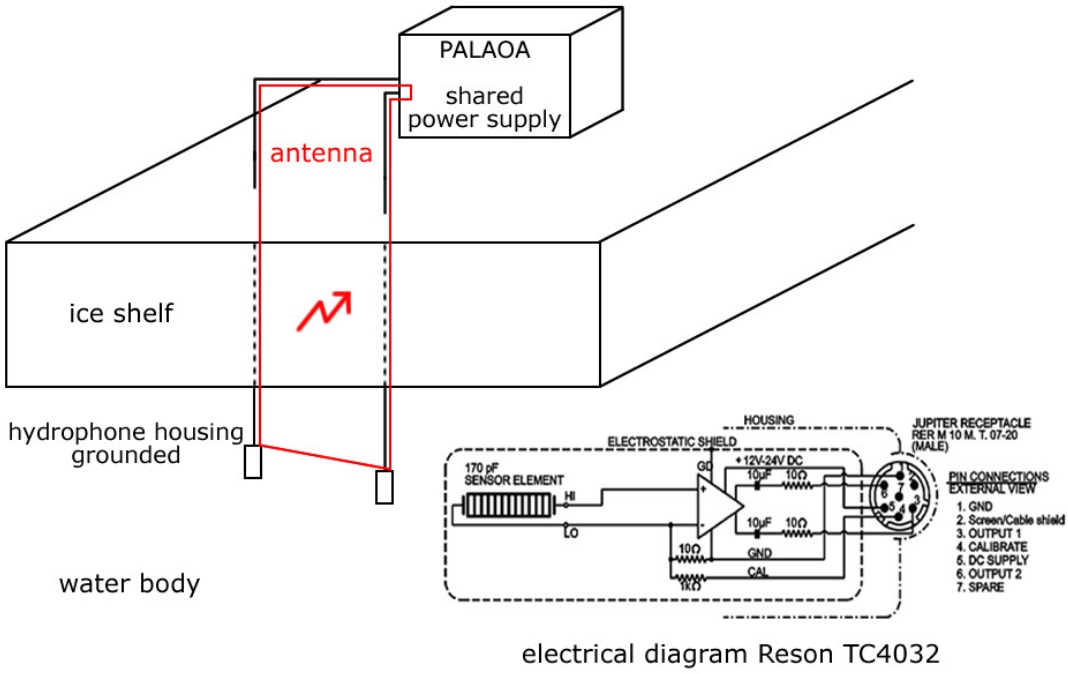


Figure 2.22: The ring antenna effect.

The cabinet's ground is connected via the ground conductor to the external MESSING housing of hydrophone A, through the water from the housing of hydrophone A to B, and again the housing of hydrophone B via its ground conductor to the cabinet ground forms a large ring antenna. This antenna most likely picks up so-called geophysical sferics, electromagnetic impulses triggered by thunderstorms at great distances (Watkins *et al.* 1998; Watkins *et al.* 2001). While this problem could not be fixed during the filed season 2005/2006, it did not compromise the goal of obtaining records of the Antarctic underwater soundscape when we connected only one hydrophone to the system. For this, hydrophone WEST (see Table 2.4 on page 69) was selected.

2.4.3.2 PALAOA 06 – results and discussion: Energy Module (EM)

The electronic noise that compromised the PALAOA 05 ice shelf records was reduced to a minimum by the new galvanic separation between batteries powering the audio and the remaining modules. While this approach is technically rather complex, it is the most reliable with regard to the reduction of electronic noise.

Surprisingly almost every device endured the harsh conditions; only the fuel cell failed. At low temperatures, water, which is one of the exhaust products, caused significant problems within the cell. The exhaust tube froze from time to time, preventing the outflow of water from the cell, which then caused the entire fuel cell to be frozen. Under such circumstances, the fuel cell had to be returned to Neumayer Base to defrost. Careful examination of the problem revealed that the control software of the cell caused this problem. When the fuel cell was on standby the software controlled the temperature of the stack. If the stack temperature reached a certain threshold the cell powered up and started to produce energy and heat to avoid freezing. After heating up, the cell goes back into standby mode and remains in this state for a certain time before a wake up can be initiated again. This interval was probably too long to avoid freezing at low temperatures (< -30°C) when the cell cooled down fast. Feedback to the manufacturer helped to improve the control software of the next generation of the fuel cell, which was to be installed in the field season 2006/2007. Furthermore, this generation of fuel cells became equipped with a tube heating element.

In 2006, PALAOA 06 was online on 204 days (56%) – though not necessarily for all 24 hours of each of these days. On an hourly basis, an uptime of 42% was reached for PALAOA 06 (see Figure 2.23). Thus, the design goal to keep PALAOA 50% for the year 2006 online was almost achieved.

Nevertheless, the fuel cells malfunctioning during austral winter 2006 caused an energy shortage which was a major cause for drop-outs during this season. Energy uptime might have been increased by up to 9% however, due to the broken WLAN link, information on the status of the energy system during this period is unavailable.

For the station's control system, a web interface was programmed (Figure 2.24) for the Barix Barionet unit. This web interface allows remote control of the PALAOA 06 station from the AWI. The web interface displays the status of each relay as well as the on- and offline status of each device installed at PALAOA 06. In case of pending energy shortages, indicated by decreasing battery voltages as displayed in the web interface, consumers of low priority (webcam, temperature sensors and GPS) were switched off. The web interface also allows modification of the control software by uploading new versions, permitting year-round development and testing of the control software.

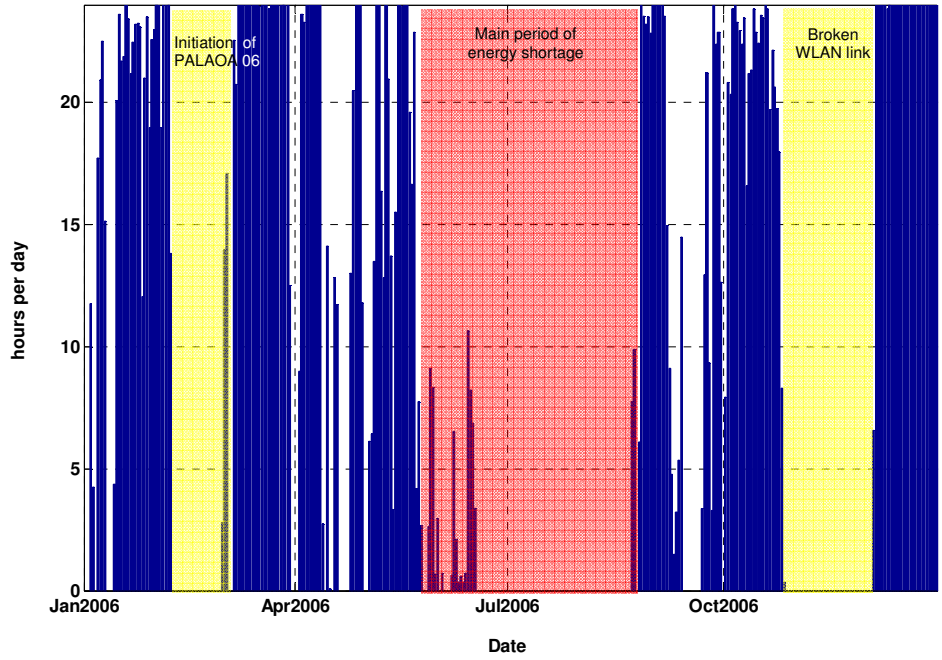


Figure 2.23: Uptime PALAOA 06.

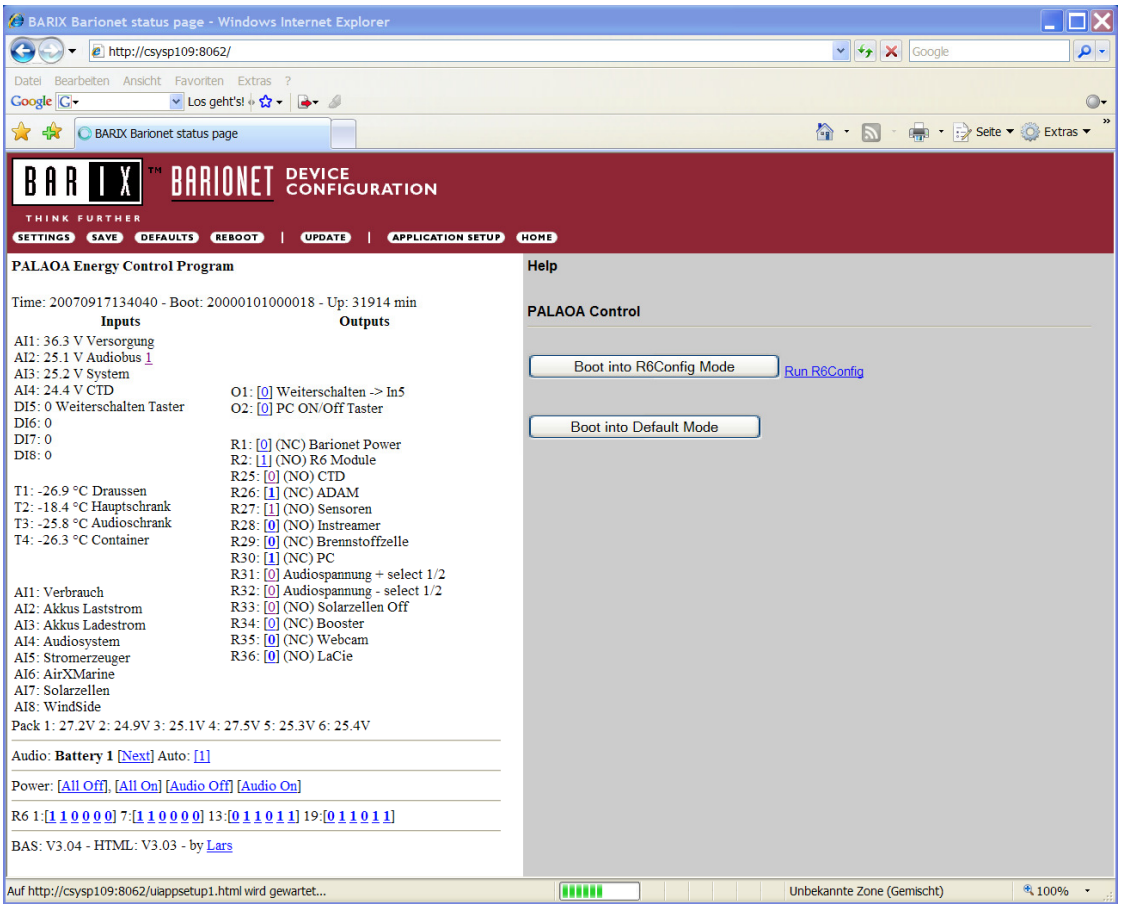


Figure 2.24: Web interface of the control system.

All data were transmitted via the WLAN link to Neumayer Base, where a Matlab™ routine saved the data locally. The operational data were transferred to AWI via FTP.

2.4.3.3 PALAOA 06 – results and discussion:

Communication Module (COM)

PALAOA 06 was equipped with a local WLAN booster. This modification was successful as the bandwidth was increased from 1.3 Mbit/s (PALAOA 05) to 3 Mbit/s.

A heavy storm at the end of October 2006 was probably the cause for a breakdown of the WLAN link. During this storm the rubber of a self-locking screw nut of the webcam clamps appeared to have dissolved. Caused by the vibrations of the pylon during heavy wind speeds, the screw nut would then have dismantled. The webcam was dangling at the Ethernet cable from the pylon, and finally destroyed the WLAN antenna and itself. This breakage caused an extended period without connectivity between Neumayer Base and PALAOA 06 because the weather remained inappropriate for maintenance for some time (see Figure 2.23).

The initiation of the livestream - particularly the installation and testing of the appropriate software - was most successful. The existing permanent satellite link between Neumayer Base and the AWI was integrated to the PALAOA 06 system to send the data as described in paragraph 2.4.2.3 in real-time to the AWI.

2.4.3.4 PALAOA 06 – results and discussion: Additional Sensors (AS)

The newly installed GPS sensor provided a time series of the station's location. Between February 2006 and September 2006 the station moved laterally by approximately 100 m towards the Weddell Sea, suggesting a mean drift of 150 m a^{-1} in N-NE direction. The displacement and daily changes of PALAOA's position have to be considered when calculating the bearing/position of underwater sound sources.

During a heavy storm at the end of October 2006, the data cable of the CTD sensor was fractured. As a consequence the CTD data could not be read and a data loss for the second half of the year 2006 occurred. However, the CTD data recorded during the first half of the year 2006 provided detailed information on the tidal currents in the vicinity of PALAOA.

2.5 PALAOA 07

2.5.1 PALAOA 07 – design goals and layout

For the planned modifications of PALAOA 07 no major structural alteration of the container was necessary. Thus the modules of PALAOA 07 remained in the container as described for PALAOA 06.

2.5.1.1 PALAOA 07 – design goals and layout: Acoustic Module (AM)

The design goal for the acoustic module of PALAOA 07 was to gain multi-channel recordings form the Antarctic underwater soundscape. This necessitated a significant reduction of the audible sferics in the recordings when connecting more than one hydrophone to the audio system. To reduce the noise in the acoustic data a complete galvanic isolation between all hydrophones had to be implemented to disable the ring antenna effect. Instead of using one central multichannel A/D converter (the MOTU Traveler soundcard) each hydrophone signal should be A/D converted by its own converter to SPDIF digital audio format. These signals would then be fed into a multichannel digital soundcard (type RME Hammerfall) in the PC using optical cables. Each of the four hydrophones, amplifier and A/D converter assemblies would have to be powered by their own switched battery pair, requiring a rather complex relay setup to allow for alternating charging and operation.

However, after detailed investigation of all installations on-site it became obvious that two of the three active hydrophones were defective and a repairi was not possible. Hence the array was reduced to only two hydrophones – the passive hydrophone located close to the station container and an active hydrophone at a distance of 300 m in N-NW direction (300°). With only two hydrophones operational, the galvanic isolation of the respective data cables became easy to be achieved with a line isolating box. This device works with isolating transducers and is often used to reduce ripple pickups. Furthermore an additional Reson VP2000 Amplifier was connected to the passive hydrophone. This additional amplifier is located in a Zargesbox outside the container. The idea was to amplify the signal of the passive hydrophone before passing all the electronic cords inside the container. This should reduce the electronic noise in the recordings of the passive hydrophone to a minimum.

By simple rewiring it was also possible to modify the battery charge/operation circuit to provide galvanically separated battery supplies for both hydrophones. Now, four of the six battery packs are charged while two packs power one hydrophone each.

The malfunctioning of the other two hydrophones was probably caused by the infiltration of sea water into the hydrophone pressure case (at a depth of 154 m they are subject to an ambient pressure of about 15 bar).

To compensate for the loss of the two hydrophones it was decided to construct a small listening station and deploy it – at least temporarily - on the sea ice. The station was called PALAOA-S(atellite). The goal was to collect data which would enable - in combination with the PALAOA data - 2D localization of sounds in the vicinity of the two stations. In particular, this requires precise synchronization of the PALAOA and PALAOA-S data streams.

For the assembly of PALAOA-S the following components were used:

Hydrophone:	Reson TC4032
Filter/Amplifier:	Reson VP2000
Recorder:	M-Audio Microtrack 96/24
GPS:	Trimble Lassen IQ
Power supply:	Panasonic rechargeable batteries 12 V and 6 V

The equipment was assembled in a Zargesbox (size 60 x 40 x 25 cm). The GPS was mandatory to provide synchronization between the PALAOA and PALAOA-S stations.

Therefore the GPS provides two kinds of information:

1.) 1pps signal: One pulse per second; an accurate signal of 4 µsec duration.
2.) NMEA signal: A RS232 string which contains the actual date, time and position.

To be able to record the short 1pps signal on one audio channel with a sample-rate of 48 kHz, the PALAOA and PALAOA-S GPS had to be modified. A monoflop circuit was implemented in the GPS to lengthen the 1 pps signal to ~200 msec duration and to synchronize the two listening stations. The procedure of calculating the position of a sound source is described in chapter 2.4.3.1 in detail.

2.5.1.2 PALAOA 07 – design goals and layout: Energy Module (EM)

The design goal for the energy module was to provide PALAOA 07 with 75% of the needed power. To reduce the impact of sferics, changes in the energy module were necessary. To calculate the energy balance the control system should be augmented with electric current sensors. The main problem during the operation of PALAOA 06 remained the provision of power during winter time. To overcome the energy shortage that occurred during PALAOA 06, an additional wind generator - type AirXmarine - was to be installed. This generator starts to produce energy at moderate winds (> 8 knots) while the Savonius wind generator is designed for high wind speeds (> 15 knots).

Also a second trial with a new methanol fuel cell was carried out because this device produces energy independently from the weather conditions.

To overcome noise caused by sferics the power supply of the audio system had to be separated into four independent systems, each providing one hydrophone with power (the abovementioned fact that two had failed was not known in the planning stage of PALAOA 07).

Furthermore a new battery type AGM (Absorbed Glass Mat) Lifeline was installed. These batteries have - compared to the standard batteries - a better capacity-temperature-curve. In addition, the AGM batteries can be more efficiently loaded at low temperatures. The energy loss during loading is only

2% (normal gel batteries = 16% energy loss) and the self-discharge of the AGM batteries is < 1%. Furthermore hall sensors were implemented to the control system. These sensors (measuring electric current) allow, in combination with the voltage measurements, calculation of the energy balance of PALAOA 07 instantaneously.

In PALAOA 06 all operation parameters were logged using a Matlab™ routine. In PALAOA 07 a syslog system was introduced. The syslog system was based on the KIWI syslog software which allows sharing the syslog with several computers in the network (in this case PALAOA, Neumayer Base and AWI). It also allows manual entries from every computer within the network. Thus, a small comment tool was programmed and installed on all computers of the ocean acoustics group. If anyone recognizes an interesting sound, it is possible to make an entry which is stored in the central syslog. The syslog is stored daily in a log file.

2.5.1.3 PALAOA 07 – design goals and layout:

Communication Module (COM)

The design goal for the communication module was to establish a secondary radio link between the observatory and Neumayer Base to be able (1) to communicate with PALAOA in case an antenna breaks and (2) to reduce energy consumption if necessary. Also a high-quality audio stream between PALAOA 07 and Neumayer Base was to be established. The communication module was retrofit with a second radio link. For this, additional Cisco AIR-ANT3338 dish antennas (+ 21 dBi) were selected. The WLAN bridge provided the possibility to connect two antennas. Consequently the Cisco dish antenna was connected via the booster with the WLAN bridge while the YAGI antenna was set up as a backup without a booster. This also reduces the energy consumption during austral winter. If the available amount of energy reaches a critical point the booster can be switched of using the Barix Barionet. After shutting down the booster the WLAN bridge recognizes that the dish antenna is not available and switches automatically to the YAGI antenna. In order to be able to stream multi-channel, high-quality audio signals, the AsioRecorder should be extended with an Icecast client module.

2.5.1.4 PALAOA 07 – design goals and layout: Additional Sensors (AS)

The design goals for the additional sensors were (1) to repair the broken CTD cable to allow long-term measurements of the tidal current and (2) to feed the meteorological data measured at Neumayer Base into the control system. The meteorological data are measured and stored at Neumayer Base every minute. The syslog routine was implemented gathering the information from the meteorological observatory and storing it into the daily syslog file.

2.5.2 PALAOA 07 – field work (January 2007 – February 2007)

Field work was conducted from early January to late February 2007 within the ANT-LAND 06/07 expedition. In advance to the modifications of the PALAOA modules, the sastrugies (snow banks) around the PALAOA 07 container were removed.

2.5.2.1 PALAOA 07 – field work: Acoustic Module (AM)

In Figure 2.25 the contour of the ice shelf edge[2.12] and the positions of PALAOA (70.5211°S - 8.2273°W) and PALAOA-S (70.6168°S - 8.1377°W) stations are given.

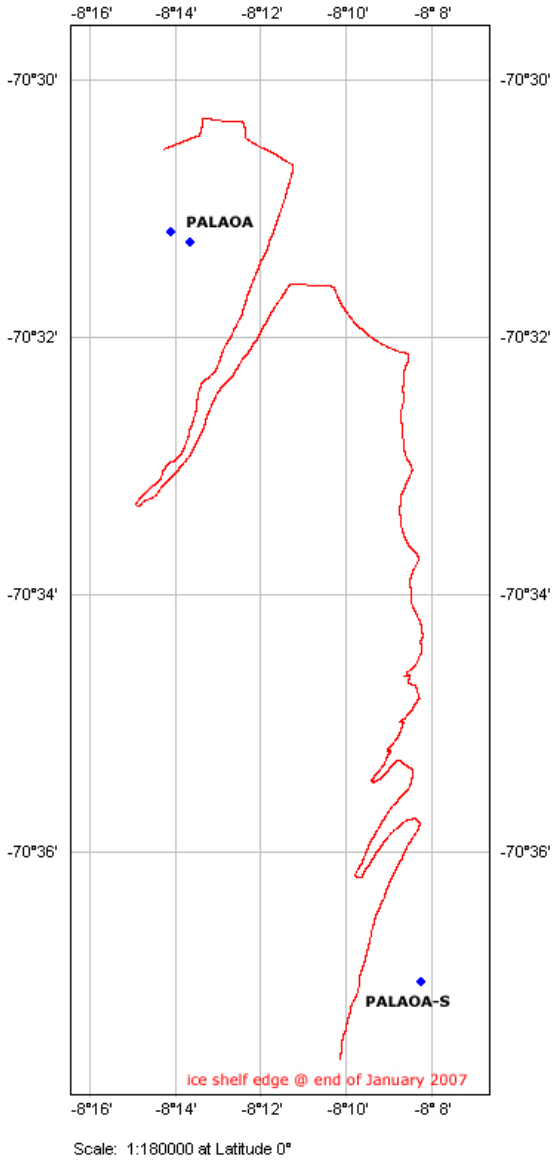


Figure 2.25: Contour of the ice shelf edge and positions of the acoustic stations in Jan. 2007.

[2.12] Thanks to Maja Petzel and Andreas Buhl, both members of the overwintering team 2006, for measuring the ice shelf contour with a Garmin GPS mounted at a Ski-Doo at 30 January and 31 January 2007.

The blue dots highlight the locations of both acoustic stations. The distance between PALAOA and PALAOA-S station was approximately 11 km. At total of 161 hours of data were recorded between 17 January and 26 January 2007. At end of January the sea ice in the Atka Bay started to break open and PALAOA-S had to be recovered. Therefore the elevated cables of the two defective hydrophones and their stands were recovered.

2.5.2.2 PALAOA 07 – field work: Energy Module (EM)

The additional wind generator (type AirXmarine) was installed on top of the flue (see Figure 2.26).

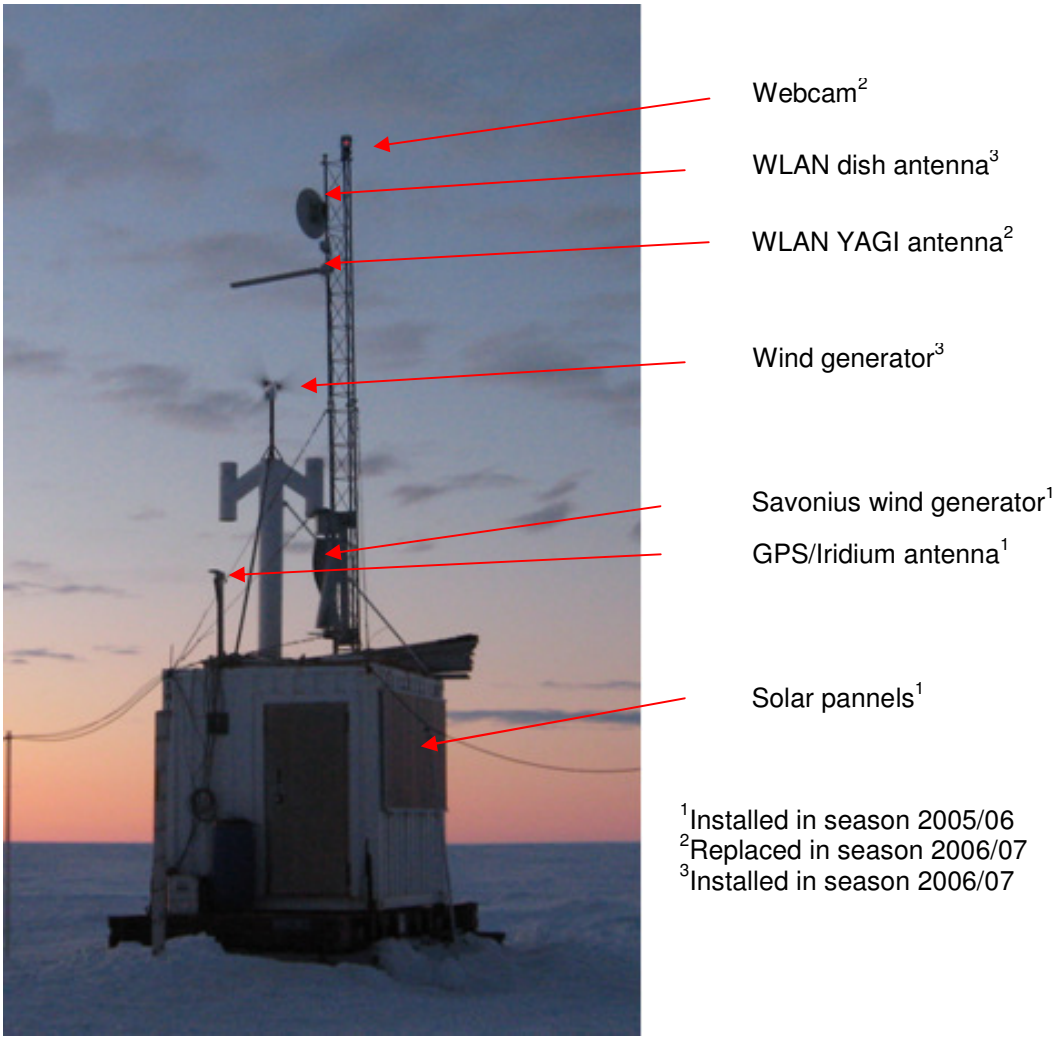


Figure 2.26: PALAOA Station (February 2007).

The new methanol fuel cell (operating voltage: 24 V) was implemented into the energy module. To increase the lifespan of the electronic components and to minimize the number of drop-outs, a voltage stabilizer was installed. This device provided the system with a constant voltage of 25 V even the input voltage varied between 12 V and 36 V. Furthermore all batteries were replaced with AGM (Absorbed Glass Mat) Lifeline batteries. A schematic wiring diagram of PALAOA 07 is given in Figure 2.27.

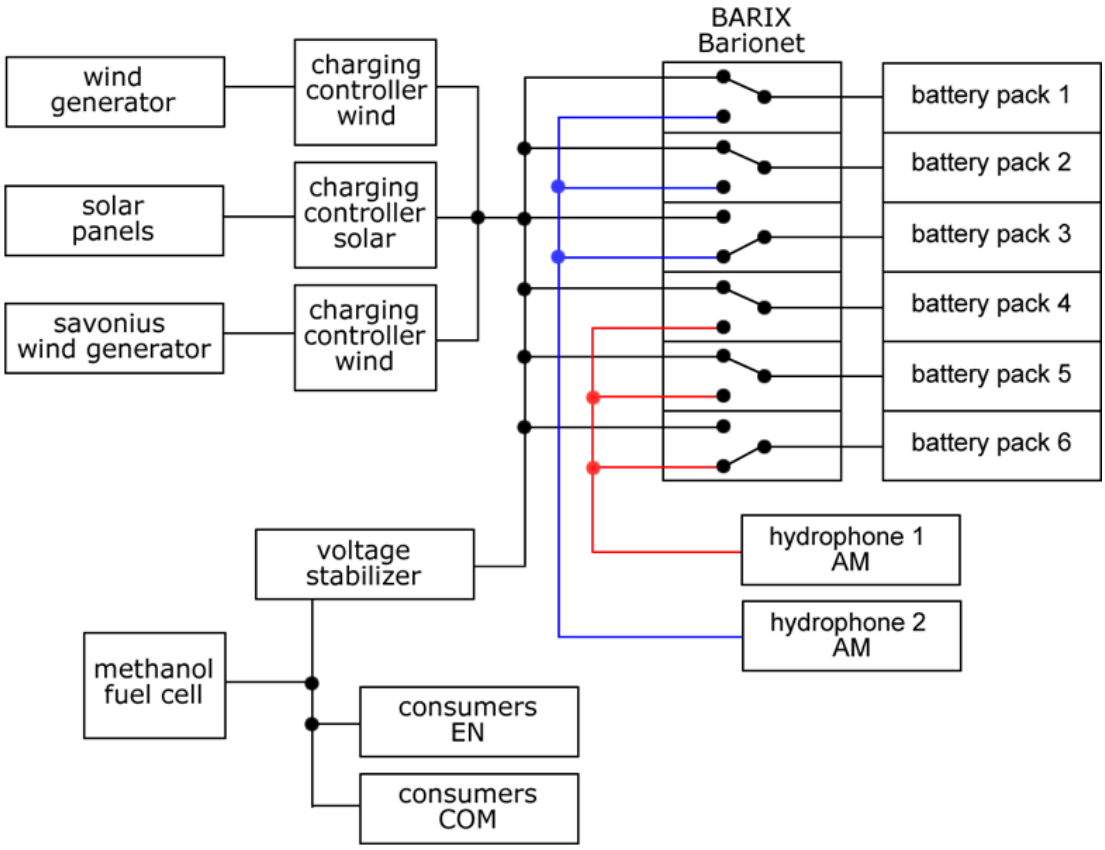


Figure 2.27: Simplified schematic wiring diagram of PALAOA 07.

The control system of PALAOA 06 monitored only the voltages of the battery packs and the system voltage of the energy and audio modules. During this season the monitoring system was extended with hall sensors. The outputs of all sensors are fed into a Barix X8 and allow monitoring the energy balance of PALAOA 07 at any time.

2.5.2.3 PALAOA 07 – field work: Communication Module (COM)

Two additional WLAN dish antennas were installed at Neumayer Base and PALAOA 07. The second radio link was successfully established.

2.5.2.4 PALAOA 07 – field work: Additional Sensors (AS)

The breakage of the CTD power/data cable was fixed. The sensor was reconnected to the system and has been in operation since January 2007. The Mobotix M10 day/night webcam was replaced by an updated version. An ADAM serial input device was installed. This device is equipped with two RS 232 inputs and one Ethernet output and allows access to the CTD and the fuel cell directly without the need to run the industrial PC.

2.5.3 PALAOA 07 – results and discussion

2.5.3.1 PALAOA 07 – results and discussion: Acoustic Module (AM)

The deployment of PALAOA-S was successful. Preliminary data analysis of the PALAOA and PALAOA-S data showed that both datasets include seal vocalizations. Because of the synchronization of the two stations the positions of the sound sources can be localized (see Figure 2.28). This is the precondition for calculating the source levels of the seal calls and also for estimating the number of seals in the vicinity of the stations. Figure 2.28 illustrates the result for a calculated position of a given sound source. The advantage of the combination of the two listening stations was the coverage of a wider area and the higher accuracy of the determined position as the accuracy decreases dramatically for sound sources outside the array.

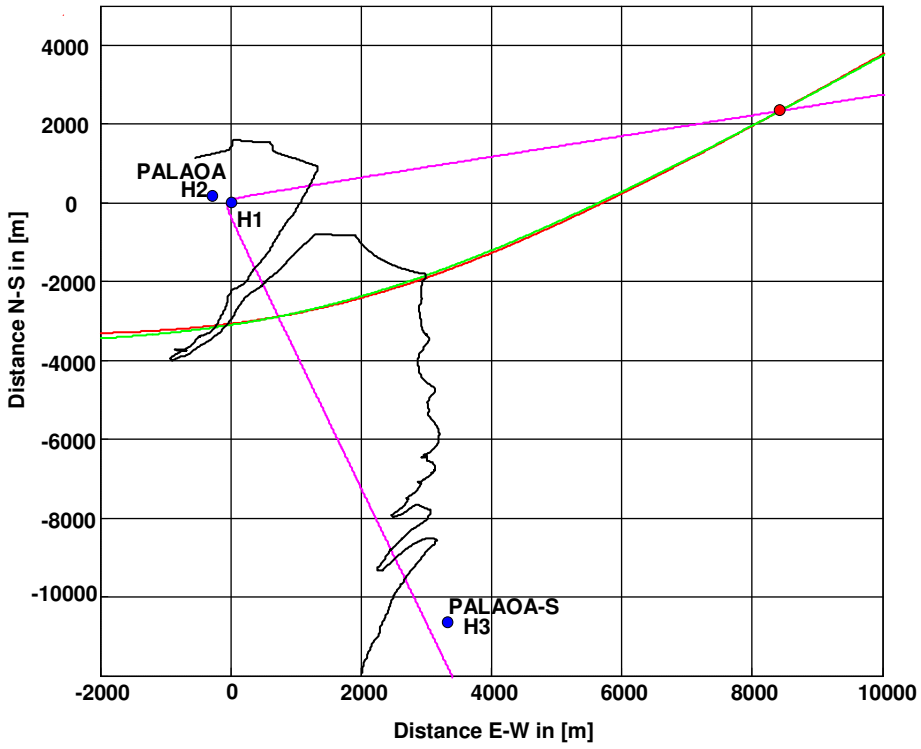


Figure 2.28: Example for a calculated position of a sound source (red dot) combining PALAOA and PALAOA-S (blue dots). Pink line: hyperbola H1-H2; red line: hyperbola H1-H3; green line hyperbola H2-H3.

2.5.3.2 PALAOA 07 – results and discussion: Energy Module (EM)

The extension of the monitoring system allowed calculating the energy balance and predicting possible periods of energy shortage. In this case unnecessary consumers were shut down to minimize the number of drop-outs due to energy shortage. The additional wind generator started to produce energy at wind speeds of around 8 knots and provided substantial amounts of power during periods with weak winds (by comparison: the Savonius wind generator starts to produce energy at around 15 knots). Unfortunately the additional wind

generator did not work at temperatures below -20°C probably because the lubricant inside the ball race gets too stiff. However, it is not clear yet if the wind generator will work again when temperatures increase.

The fuel cell worked smoothly during the first half of 2007. After a cold spell in July 2007, with temperatures as low as -50°C the cell malfunctioned and had to be taken to Neumayer Base for repair. Unfortunately, even there it was not possible to fix the problem without spare parts from Germany.

Overall, the station still powered down from time to time during the 2007 winter season. However, the system was online more often (271 days, 74% of all days) than in the previous winter season (see Figure 2.29). Most electronic components - whether part of the power, audio, communication or additional sensor module - worked reliably and only the WLAN bridge had to be exchanged. The main remaining problem is to provide the station during winter time with sufficient power. This problem can probably only be solved by a reliable fuel cell.

Also, the control software should be extended to automatically shut down devices with high power consumption during periods with an energy shortage. The available weather forecast could feed into a decision support system, and shut down decisions can be made as early as possible.

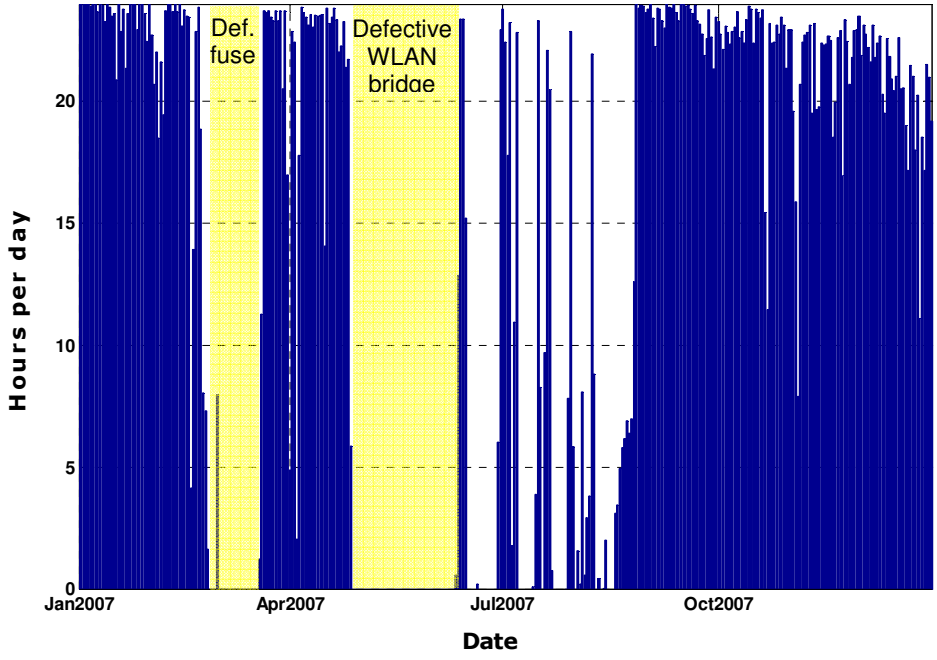


Figure 2.29: Uptime PALAOA 07.

2.5.3.3 PALAOA 07 – results and discussion: Communication Module (COM)

The AsioRecorder was equipped with a Streaming function (Icecast client module). This allowed streaming high-quality data directly to Neumayer Base without the need to store the data first on a local hard-disk at PALAOA Station.

2.5.3.4 PALAOA 07 – results and discussion: Additional Sensors (AS)

The breakage of the CTD cable was successfully fixed. This sensor has been recording oceanographic data continuously since January 2007 without any gaps. Figure 2.30 displays the CTD pressure data versus sound intensity as recorded in early April 2007.

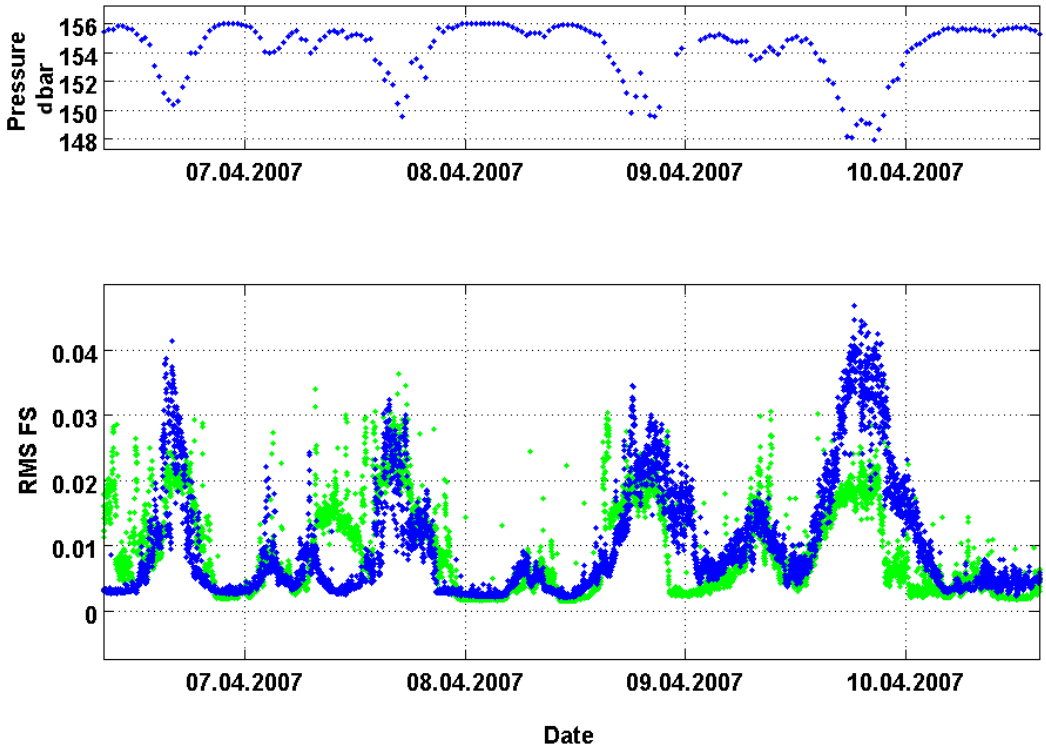


Figure 2.30: CTD pressure data [dbar] (upper panel) vs. sound intensity [voltage RMS FS] (lower panel; y-axis linear scaled) as recorded in early April 2007. Blue line represents hydrophone WEST, green line represents hydrophone CENTRAL.

The two curves shown in Figure 2.30 demonstrate the influence of the current on the full band sound intensity measured at the two hydrophones. The decreasing pressure indicates a deflection of the hydrophone cable at point where the CTD is mounted. Most likely, the increased noise level is caused by the movement of the cables as the hydrophones will be deflected as well. The deflection is caused by currents and can therefore indirectly provide additional information on the current system in the vicinity of PALAOA.

Temperature, and hence sound speed showed little variation during the periods observed. A mean sound speed of 1419 m s^{-1} was calculated. Extremes were determined to be 1424 m s^{-1} (maximum) respectively 1415 m s^{-1} (minimum).
Furthermore, the interface with the meteorological data base was successfully established. The data are stored in the daily syslog file and allow (1) development of decision rules for shutting down consumers of low priority and (2) integration of these data in the analysis of the vocal behavior of marine mammals.

2.6 Data management

As described in detail in the previous chapter, the PALAOA recordings are being sent continuously via the IntelSat satellite link to the AWI in Bremerhaven. The Ogg-Vorbis compressed audio stream (24 kBit) is stored on a PC located at the institute. A backup of the stream is copied daily to the institute's data silo. Once a year, the highly compressed audio stream data are replaced by the high quality stream data, which were stored at Neumayer Base and are shipped on LTO-2 tapes. A high quality Mp3 file of one minute duration has a size of about 1.25 MB. Thus, during one day (year) of operation 1440 files (525600 files) and 1.7 GB (620 GB) of data are generated. This estimation comprises the mp3 stream data only. Broadband data recorded occasionally during polar summer with the high quality recording system or records of PALAOA-S are not included in this budget.

The amount of additional data (webcam pictures, CTD data, operating data, meteorological data and network statistics) sent over the satellite link is about 30 MB (10 GB) per day (year) and carries little weight. Between 2005 and 2007, a total 4 TB of data were collected and stored in the AWI data silo (which has a total capacity of 1 PByte). All PALAOA data sets will be published in the near future under open access license in PANGAEA (http://www.pangaea.de) - a public digital library for science aimed at archiving, publishing and distributing geo-referenced data with special emphasis on environmental, marine and geological basic research. PANGAEA guarantees long-term availability of scientific primary data related to publications. Each dataset can be identified, shared and published by a persistent Digital Object Identifier (DOI). In addition a full citation for sets of data can be defined on request. The PANGAEA policy of data management and archiving follows the Principles and Responsibilities of ICSU World Data Centres and satisfies the conditions of the Berlin Declaration on "Open Access to Knowledge in the Sciences and Humanities".

2.6.1 Data processing

The PALAOA acoustic data set consists of standard multimedia files (Wav, Flac, Pac, Mp3 and Ogg-Vorbis). Additional metadata are stored in text files. All data are accessible on a network drive allowing standard software to preview the dataset. However, the large amount of files severely hampers the analysis of longer periods as standard software cannot load and handle thousands of files at once.

For this reason a Matlab™ based data interface (named PALAOAdb) was developed, to allow easy access to the acoustic and metadata files, using a timeline or event oriented (entries in the log file made by users) GUI (see Figure 2.31).

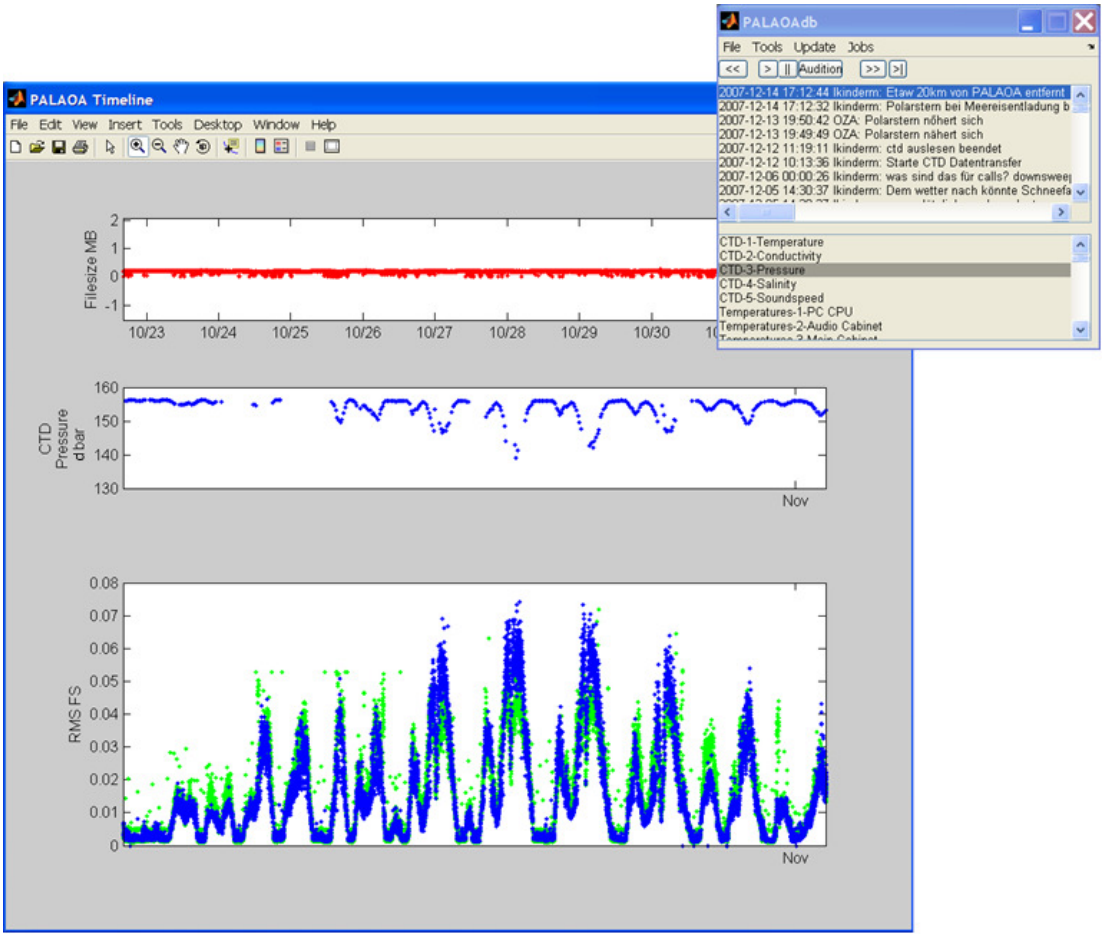


Figure 2.31: PALAOAdb interface and timeline display. Panel on top: file size [MB]; panel in the middle: CTD pressure [dbar]; Panel on bottom: sound intensity [voltage RMS FS] (full scale) of hydrophone WEST (blue) and hydrophone CENTRAL (green). Display in the upper right shows the control interface.

Once daily, PALAOAdb analyzes the data recently received and updates the associated database. Upon start-up, PALAOAdb generates, from a set of selectable parameters, a plot of the entire recording period. Currently sound specific measures (e.g. RMS or peak sound levels) and metadata information (e.g. air and water temperature or pressure indicating currents) may be selected. PALAOAdb permits zooming in or clicking on specific data points. In the latter case, the corresponding sound file is opened, either by a built-in player or by an external program like Adobe Audition™, Audacity, Ishmael or XBAT. Loading and processing a single mp3 file with the basic statistics like calculating RMS and peak sound levels takes about 1s. Additional analysis modules, such as pattern recognition algorithms, can be (and have been) embedded into PALAOAdb. The algorithms are implemented as Matlab™ functions, with a one-minute waveform array and a file information structure as input and a structure containing the result as output. PALAOAdb displays the

results of the analyses as timelines. This allows human operators to explore the correlation of the vocalization behavior of target species with the tidal current.
A (basic) offline analysis can be processed at up to 50 times real-time, but processing speed strongly depends on the complexity of the analysis algorithms. The amount of available data, for example, prohibits running complex (time consuming) detection algorithms across the entire data set in a reasonable amount of time while using a single PC. To shorten the processing time, a distributed computing system that allows sharing tasks between multiple computers was implemented. The main task (i.e. applying a detection algorithm to the entire data set) is split into smaller jobs (defined by m-files and containing information on for example 200 single files to be analyzed) and placed in a central network directory which can be accessed by the ”workers”. These ”workers” are executables, which are generated by PALAOAdb and contain the analysis algorithms. The ”workers” can be executed on any PC, without the need of having Matlab™ installed.

All communication between the workers and the central network directory containing the jobs is executed via the file system: any computer with access to the institute’s computer network can potentially contribute to the analysis by executing the worker executable. After executing the exe-file all necessary routines are installed automatically and the processing of jobs starts. The results of the analysis are appended to the m-file describing the job. After the computation is finished the worker renames the m-file located on the central network drive to indicate that this job is completed and starts a new job. Finally PALAOAdb collects all m-files containing the computational results and generates a single structure array which is added to the data base. This structure can be accessed in the same way as the raw audio data or the metadata.

As there is no ”interprocess communication” between the workers, the speed scales linearly with the number of PCs as long as the network and the file server are not saturated by fetching the audio files. With a size of 1.25 MByte per mp3 file of one minute duration, a 1 GBit network as currently installed in the AWI can handle up to 100 files per second. Thus, the processing speed can be increased 100 times using 100 PCs simultaneously. The AWI network is currently being upgraded to a bandwidth of 10 GBit which would (theoretically) allow running 1000 PCs simultaneously in the future. Thus, complex and time consuming detection algorithms could be applied to the entire PALAOA data set, and results could be expected within a few hours or days.

3. Leopard seal underwater vocalizations and their characteristics

3.1 Call types and frequency of occurrence

3.1.1 Manual analysis of the audio stream

The PALAOA audio stream was analyzed manually to identify which of the various leopard seal call types – as known from the literature – occur in the vicinity of PALAOA and to determine call frequency of occurrence. A one week data set - spanning 21 - 27 December 2006 - was manually analyzed. Listening to the PALAOA acoustic livestream, it became apparent that this period included very high levels of leopard seals calling activity. This period coincides with the leopard seal mating season, i.e. late December, during which the animals are known to be highly vociferous (see chapter 1.3.3.6).

The data set used for manual scanning consisted of 1344 mono wav files of 1 minute duration (sampled with 11025 Hz/16 bit). For each hour, 8 files were selected (minutes 00 and 01; 15 and 16; 30 and 31; 45 and 46) within which the number and type of vocalizations were counted. The resulting time series of counts were used to determine the overall diurnal pattern of calls, to develop the detection algorithms, and to evaluate detector performance.

In order to ascertain whether the acoustic repertoire of the leopard seal differs between geographic locations, a comparison between recordings made at Atka Bay and Drescher Inlet was made. The DIPS (Drescher Inlet Pilot Study) data set was recorded in the Drescher Inlet (72°50'S, 19°02'W), a 25-km-long crack in the Riiser Larsen Ice Shelf between 21 and 25 December. A sub data set consisting of 960 audio files of 1 minute duration (each hour the minutes 00 and 01; 15 and 16; 30 and 31; 45 and 46) and sampled with 48000 Hz/16 bit was extracted from the DIPS data set and manually analyzed.

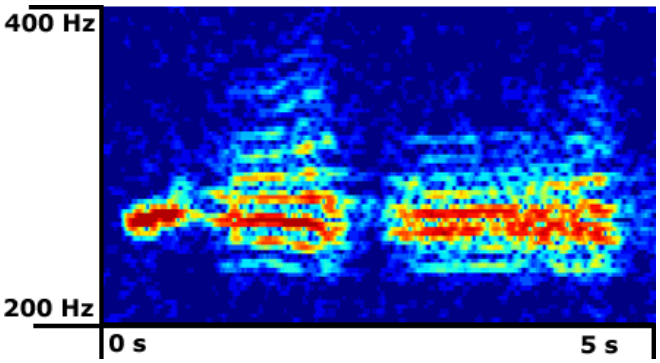
3.1.2 Results and discussion

3.1.2.1 Leopard seal call types in the vicinity of PALAOA

The main underwater acoustic repertoire of the leopard seal - as known from the literature - consists of eight different calls. In the vicinity of PALAOA all eight call types have been detected by verifying recorded spectrograms with those published in literature. However, it is unknown, whether or not any additional calls of leopard seals exist in the PALAOA data as no measurements were performed on individual animals directly. The assignment of the various call types to leopard seals was hence indirect by comparing spectrograms from the PALAOA recordings with spectrograms published in literature.

Figures 3.1 - 3.8 provide an overview of the eight call types as obtained from the PALAOA data set. The parameters on the right provide information on the maximum range of each call type. Individual calls – as depicted on the left – however, do not necessarily span this entire range. Note that varying scales have been used in the spectrograms to best resolve details of each call type.

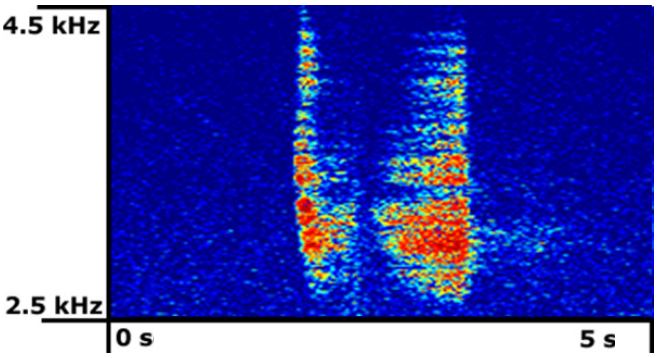
(a) The low double trill (LDT):



General description:	Low frequency call which consists of two pulsed elements and a initial narrow-band component.
Min. freq.	230 Hz
Max. freq.	470 Hz
Min. dur.	1.9 s
Max. dur.	9.5 s

Figure 3.1: Typical spectrogram of a low double trill.

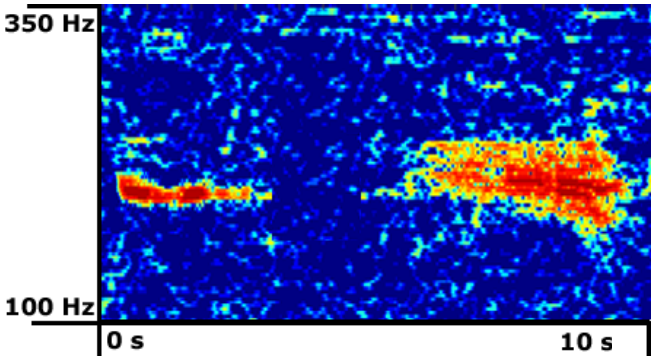
(b) The high double trill (HDT):



General description:	Broadband high frequency call which consists of two pulsed elements.
Min. freq.	2.500 Hz
Max. freq.	4.450 Hz
Min. dur.	1.9 s
Max. dur.	9.0 s

Figure 3.2: Typical spectrogram of a high double trill.

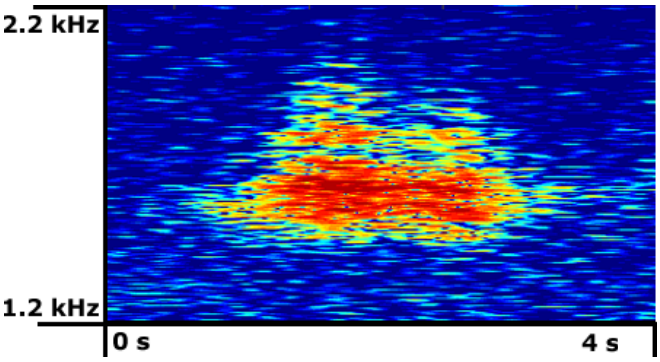
(c) The hoot with single trill (HST):



General description:	Low frequency call which consists of a narrow-band component followed by a single pulsed element.
Min. freq.	170 Hz
Max. freq.	310 Hz
Min. dur.	3.0 s
Max. dur.	9.5 s

Figure 3.3: Typical spectrogram of a hoot with a single trill.

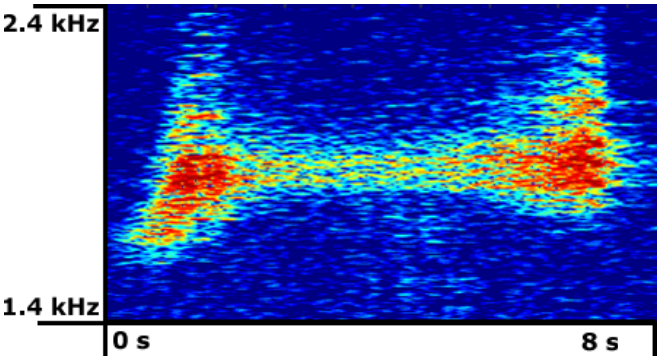
(d) The medium single trill (MST):



General description:	Mid frequency call which consists of a single pulsed element.
Min. freq.	1.300 Hz
Max. freq.	2.400 Hz
Min. dur.	2.9 s
Max. dur.	6.7 s

Figure 3.4: Typical spectrogram of a medium single trill.

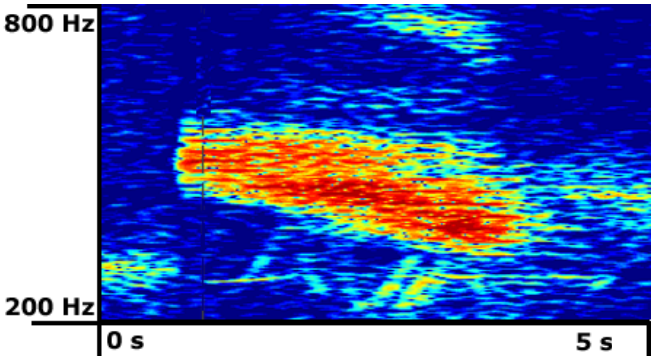
(e) The medium double trill (MDT):



General description:	Mid frequency call which consists of two pulsed elements.
Min. freq.	1.400 Hz
Max. freq.	2.500 Hz
Min. dur.	6.0 s
Max. dur.	9.0 s

Figure 3.5: Typical spectrogram of a medium double trill.

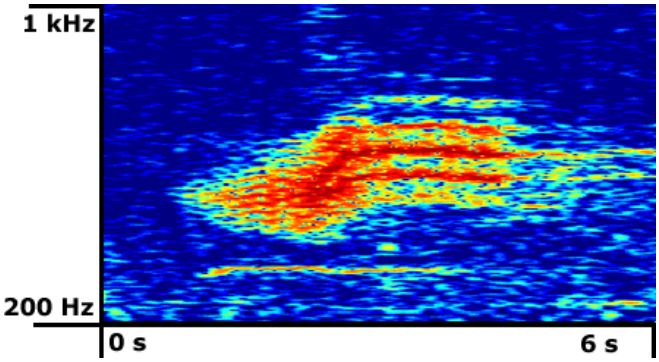
(f) The low descending trill (LDST):



General description:	Low frequency call which consists of a single pulsed element of decreasing frequency.
Min. freq.	270 Hz
Max. freq.	800 Hz
Min. dur.	4.0 s
Max. dur.	6.8 s

Figure 3.6: Typical spectrogram of a low descending trill.

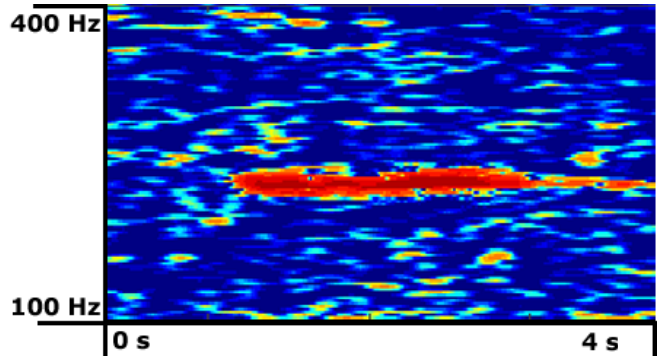
(g) The low ascending trill (LAST):



General description:	Mid frequency call which consists of a single pulsed element of ascending frequency.
Min. freq.	950 Hz
Max. freq.	400 Hz
Min. dur.	3.4 s
Max. dur.	6.2 s

Figure 3.7: Typical spectrogram of a low ascending trill.

(h) The hoot (H):



General description:	Low frequency call which consists of a single tonal element.
Min. freq.	290 Hz
Max. freq.	150 Hz
Min. dur.	3.5 s
Max. dur.	1.5 s

Figure 3.8: Typical spectrogram of a hoot.

3.1.2.2 Geographic occurrence of call types

A comparison of the leopard seal's acoustic repertoire at different geographic locations is given in Table 3.1.

Table 3.1: Leopard seal call types at different geographic locations.

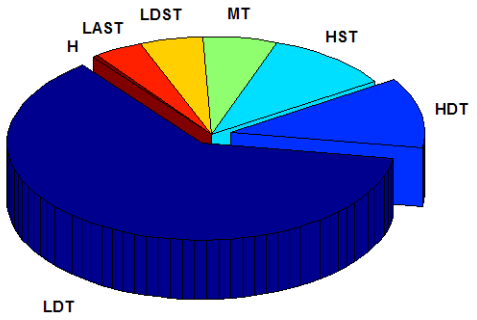
Location →	***Atka Bay (PALAOA)***	***Drescher Inlet (DIPS)***	***Prydz Bay***[3.1]	***Mc Murdo***[3.2]	***Palmer Peninsula***[3.2]	***South Shetland Islands***[3.3]
Period covered →	*21 - 27 Dec. 2006*	*21 - 25 Dec. 2003*	*Nov. - Jan. 93/94 & 94/95*	*Austral spring '79*	*Austral spring '78*	*6 Oct. - 2 Nov. '76 22 Oct. - 20 Nov. '77*
Call: LDT	**yes**	**yes**	**yes**	**yes**	**yes**	**yes**
HDT	**yes**	**yes**	**yes**	**yes**	**yes**	**yes**
HST	yes	yes	yes	no	no	yes
MST	yes	yes	yes	yes	yes	no
MDT	yes	yes	no	no	no	yes
LAST	yes	yes	yes	yes	yes	no
LDST	yes	yes	yes	yes	yes	no
H	yes	yes	no	no	no	no

[3.1] Rogers *et al.* (1995).
[3.2] Thomas & Golladay (1995).
[3.3] Stirling & Siniff (1979).

The table illustrates that only LDTs and HDTs occur at every location listed (which – to my knowledge – comprises all locations where leopard seals have so far been studied acoustically).

3.1.2.3 Frequency of occurrence of call types

Figure 3.9 gives an overview of the frequency of occurrence of each call type recorded at PALAOA (derived from the data set described in chapter 3.1.1).



(LDT): Low double trill – 62.2%

(HDT): High double trill – 11.8%

(HST): Hoot with single trill – 10.2%

(MT*): Medium trill – 6.1%

(LDST): Low descending trill – 5.2%

(LAST): Low ascending trill – 4.1%

(H): Hoot – 0.4%

Figure 3.9: Call types of the leopard seal recorded in the vicinity of PALAOA.
*MT (medium trill) = MST (medium single trill) + MDT (medium double trill)

By far most frequently recorded vocalization is the low double trill (62.2%) followed by the high double trill with 11.8% (DIPS: 52% LDTs and 9% HDTs). Together, they comprise nearly three quarters of all leopard seal calls which makes these two call types promising candidates for a reliable automated detection. The HDT also offers the possibility to detect leopard seals in the presence of (ship) noise because of its higher frequency, where ship noise typically declines.

To estimate an error for the derived frequency of occurrence of the LDT and HDT in the PALAOA recordings, a bootstrapping approach was conducted. Bootstrapping is a statistical method which repeatedly calculates statistical parameters on randomly chosen sub data sets of a primary data set. The range of the calculated parameters provides information on the accuracy of the derived frequency of occurrence. For the PALAOA data set it worked as follows: The manually analyzed PALAOA data set consists of 1343 entries containing the associated number of vocalizations of each call type. The bootstrapping routine randomly extracted 200 entries of the data set and calculated the frequency of occurrence of the LDT and HDT for this sub data set. This procedure was repeated 1000 times. The results of the bootstrapping analysis are comprised in Table 3.2.

Table 3.2: Results of the Bootstrapping.

Frequency of occurrence	*Min.*	*Max.*	*Mean*	*Median*	*Std. dev.*
LDT [%]	57.9	65.7	62.0	62.0	1.3
HDT [%]	9.5	14.0	11.7	11.7	0.7

The calculated minimum and maximum values as well as the standard deviation values indicate that the error in the calculated frequency of occurrence of the LDT and HDT is low.

3.2 Characteristics of the most prominent leopard seal vocalizations

3.2.1 Manual analysis of audio samples

To develop an effective and robust (against masking noise) automated detection algorithm, typical call type characteristics including their respective ranges need to be known. To evaluate the range of acoustic characteristics it is necessary to analyze a large set of samples for both call types. To facilitate this process, a Matlab™-based GUI (graphical user interface) was programmed (see Figure 3.10).

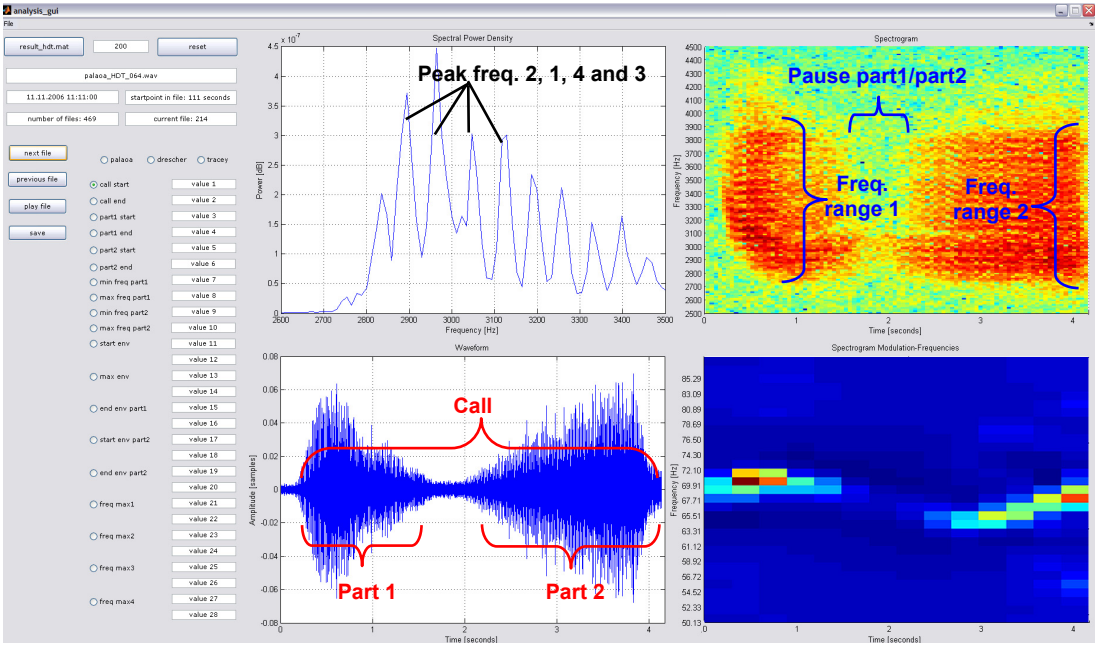


Figure 3.10: The Matlab™ based GUI. Upper left panel: spectral power density of a HDT; upper right panel: spectrogram of a HDT; lower left panel: waveform of a HDT; lower right panel: spectrogram of the pulse repetition rate of the HDT.

This GUI provides concurrent displays of spectral power density (upper left), waveform (lower left), spectrogram (upper right) and spectrogram of the modulation frequency (lower right) panel. It simplified and accelerated the call measurement by automatically assigning graphically selected values to a preselected set of characteristic parameters and immediately calculating derived values. In Figure 3.10, for example, a HDT is displayed and annotated

with its corresponding characteristic parameters. In detail the following characteristic parameters were extracted:

HDTs and LDTs:

(01) Call start time [s]

(02) Call end time [s]

(03) Call duration [s] = (02) – (01)

(04) Part 1 start time [s]

(05) Part 1 end time [s]

(06) Part 1 duration [s] = (05) – (04)

(07) Part 2 start time [s]

(08) Part 2 end time [s]

(09) Part 2 duration [s] = (08) – (07)

(10) Duration of the pause between Part 1 and 2 [s] = (07) – (05)

(11) Part 1 minimum frequency [Hz]

(12) Part 1 maximum frequency [Hz]

(13) Part 1 frequency bandwidth [Hz] = (12) – (11)

(14) Part 2 minimum frequency [Hz]

(15) Part 2 maximum frequency [Hz]

(16) Part 2 frequency bandwidth [Hz] = (15) – (14)

(17) Peak frequency 1 [Hz] of most energetic peak (highest peak)

(18) Peak frequency 2 [Hz] of second highest peak

(19) Peak frequency 3 [Hz] of third highest peak

HDTs only:

(20) Peak frequency 4 [Hz] of fourth highest peak

LDTs only:

(21) Frequency of the narrow-band component [Hz]

(22) Start time of the narrow-band component [s]

(23) End time of the narrow-band component [s]

(24) Duration of the narrow-band component [s] = (23) – (22)

(25) Duration of the of the pause between the narrow-band comp. and Part 1 [s] = (04) – (22)

In the following section/paragraph, the results of this analysis are described in detail.

3.2.2 Results and discussion

3.2.2.1 The low double trill (LDT)

A total of 150 LDTs recorded at PALAOA station were analyzed to gain information about the characteristics of this call type. Waveform and spectrogram of a low double trill call are depicted in Figure 3.11.

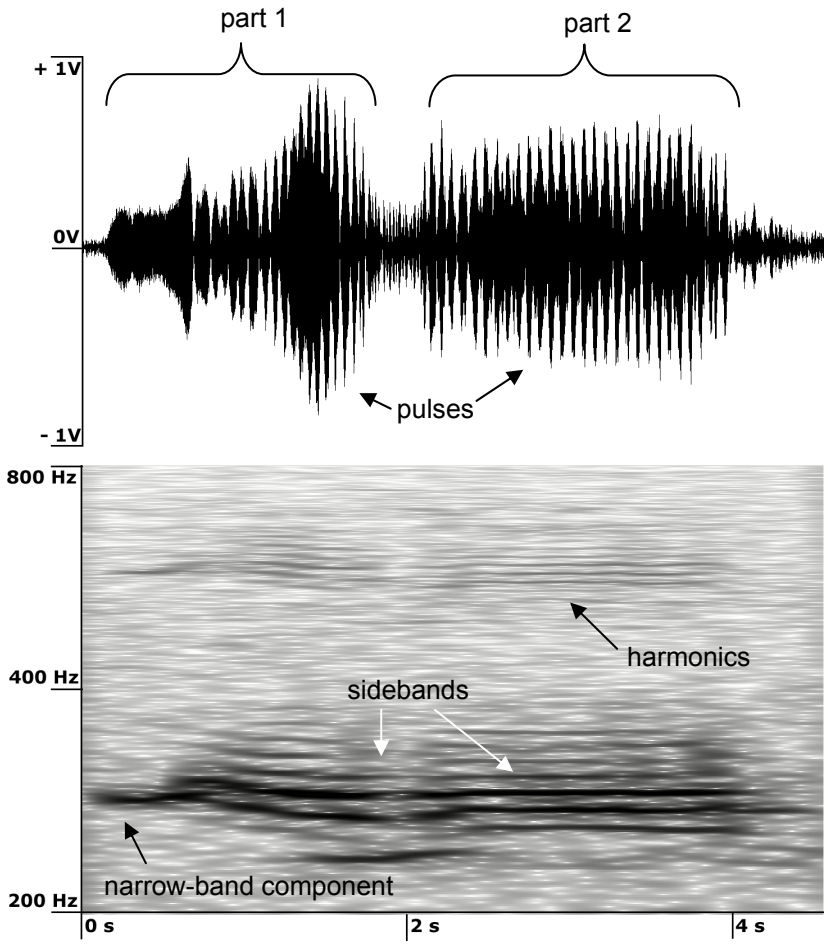


Figure 3.11: Waveform and spectrogram of a low double trill.

The call starts with a narrow-band component followed by two segments, each consisting of a series of short pulses. These pulses cause an amplitude modulation of the main signal which generates so-called sidebands. The sideband structure is revealed by the spectrogram in Figure 3.11 (bottom). In general, the number of sidebands is determined by the type of amplitude modulation. In the case of a sinusoidal modulation, only two sidebands are generated while the primary frequency is rendered invisible in the spectrogram. By contrast, triangular or rectangular modulations cause multiple (> 2) sidebands (primary frequency also invisible). The type of amplitude modulation of the LDT is in transition between a sinusoidal and triangular modulation. The exact number of visible sidebands in the spectrogram is also dependent on the SNR of the call. Furthermore, the spectrogram in Figure 3.11 displays the 1st harmonic (double frequency) of the LDT. In Table 3.3 the characteristic parameters of the LDTs are summarized.

Table 3.3: Acoustic features of the LDTs recorded at PALAOA.

Parameter	*Min.*	*Max.*	*Mean*	*Median*	*Std. dev.*
Call duration [s]	1.90	9.52	3.77	3.59	1.03
Duration NBC [s]	0.12	1.14	0.39	0.37	0.16
Part 1 duration [s]	0.13	2.94	1.23	1.15	0.49
Part 2 duration [s]	0.89	4.30	1.90	1.81	0.49
Duration pause 1 [s]	0.00	2.22	0.07	0.01	0.24
Duration pause 2 [s]	0.02	1.27	0.25	0.25	0.16
Freq. NBC[3.4] [Hz]	272	361	330	334	21
Min. freq Part 1 [Hz]	232	359	304	311	22
Max. freq Part 1 [Hz]	270	472	399	397	38
Bandwidth Part 1 [Hz]	34	156	97	93	27
Min. freq Part 2 [Hz]	230	401	298	301	23
Max. freq Part 2 [Hz]	297	471	396	395	35
Bandwidth Part 2 [Hz]	40	166	100	100	28
Freq. max 1 [Hz]	270	363	329	332	21
Freq. max 2 [Hz]	258	371	334	340	27
Freq. max 3 [Hz]	246	399	331	332	30

Specific for this call type is its highly variable (1.9 s - 9.5 s) duration, which complicates its automatic detection with standard methods such as matched filtering (see chapter 1.2.2.2). The fundamental frequencies of the analyzed calls (incl. the sidebands) occurred in a frequency band between 230 Hz and 470 Hz, a frequency range that unfortunately is easily masked by ship noise.

3.2.2.2 The high double trill (HDT)

A total of 150 HDTs were analyzed to gain information about the temporal structure and the frequency characteristics of this call type (Table 3.4).
Figure 3.12 (top) depicts the waveform and spectrogram of a high double trill. The waveform clearly shows that the call is separated into two segments which consist of a series of short pulses. These pulses cause - equivalent to the LDT - an amplitude modulation of the main signal. The type of amplitude modulation of the HDT is in-between a sinusoidal and triangular modulation.

[3.4] Freq. NBC = frequency of the narrow-band component.

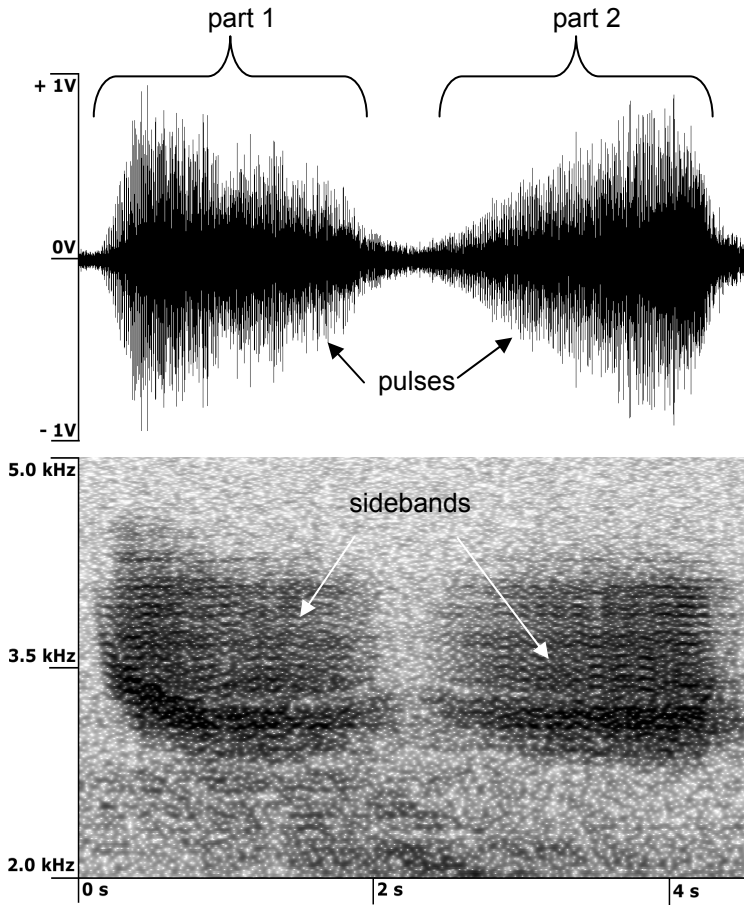


Figure 3.12: Waveform and spectrogram of a LDT.

The HDTs features, like the LDT, a highly variable call duration ranging from 1.9 s to 9.0 s. The calls cover a frequency range between 2500 Hz and 4450 Hz (Table 3.4).

Table 3.4: Acoustic features of the HDTs recorded at PALAOA.

Parameter	*Min.*	*Max.*	*Mean*	*Median*	*Std. dev.*
Call duration [s]	1.92	9.02	3.92	3.74	1.22
Part 1 duration [s]	0.67	3.88	1.69	1.65	0.68
Part 2 duration [s]	0.72	3.29	1.64	1.61	0.48
Duration of pause [s]	0.09	1.80	0.61	0.52	0.36
Min. freq. part 1 [Hz]	2533	2835	2673	2675	62
Max. freq. part 1 [Hz]	3979	4462	4202	4203	97
Bandwidth. Part 1 [Hz]	1197	1858	1529	1519	123
Min. freq. part 2 [Hz]	2507	2799	2610	2613	51
Max. freq. part 2 [Hz]	3859	4360	4065	4063	102
Bandwidth. Part 2 [Hz]	1202	1805	1454	1435	108
Frequency max 1 [Hz]	2778	3057	2904	2907	52
Frequency max 2 [Hz]	2740	3132	2913	2905	69
Frequency max 3 [Hz]	2728	3284	2915	2908	97
Frequency max 4 [Hz]	2670	3209	2938	2940	125

These results provide basic information on the acoustic features of the LDT and the HDT which are important for the development of the detection algorithms. The analysis revealed that LDTs and HDTs are highly variable in their duration (between ~ 2 and 9 s). This has to be considered when choosing the detection method, as common detection methods such as matched filtering are not suitable for call types with a highly variable duration.

4. Automated detection of leopard seal vocalizations

4.1 Identifying leopard seal calls in the Antarctic underwater soundscape

The PALAOA recordings are dominated by a high degree of natural abiotic and biotic activity. During austral summer in particular, vocalizations of Weddell seals (*Leptonychotes weddellii*), Ross seals (*Ommatophoca rossii*), crabeater seals (*Lobodon carcinophaga*), leopard seals (*Hydrurga leptonyx*) and various baleen (Mysticeti) and toothed (Odontoceti) whale species are omnipresent, and – on occasion – result in a nearly inextricable cacophony of sounds (see Figure 4.1). Abiotic contributions to the Antarctic underwater soundscape are predominantly caused by the motion of ice. The most prevalent of these signals is due to the calving of glacial ice, which happens throughout the year. Calving generates broadband sounds, which last for a few seconds and range from a few Hertz to several Kilohertz.

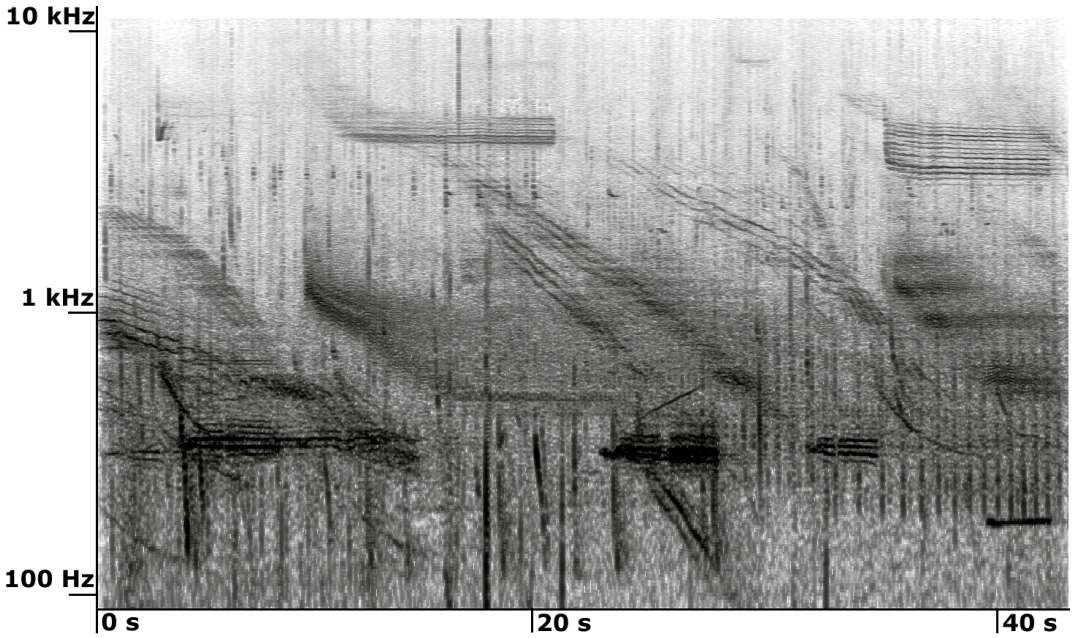


Figure 4.1: Spectrogram of a sound file recorded at PALAOA in austral summer 2006. The dark horizontal and vertical lines represent vocalizations of marine mammals.

The overlapping biotic and abiotic sounds complicate the detection of particular “target” sounds - such as specific mammalian call types - significantly.

4.2 Automated detection of LDTs using a zero-crossing rate (ZCR) detector

The manual analysis of the LDTs (see chapter 3.2.2.1) revealed a high variability in their duration, while spanning a narrow band of rather low frequencies (230 Hz – 470 Hz). A detection algorithm was developed for LDTs, based on the zero-crossing rate (ZCR) of the (band-passed) signal. In the following sections, the ZCR detector, its performance on a test data set, and a

comparison with a commonly used detector based on spectrogram correlation (SC detector) will be described in detail.

4.2.1 Methods

4.2.1.1 Introduction to the zero-crossing rate (ZCR)

An averaged zero-crossing rate (ZCR) is often used in human speech recognition to detect vocalizations and also has applications in simple pitch detectors (Rabiner & Juang 1993). The ZCR rate is – for a given time-frame of length Δt – directly derived from the (band-passed) waveform of the recorded signal. Whenever the waveform of a signal crosses the mean (DC) signal level of the respective time-frame, a counter is incremented. The final number of counts for this time-frame represents its zero-crossing rate:

$$ZCR(t) = \frac{1}{\Delta t} \sum_{n=1}^{N} I\left\{\left(s_n * s_{n+1}\right) \leq 0\right\}$$

with *s* = signal of length Δt, N = sampling rate * Δt and the indicator function $I\{A\}$ is 1 if its argument *A* is true and 0 otherwise. The calculation of the ZCR (unit: Hz) is rather fast due to the fact that the calculations are based on the signal waveform, avoiding the need of transforming the signal into the frequency domain (FFT).

The ZCR depends strongly on the characteristics of the investigated signal. Signals of constant frequencies (like the LDT), for example, are characterized by a constant zero-crossing rate. The signals' frequency is a key factor in determining the zero-crossing rate, with high/low frequencies generating high/low zero-crossing rates. Usually, background noise is broadband and contains frequencies between 1 Hz and the Nyquist frequency. Hence the ZCRs of both a high frequency sound and the common background noise are high and their difference small. By contrast, zero-crossing rates of narrow-band, low frequency signals such as the LDT are low, and clearly distinguishable from the higher ZCR typical for noise.

Important note: The zero-crossing rates described in the following represent normalized values. That means the ZCR of a signal with the Nyquist frequency (in this case 11025 Hz) equals 1 Hz. This was necessary as the ZCR detector was implemented in the XBAT[4.1] (Extensible Bioacoustic Tool) software which is designed to handle feature values between 0 and 1.

4.2.1.2 Design of the zero-crossing rate (ZCR) detector

First, a (software) filter bank consisting of a high-pass and a band-stop filter was applied to the acoustic data. The filtering scheme is shown in Figure 4.2. The acoustic data are classified into four parts. Part 1 includes the low frequency

[4.1] www.xbat.org

noise. In the PALAOA recordings, this part (in which noise is mostly caused by cable strumming due to tidal flow) must be removed by applying a high-pass filter. Because of their high amplitude, these low frequency signals act similar to a carrier wave which inhibits zero-crossings of signals of higher frequency and lower amplitude.

A final stage of the detector searches for periods of time in which the ZCR function is relatively constant. Such periods are found with the relation:

$$|ZCR(k + 1) – ZCR(k)| < tol$$

where k is a sample number and tol is a tolerance factor, which is discussed further below. A call detection occurs when this relation is true for a specified duration or longer, here 1.7 s.

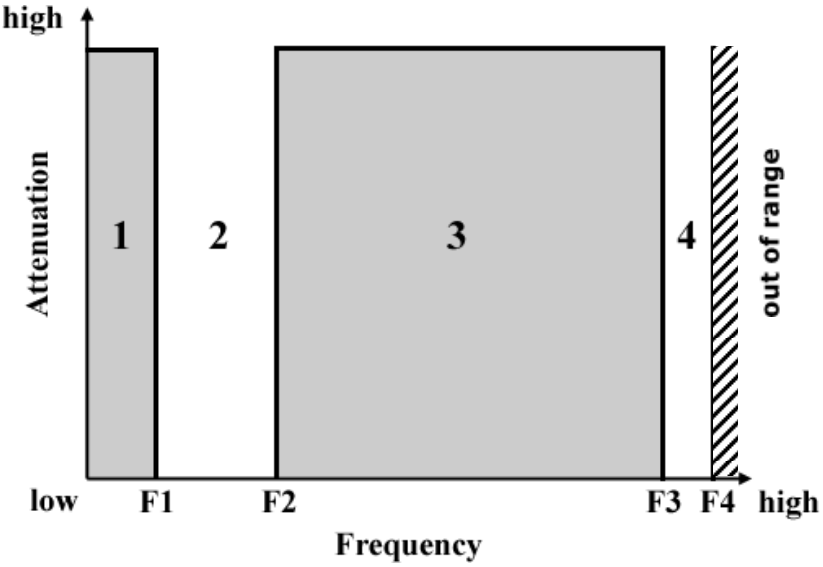


Figure 4.2: Filtering scheme of the zero-crossing detector. Pass-bands are shown in white, stop-bands in gray.

Part 2 is described as the “target frequency band” (between frequency F1 and F2), and contains the call type of interest. Part 3 contains signals which could interfere with the target vocalizations. Signals of high amplitude (compared to the target signal) occurring concurrently in this frequency band would mask the lower ZCR of the target signal and therefore have to be removed by applying a band-stop filter. Part 4 is the so-called noise gate. The noise gate is added to the target frequency band. This gate is located between frequency F3 and F4 (F4 equals the Nyquist frequency of the sound file). The noise gate is described by a high frequency band because: (a) high frequencies imply high ZCR which generate a high contrast to the ZCR in the target frequency band and (b) this frequency band is normally of low energy and likely does not contain interfering signals such as vocalizations from other marine mammals.

In case of the leopard seal vocalization, the target frequency was set between 250 Hz (F1) and 400 Hz (F2). The noise gate is located between 5250 Hz (F3) and 5512.5 Hz (F4 = Nyquist frequency @ 11025 Hz sampling rate). An

example of the ZCR detector output is shown in Figure 4.5. The data (spectrogram) depicts a typical low double trill vocalization starting with a tonal component followed by two pulsed segments. As expected, the ZCR (Figure 4.3 bottom) is low and nearly constant throughout the call, due to the calls' tonal character. By contrast, the gap between the two pulsed elements is characterized by a significantly higher ZCR.

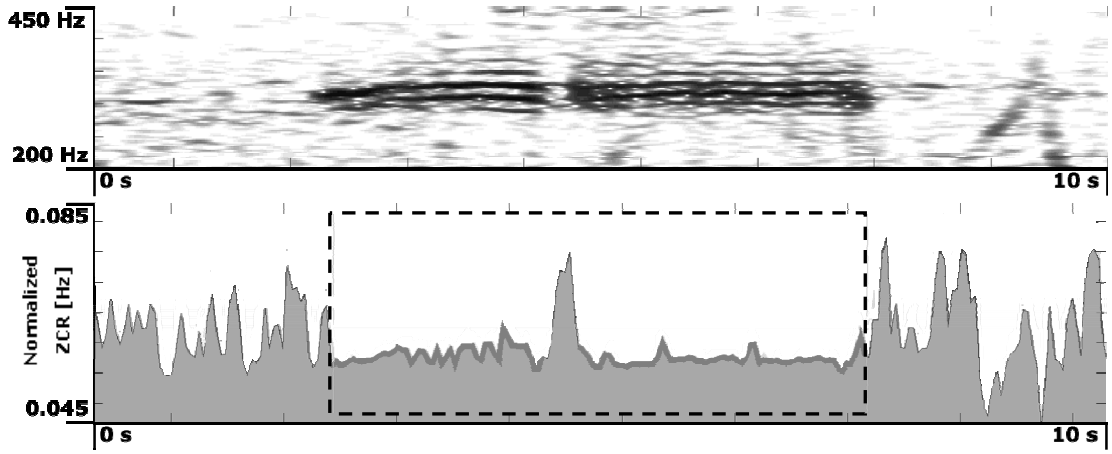


Figure 4.3: Spectrogram (top) and ZCR rate (bottom) of a LDT. The dashed box indicates a period with a relatively constant ZCR.

The ZCR detector was developed and implemented into XBAT (Extensible Bioacoustic Tool). XBAT is a sound analysis application and Matlab™ platform, which can be extended with user defined Matlab™ scripts for advanced sound analysis. It is open-source, and licensed under the GPL[4.2]. The XBAT software platform is already equipped with efficient data I/O interface, filters and other useful functions, enabling the developer to focus on the development of the proper detection method.

The detector was applied to a test data set (mono wav files sampled with 11025 Hz @ 16 bit) executed on a desktop PC (single Intel Pentium IV processor with 3.4 kHz and 2 GB RAM) at 60 times real-time. This high speed allows screening extensive data sets within a reasonable amount of time.

4.2.1.3 Problems and limitations

The analysis of the output of the detector revealed that its performance depends critically on the SNR (to be discussed in chapter 4.2.2) and frequency of the target signal. As the detector is searching for events where the ZCR stays constant for a certain period (at least 1.7 sec), calls with a fragmented ZCR were missed. To reduce the number of missed calls, a morphological filter (which was already implemented in XBAT and is available via the main filter pallet - for more information on morphological filter see Heijmans 1997) was integrated to bridge gaps between fragments of constant ZCR. This morphological filter can be adjusted by changing two threshold values. The first threshold defines the maximum admitted duration of the gap between two neighboring fragments (selected to be < 1.5 s), while the second parameter

[4.2] http://www.gnu.org/licenses/gpl-3.0.html

defines the maximum admitted difference between the mean values of the ZCR of two neighboring gaps (tolerance factor (tol) selected to be < 0.005 Hz). If the parameters of two neighboring fragments are in the range of both thresholds, they are joined.

There is also the possibility of false positive detections[4.3] caused by interfering biotic and abiotic narrow-band sounds in the target frequency band. This problem can be addressed by:

- Manual validation of the detection in questions by looking at the corresponding spectrogram and listening to the sound snippet, or by
- the implementation of a scoring scheme.

In this study, a score was calculated on the fly from the following characteristic parameters of the time series of the ZCR, whenever the ZCR was stable (+/- 0.01 Hz) for a segment of at least 1.7 seconds:

- The segment averaged zero-crossing rate,
- the variability of the zero-crossing rate with the segment,
- the difference between ZCRs at segment start and end time and
- the duration of the segment.

The calculation of the score consisted of a two step process. First, the score of all segments fulfilling the following conditions was set (irrevocably) to zero if:

- The mean zero-crossing rate was < 0.04 Hz or > 0.07 Hz
- or the absolute difference between start and end values of the ZCR was > 0.05 s^{-1}
- or the length of the fragment was > 9.5 s.

For the remaining segments, the score was calculated from the variability of the ZCR – which is a proxy for the SNR - for the respective segment. The score (*S*) ranges from 0 to 1 and is calculated (formula empirically derived) as:

$$S = 1 - \left(\frac{5 * \sigma(ZCR)}{100} \right)$$

Zero variation within the ZCR (i.e. constant ZCR) results in a high score (e. g. > 0.8) while high variability will cause lower scores. Thus it is possible after a detection run to select only calls of high score. Other factors (e.g. call duration) were not included into the calculation of the score because of the high variability of this parameter (see chapter 3.2.2.1). For those parameters it is more suitable to set an absolute upper and lower limit.

[4.3] False positive means the system detected a signal which is not the right one. By contrast, false negative means the number of missed calls which were not detected.

4.2.2 Results and discussion

In the following chapter, the performance of the ZCR detector will be evaluated. In addition, a second well-established spectrogram correlation (SC) detector was applied to the test data set. Finally, the performances of both detectors are compared and discussed.

Prior to the test run[4.4], the SNRs (in the frequency band between 250 Hz and 450 Hz) of 12 LDTs per day (every two hours - one LDT at hourly minute 00 or 01) were analyzed. SNRs were determined from the ratio of the calls RMS sound pressure level and the RMS sound pressure level of the 0.5 s prior to the call. This analysis showed that the SNRs of LDTs were significantly higher in the period 21 December to 23 December 2006, as compared to the period 24 December to 27 December 2006. For this reason the results of the test run are discussed separately for these two periods (see Table 4.1).

Table 4.1: SNR of selected LDTs.

	No. of samples	*Min. SNR [dB]*	*Max. SNR [dB]*	*Mean SNR [dB]*	*Median SNR [dB]*	*Sigma SNR [dB]*
Period 1 (21.12 – 24.12)	36	5.7	31.6	16.2	16.8	6.4
Period 2 (25.12 -27.12)	48	3.4	24.4	11.3	11.0	4.9

To get information on the possible reasons for the drop in the SNR of the LDTs, the background noise level in the target frequency band and the meteorological conditions measured at Neumayer Base were analyzed. The mean background noise level in the target frequency band was calculated with -58 dBFs[4.5] (sigma: 3 dB) for the first period, and -57 dBFs (sigma: 2 dB) for the second period. The analysis of the meteorological data revealed that at 24 December 2006 between 12 am and 2 pm the wind speed increased from 2 m s^{-1} to 12 m s^{-1}. Although there was no observable significant difference in the background noise level (250 – 450 Hz) measurements of period one and two, the increased wind speed might have influenced the SNR indirectly and the seals could have moved away with the drifting ice floes.

4.2.2.1 Performance of the ZCR detector

In a first step the performance of the ZCR detector was compared against the manually analyzed test data set described in paragraph 3.1. The results of the test run are presented in Figure 4.4. The exact numbers of detections are provided in Table 4.2.

[4.4] The term ”test run” hereinafter describes the application of the detector to the test data set.
[4.5] dBFs means "decibels full scale". 0 dBFS is assigned to the maximum possible level.

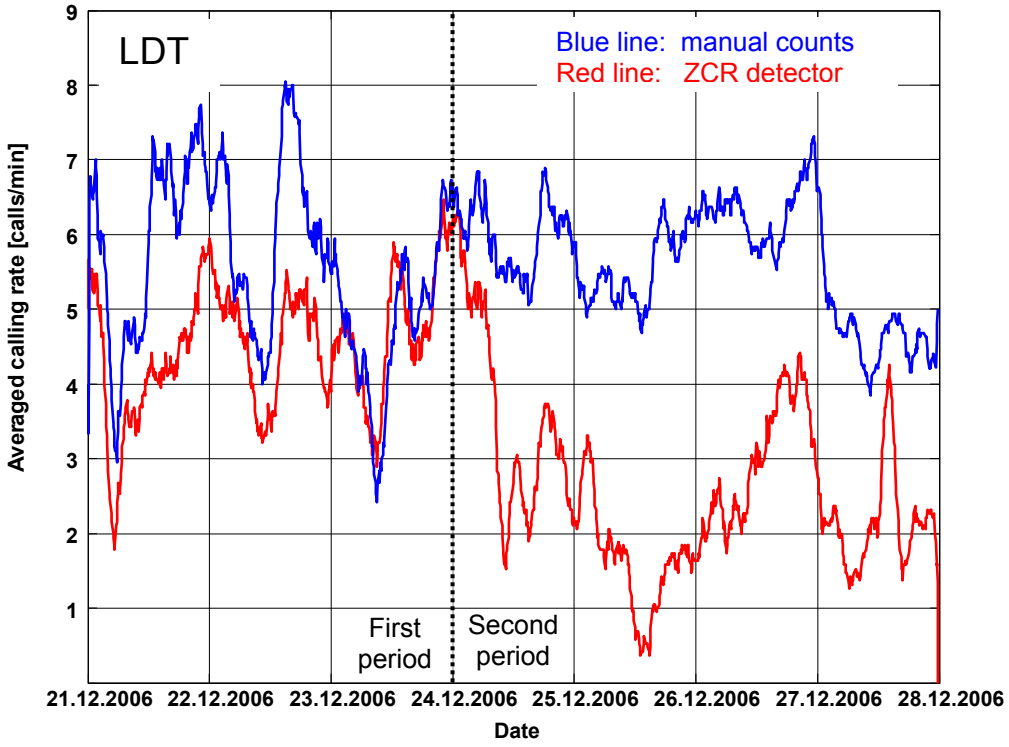


Figure 4.4: Averaged call rates of LDTs as obtained by the ZCR detector (red) and manual counting (blue). The time series shown gives 3 hour running mean values (i.e. the average over 24 files of one minute duration - see chapter 3.1.1 for details).

The detection performance of the ZCR detector for the entire period was 57%. The percentage of false detections was 5%. For the first period which was characterized by a mean SNR of the LDTs of 16.2 dB the ZCR detector performed better as 73% of all manual counts were detected (false positives: 8%).

Table 4.2: Detailed results of the ZCR detector on LDTs.

Method	***Total***		***Period 1 (21.12-23.12)***		***Period 2 (24.12.-27.12)***	
	Correct detections[4.6]	*False positives*[4.7]	*Correct detections*	*False positives*	*Correct detections*	*False positives*
Mean SNR (250 – 450 Hz)	---		16.2 dB		11.3 dB	
Manual counts	7538 100%	---	3334	---	4204	---
Zero-crossing rate detector (ZCR)	**4296 57.0%**	**238 5.2%**	**2448 73.4%**	**221 8.3%**	**1848 44.0%**	**39 2.0%**

The lower SNR during the second period (mean SNR 11.3 dB) caused a significant decrease in the detection performance. Only 44% of all calls were

[4.6] The number of correct detections is already corrected by the number of false positives.
[4.7] The percentage of false positives is derived by:
false positives /((correct detections + false positives)/100).

detected in this period. The percentage of false positives was calculated to be only 2%.
These results clearly indicate that the performance of the ZCR detector is highly dependent on a high SNR because a low SNR causes a waveform of high-frequency, thereby generating a high ZCR which masks the call specific, low ZCR.

4.2.2.2 Performance of a spectrogram correlation (SC) detector

To pursue the question, whether the ZCR detector performs any better than the commonly used spectrogram correlation detector, the latter - already implemented in the XBAT software - was tested on the data set. As already described in chapter 1.2.2.3, the spectrogram correlation (SC) detector calculates the correlation between a detection kernel (representing the target vocalization) and a sound snippet. The most critical parameter regarding a reliable detection of the LDTs (and also HDTs) is its varying duration because a mismatch in duration causes a lower correlation coefficient. Based on the results of the manual analysis (see chapter 3.2.2.1) a kernel with the following characteristic was chosen:

Table 4.3: Characteristics of the LDT kernel and statistical parameters for the respective call type derived from the manual analysis.

Parameter	*Value kernel*	*Mean value call*	*Sigma value call*
Duration [s]	3.8	3.8	1.0
Duration part 1 [s]	1.4	1.2	0.5
Duration part 2 [s]	1.8	1.9	0.5
Min. frequency part 1 [Hz]	310	304	22
Max. frequency part 1 [Hz]	390	399	38
Min. frequency part 2 [Hz]	300	298	23
Max. frequency part 2 [Hz]	381	396	35

The XBAT SC detector was fed with a kernel (parameters see Table 4.3) and the detection threshold (based on the correlation coefficient) was set to 0.3 as lower threshold levels caused significant higher false positive detections (> 10%). A call was detected by the SC detector whenever the correlation exceeded the set threshold level. However, the XBAT SC detector was used as a frontend user, thus more detailed information on the “internal” detection process (e.g. sliding step size of the kernel) are not available.

The results of the analysis are shown in Figure 4.5 and summarized in Table 4.4. The difference in detection performance between the two periods was negligible (period 1: 45%; period 2: 43%) and so the number of false positive detections (3% for the first period and 6% for the second period).

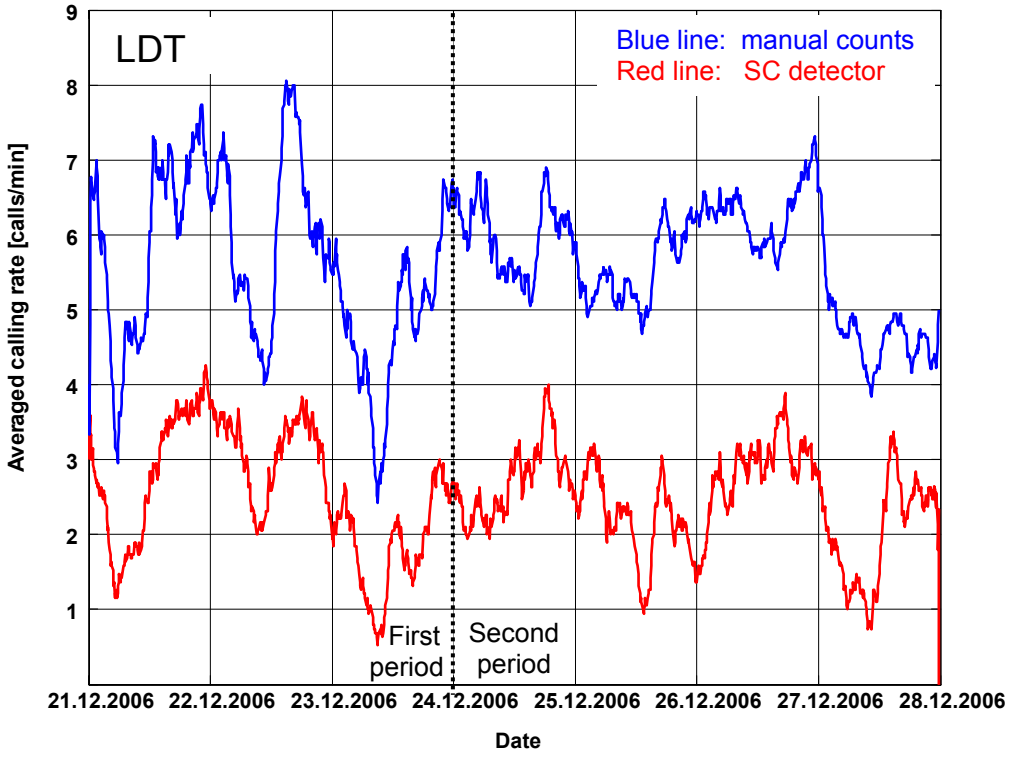


Figure 4.5: Averaged call rates of LDTs as obtained by the SC detector (red) and manual counting (blue). The time series shown gives 3 hour running mean values (i.e. the average over 24 files of one minute duration - see chapter 3.1.1 for details).

Table 4.4: Detailed results of the SC detector on LDTs.

Method	***Total***		***First period (21.12-23.12)***		***Second period (24.12-27.12)***	
	Correct detections	*False positives*	*Correct detections*	*False positives*	*Correct detections*	*False positives*
Mean SNR (250 – 450 Hz)	---		16.2 dB		11.3 dB	
Manual counts	7538 100%	---	3334	---	4204	---
Spectrogram correlation detector (SC)	**3297 43.7%**	**68 2.0%**	**1483 44.5%**	**48 3.1%**	**1814 43.1%**	**113 5.9%**

4.2.2.3 Comparison of the performance of the ZCR and the SC detector

The ZCR detector performed better for both periods. For the first period in particular, the differences between the ZCR detector and the SC detector are significant (ZCR: 73%; SC: 45%; Diff: +28%). For the second period, the results of both detectors are comparable (ZCR: 44%; SC: 43%; Diff: +1%). The results clarify that the ZCR detector depends strongly on the SNR of the calls. If the SNR is high (> 11 dB), the detector performed well. In contrast to the ZCR detector, the SC detector was less dependent on the SNR, as its performance is similar for both periods (see also Figure 4.6).

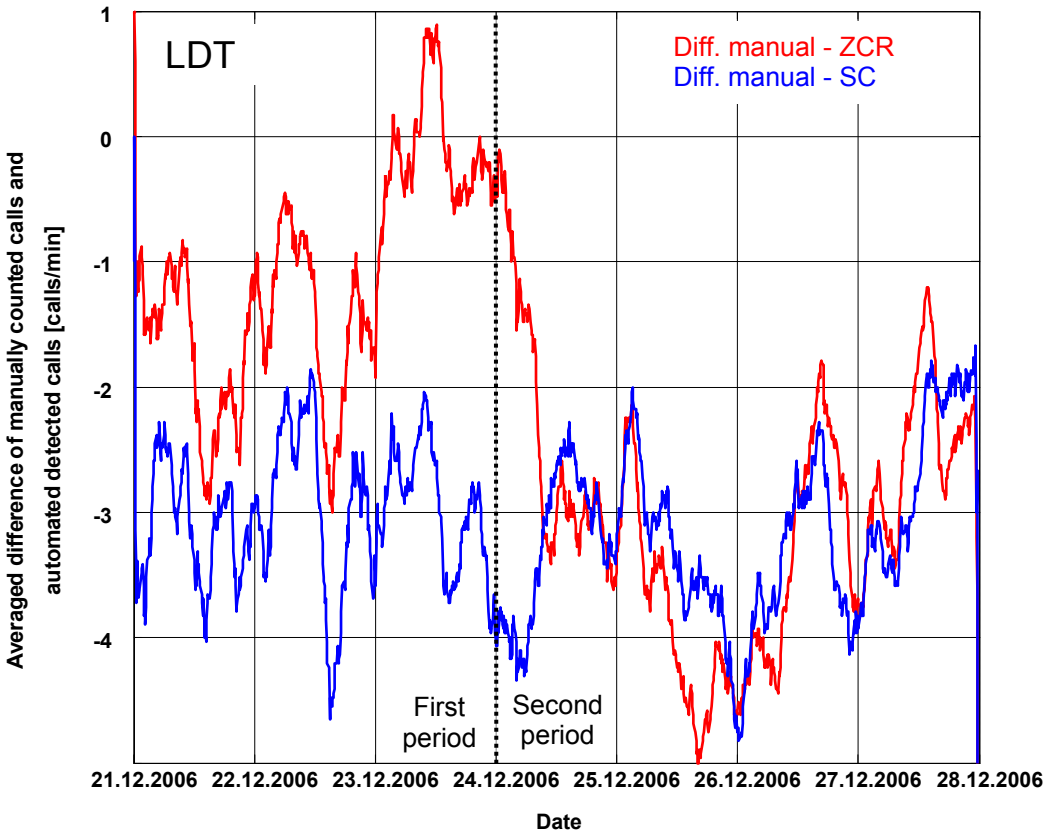


Figure 4.6: Averaged difference of manually counted calls and automated detected calls (ZCR - red; SC - blue). Numbers > 0 indicate false positives; numbers < 0 indicate missed calls. The time series shown gives 3 hour running mean values (i.e. the average over 24 files of one minute duration - see chapter 3.1.1 for details).

However, the lower number of correct detections indicates the SC detector is less adaptive to the high variability in call duration. The number of false positives derived for both detectors and periods was below 10%.

The constant detection rate (~44%; independently from the SNR) of the SC detector could be used to correct the number of detections by a factor 1/0.44. However, that ratio could change with time (see chapter 4.4.2.2). Thus one has to be careful if applying such a correction factor on extended data sets. In conclusion, in the vicinity of PALAOA, using the ZCR detector appears advantageous, as the SNRs of LDTs are most of the time above 11 dB.

4.3 tEST (the Envelope-Spectrogram Technique)

Human listeners can easily distinguish leopard seal vocalizations from those of other marine mammals because of their specific ”sound”. This specific sound is a result of the pulsed structure of the leopard seal vocalizations, a feature that lends itself to the development of automated pattern recognition algorithms. To exploit this feature for a detection algorithm for HDTs, the temporal evolution of the pulse repetition rate (PRR) within the call was analyzed and exploited in detail. So far, descriptions of HDTs and LDTs assumed a constant PRR throughout these calls (Rogers 2007; Rogers *et al.* 1995). By contrast, spectrograms of calls recorded by PALAOA reveal varying distances between side-bands over the duration of HDTs and LDTs, suggesting a variation of the PRR in the course of the call. To accurately analyze the temporal structure of the PRR, the respective sound snippet was first band pass filtered with the frequency range of the target signal (LDT: 250 – 450 Hz; HDT: 2.5 - 3.5 kHz). Then, the envelope of the absolute values of a band passed waveform was calculated by detecting all local maxima (peak values) in the waveform and interpolating (1-D) the detected points. The resulting waveform was then down-sampled to a sampling rate of 1000 Hz (HDT) or 100 Hz (LDT) and transformed into the frequency domain by means of a Fast Fourier Transformation (FFT-Parameters LDT: Hamming window 128 points; 50% overlap; FFT-Parameters HDT: Hamming window 256 points; 50% overlap;).

This algorithm, named tEST (the Envelope-Spectrogram Technique) hereinafter, provides the spectrogram of the envelope, i.e. the temporal evolution of the PRR.

4.3.1 Applying tEST to HDTs

Figure 4.7 shows the result of tEST applied to a HDT. The signal was processed as described in the previous paragraph.

For the selected sample, the pulse repetition rate (bottom panel) of the envelope varies between 52 Hz and 72 Hz (20 Hz bandwidth with a 2 Hz resolution). The first part of the call is characterized by descending rates. In the second part the pulse repetition rate is ascending.

The same 150 HDTs as used in characterization of this call type were analyzed with tEST. All vocalizations showed descending repetition rates in the first part of the call and ascending rates in the second part. The observed frequencies ranged between 45 Hz and 75 Hz.

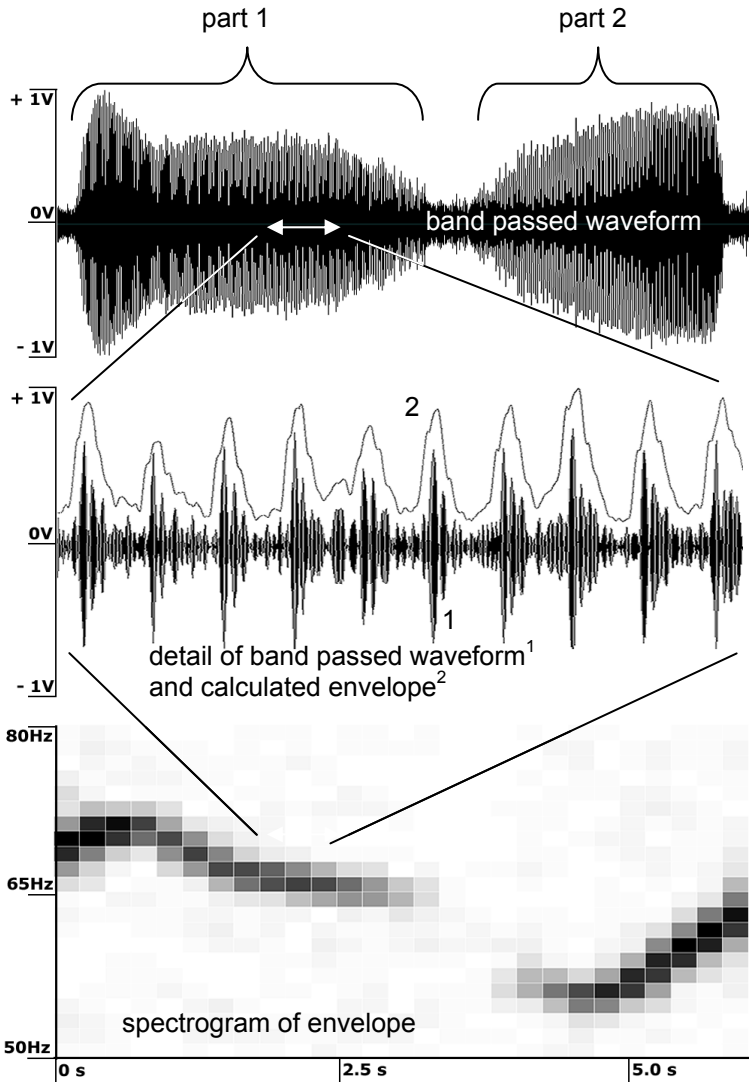


Figure 4.7: Results of tEST applied on a HDT. The top panel shows the band passed waveform, the mid panel a zoom into the waveform and the calculated envelope, while the bottom panel shows the spectrogram of envelope.

4.3.2 Applying tEST to LDTs

Figure 4.8 shows the result of tEST applied on a LDT. The PRR measured for this sample varies between 13 Hz and 20 Hz (7 Hz bandwidth with a 0.4 Hz resolution) with lowest PRRs in between the two segments. At the beginning and end of each segment, the pulse repetition rate is significantly increased.

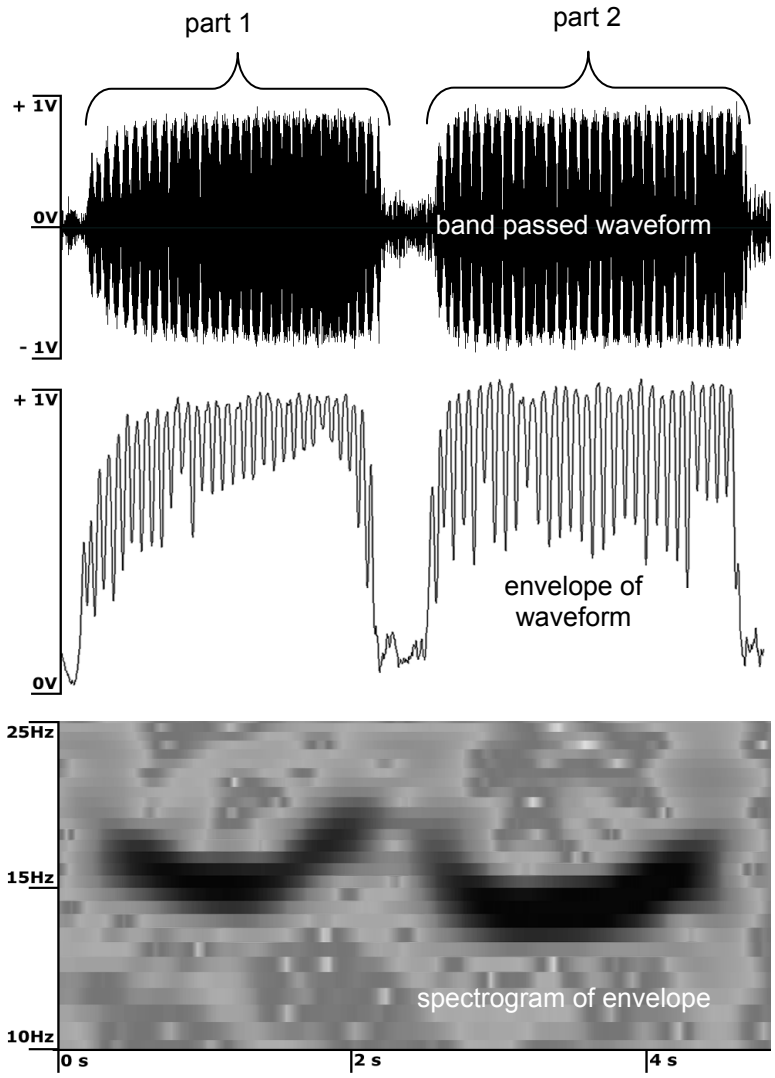


Figure 4.8: Results of tEST applied on a LDT. Top panel shows the band passed waveform, mid panel a zoom into waveform and calculated envelope and the bottom panel the spectrogram of envelope.

Analyzing the LDTs with tEST is more difficult than analyzing HDTs, because of the lower PRRs of the LDTs and the lesser variations of the PRR throughout the call. The call length (1.9 s - 9.5 s) limits a simultaneous high temporal and spectral resolution of the calculated spectrogram. The observed range of the PRR varied between 10 Hz and 25 Hz.

4.3.3 Results and discussion

The tEST algorithm as developed in this study is a useful tool to analyze the temporal evolution of pulse repetition rates in amplitude modulated animal calls. The analysis of the high and low double trill of the leopard seal revealed for the first time a systematic, temporal variation in the repetition rate of the pulses.

The PRR appears to be a unique feature, exclusively linked to the leopard seal vocalizations (at least in the vicinity of PALAOA), and is therefore suitable for their detection because the feature allows to distinguish leopard seal vocalizations from other marine mammal sounds. However, because of the difficulties in deriving the PRR feature in LDTs it might only be suitable for the detections of HDTs.

4.4 Automated detection of HDTs using a Hidden Markov Model

On the extracted PRR data, a Hidden Markov Model (HMM) was applied for the detection of HDTs.

4.4.1 Methods

Hidden Markov Models (HMM) are statistical models for the detection and classification of transient patterns and represent state-of-the-art tools in human speech recognition (Rabiner and Juang 1993). HMMs are particularly well adapted to call types of variable duration, as HMMs allow detection of temporally changing structures. For leopard seal vocalizations, this implies that the detection probability is independent of call duration as long as the envelope of the call follows a specific temporal evolution (see Figure 4.9). A short introduction and a description on building a Hidden Markov Model is given in the following paragraphs. For detailed information see Rabiner & Juang 1993 and Deller *et al.* 2000. All model parameters used in this study are given in the appendix.

The basic functions of the detection system were developed by Deller *et al.* (2000) and are available at: http://www.itu.dk/courses/TKG/E2005/exercises.html. However, for the usage of these functions in the HDT detection system a modification/adaptation by the author was necessary. The three main steps in setting up a HMM detector are described in detail.

(1) Feature extraction: From the available 150 samples, the best 100 samples (high SNR) were selected manually. For each of these sample files (of a variable duration of 2 - 9 s, depending on call duration), a set of feature vectors was calculated by segmenting the file into time-frames of 256 ms. For each time frame, a feature vector was calculated (see Figure 4.9), representing the respective energy distribution as a function of frequency.

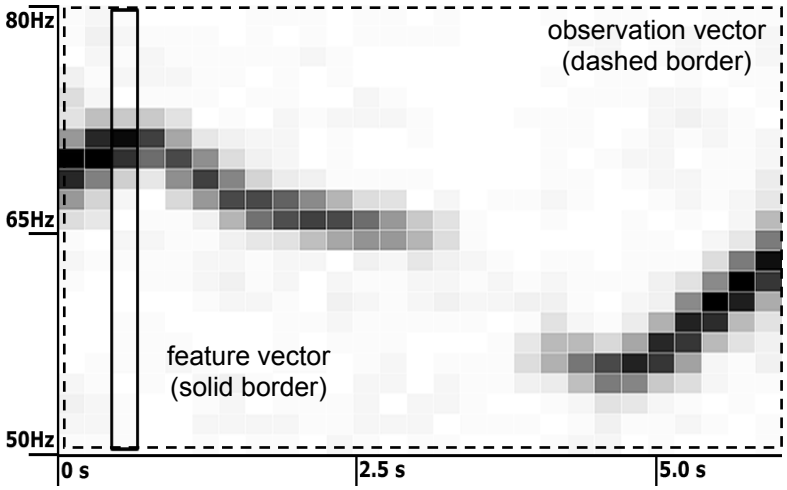


Figure 4.9: Spectrogram of pulse rep rate (PRR) for a sample file. Feature vector (enclosed by solid box) and observation vector (enclosed by dashed box) of a HDT.

All n feature vectors of one sample file (with n = duration of sample file / 256 ms) comprise the so-called observation vector (which actually is a matrix - see Figure 4.9). Hence, from the 100 sample files, 100 observation vectors (i.e. sets of spectral vectors) are extracted from the spectrograms of the envelopes.

(2) Information reduction: To condense the numerous ensuing feature vectors to a set of 'most significant' feature vectors, a "k-mean (squared Euclidean distance)" cluster algorithm (Deller *et al.* 2000) was applied to the training set. The association of a particular feature vector with its respective observation vector is thereby disregarded. This k-mean cluster algorithm creates the so-called codebook of 10 (number empirically chosen) codebook vectors representing the target vocalization's most significant (sub-) set of 10 feature vectors. Based on these 10 feature vectors, an observation sequence is calculated for each observation vector by determining the best matching (minimal distance) codebook vector for the sequence of feature vectors within an observation vector. Each of the resulting sequences consists of a series of integers, representing the succession of IDs of the best fitting codebook vectors. Thus, each set of the 100 spectral vectors is quantized to a one dimensional array of integers. The resulting set of 100 quantized vectors represents the quantized training set.

(3) Generate the HMM: Evaluating the model parameters that best describe the quantized training set. A Hidden Markov Model is a quintuple, comprising (a) the number of (hidden) states *S*; (b) the state transition matrix *A* (transition probabilities between the states); (c) the observation probability matrix *B*; (d) the state probability vector at time t=1, π (1); and (e) the number of observable outputs *Y* (number of codebook vectors), or in short:

$$HMM = \{S, \pi(1), A, B, Y\}$$

The number of states (*S*) was empirically assigned to 5. The number of 10 observable outputs (*Y*) is given by the size of the codebook. The state transition matrix (A), the observation probability matrix (*B*) and the state probability vector at time=1 (π (1)) were determined by applying a forward/backward algorithm (Deller *et al.* 2000) on the quantized data set.

In a first step, the model parameters *A*, *B* and π (1) are initialized (see below) and the algorithm calculates the match between the model and the training set. In a second step, the algorithm starts to modify the model parameters. The algorithm guarantees that every iteration has a matching likelihood that is greater or equal to the previous one. Once the matching likelihood converges, the training is completed.

Critical to this process is the initial guess of the model parameters' values. In the case of the described target signals, meaningful initial guesses of the parameter values were unknown. If the initial guess is too far away from the optimal parameter values, then the algorithm will only find a local maximum which may not necessarily coincide with the global maximum. For this reason the choices for the initial parameter values were randomized and the resulting HMMs used to repeatedly analyze one sample file with a known number of target signals. The best fitting model parameters' values (giving the highest detection probability for the target signals) were then chosen for further processing.

To detect HDTs with the optimized HMM (5 states), a 6 s window is stepped across the data stream in increments of 1.0 s (~ 83% overlap). The respective window content is first band pass filtered (2500 Hz - 3500 Hz) and then used to calculate the waveform envelope, which is used to derive the observation sequence as described above. In a final step the probability of the observation sequence of each window under the assumption of the model $P_{(window|model)}$ is calculated. To detect HDTs of low SNR a wavelet based denoising technique (Kovesi 2000; Kovesi 1999) to manipulate the spectrogram of the envelope was tested. Also, an anisotropic diffusion was performed on the spectrogram to enhance the contrast at sharp intensity gradients (Kovesi 2000). Both techniques were applied to the data using existing Matlab™ functions (developed by Peter Kovesi[4.8]). The use of the denoising and the anisotropic diffusion algorithm increased the detection performance significantly. Thus, the algorithms were directly integrated into the system.

Figure 4.10 exemplifies the output of the HMM detection algorithm for different signals: two HDTs of different duration (a), an artificial signal (b) and a Weddell seal call (c). The detector output is shown in the lower part of Figure 6. The detection threshold is indicated therein by a horizontal line, which is automatically set by the algorithm and dependent on the overall SNR. If the detector output reaches the detection threshold, a call is "detected". In this sample the HDTs are clearly detected by the system. By contrast, the Weddell

[4.8] http//www.csse.uwa.edu.au/~pk/Research/MatlabFns/index.html

seal call or an artificial sinusoidal unmodulated signal in the same frequency band as the HDTs remain undetected.

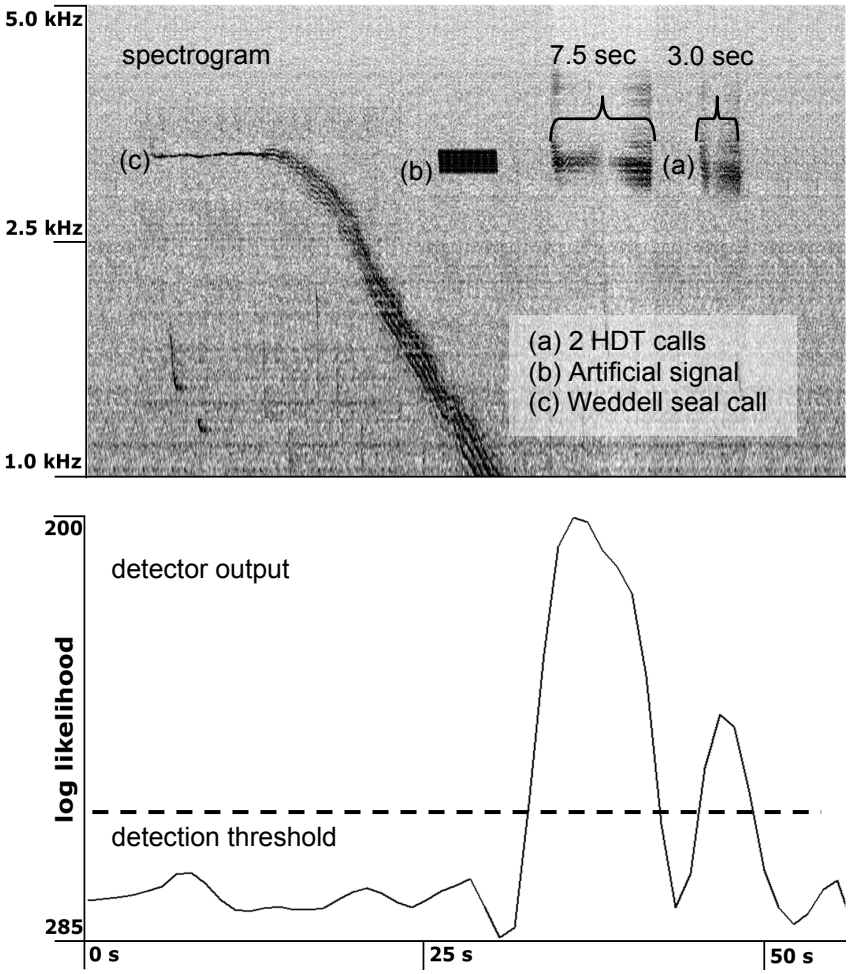


Figure 4.10: Top panel shows the spectrogram of the test signals: (a) two HDTs of different duration, (b) an artificial signal and (c) a Weddell seal call. The detector output for each test signal is displayed in the bottom panel.

With this method, analyzing a sound file of 2 minutes duration (48 kHz - 16 bit) takes about 40 seconds (Desktop PC with single Intel Pentium IV 3.4 kHz processor and 2 GB RAM). Thus the HMM based detection system is suitable for real-time applications.

4.4.2 Results and discussion

Analyzing the SNR (in the frequency band between 2500 Hz and 3500 Hz) of 12 HDTs per day (every two hours one HDT; because of the low frequency of occurrence the hourly minute had to be varied) showed that the period can be divided into three shorter periods of characteristic SNR (see Table 4.5). The first period (21 December – 22 December) is characterized by a high SNR of the HDTs (mean SNR 17.4 dB) while the mean SNR measured for the second (23 December – 25 December) and third period (26 December – 27 December) are 4.0 dB and 10.7 dB respectively.

The calculated background noise levels (frequency band 2500 Hz – 3500 Hz) were for period 1: -77 dBFs (sigma: 3 dB), period 2: -67 dBFs (sigma: 5 dB) and period 3: -73 dBFs (sigma: 3 dB). As already mentioned in the LDT section the wind speed significantly increased during 24 December 2006 from 2 m s^{-1} to 12 m s^{-1}. This increase in wind speed probably caused a higher background noise level in the target frequency band of the HDT.

Table 4.5: SNR of selected HDTs.

	No. of samples	*Min. SNR [dB]*	*Max. SNR [dB]*	*Mean SNR [dB]*	*Median SNR [dB]*	*Sigma SNR [dB]*
Period 1 (21.12 – 22.12)	24	11.3	27.2	17.4	15.1	5.2
Period 2 (23.12 -25.12)	36	0.3	9.3	4.0	3.2	3.0
Period 3 (26.12 -27.12)	24	5.3	18.7	10.7	9.9	4.3

Furthermore the variation in the SNR could be also be explained by the movements of the seals away from the hydrophones, as the SNR decreases with distance. The absorption of high frequency sounds (like the HDT) is higher compared to low frequency sounds (like the LDT). The absorption of a 3000 Hz (300 Hz) signal in sea water over a distance of 1 km is (assuming a water temperature of -2°C, a depth of the sound source of 50 m, a salinity of 34 ppt and a pH value of 8) 0.195 dB km^{-1} (0.012 dB km^{-1}). The absorption was calculated using the formula described in Ainslie and McColm (1998). Thus, the SNR of the HDTs is more influenced by the movements of the animals compared to the LDTs.

4.4.2.1 Performance of the HMM detector

The performance of the detector was verified - similar to the ZCR detector - using the manually counted test data set described in paragraph 3.1. The results are given in Figure 4.11 and Table 4.6. Overall, 56% of all manually counted calls were detected by the HMM detector. The number of false positives was 7%. The best result was achieved for the first period which is characterized by a mean SNR of the HDTs of 17.4 dB. For this period the HMM detector detected nearly 85% of the manual counts. For the second and third period the detection performance was 28% respectively 69%. The number of false positives was acceptable for both periods (first period: 4%; second period: 9%).

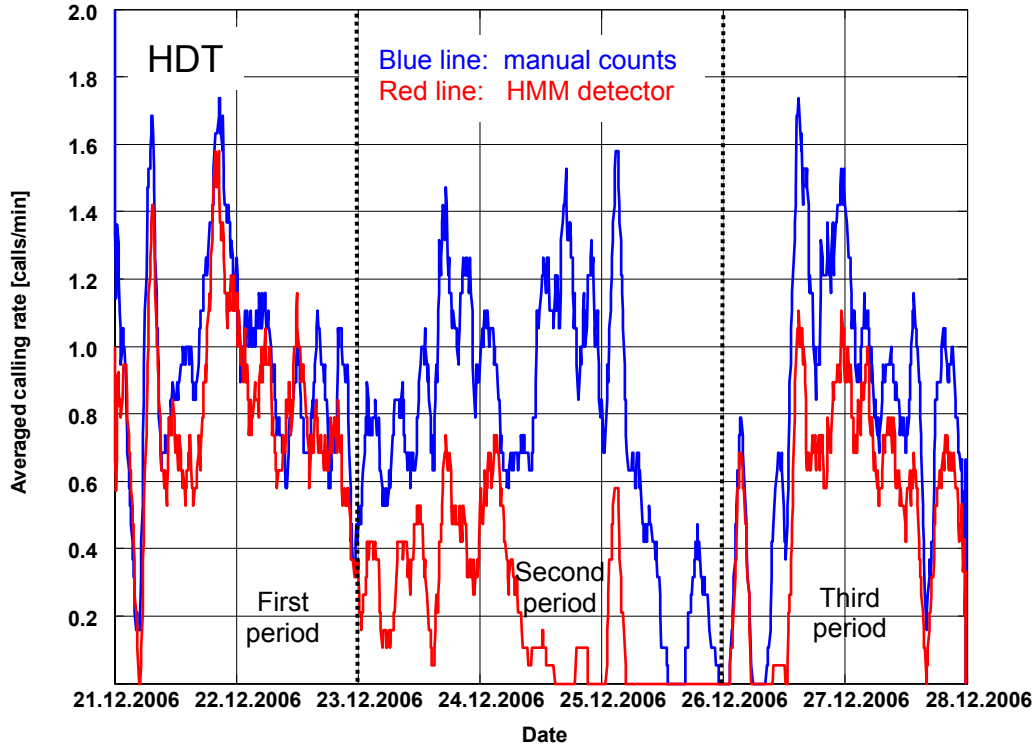


Figure 4.11: Averaged call rates of HDTs as obtained by the HMM detector (red) and manual counting (blue). The time series shown gives 3 hour running mean values (i.e. the average over 24 files of one minute duration - see chapter 3.1.1 for details).

Table 4.6: Detailed results of the HMM detector on HDTs.

Method	***Total***		***First period (21.12-22.12)***		***Second period (23.12-25.12)***		***Third period (26.1-27.12)***	
	Correct detections	*False positives*	*Correct detections*	*False positives*	*Correct detections*	*False positives*	*Correct detections*	*False positives*
Mean SNR (2.5-3.6 kHz)	---		17.4 dB		4.0 dB		10.7 dB	
Manual counts	1118 100%	---	366 100%	---	489 100%	---	263 100%	---
HMM detector	**628 56.2%**	**46 6.8%**	**310 84.7%**	**26 7.7%**	**137 28.0%**	**5 3.5%**	**181 68.8%**	**17 8.6%**

These results show that the detection performance is dependent to the SNR of the calls. For calls with a SNR > 10 dB the detection performance was acceptable. When the mean SNR of the calls dropped below 10 dB (second period) the detection probability was reduced to 28%. This decrease in the detection rate is caused by the irregular waveform of the HDTs which masks the PRR of the calls (pulses cannot be clearly identified in the waveform) and prevents the detector to pick up such calls.

4.4.2.2 Performance of a spectrogram correlation (SC) detector

For comparison, a detector based on spectrogram correlation was applied to the data set. Based on the results of the manual analysis (see chapter 3.2.2.1) a kernel with the following characteristic was chosen:

Table 4.7: Characteristics of the HDT kernel and statistical parameters for the respective call type derived from the manual analysis.

Parameter	*Value kernel*	*Mean value call*	*Sigma value call*
Duration [s]	3.9	3.9	1.2
Duration part 1 [s]	1.8	1.7	0.7
Duration part 2 [s]	1.5	1.6	0.5
Min. frequency part 1 [Hz]	2640	2673	62
Max. frequency part 1 [Hz]	4200	4202	97
Min. frequency part 2 [Hz]	2630	2610	51
Max. frequency part 2 [Hz]	4055	4065	102

The detection threshold was empirically set to 0.3. Lower thresholds caused significantly higher false positive detections (> 15%). The results are shown in Figure 4.12 and summarized in Table 4.8.

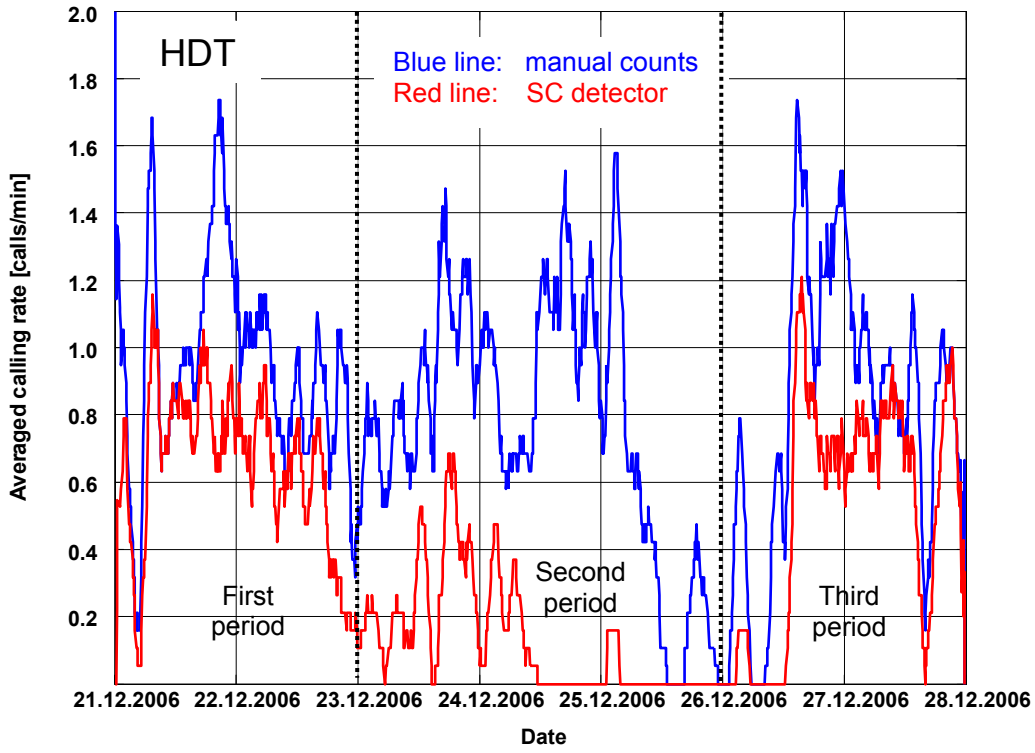


Figure 4.12: Averaged call rates of HDTs as obtained by the SC detector (red) and manual counting (blue). The time series shown gives 3 hour running mean values (i.e. the average over 24 files of one minute duration - see chapter 3.1.1 for details).

In total, 41% of all manually counted calls were detected by the SC detector. The total number of false positives was 10%. The highest detection rate of 70% was achieved for the first period with mean SNRs of 17.4 dB. In the second and

third period 17% and 63% of the manual counts were detected respectively. The false positive rate for the first and third period - for which a sufficient number of calls were detected - was above 10%. In the second period the number of false positives was low (2%) but during this period only 17% of all calls were detected.

Table 4.8: Detailed results of SC detector on HDTs.

Method	***Total***		***First period (21.12-22.12)***		***Second period (23.12-25.12)***		***Third period (26.12-27.12)***	
	Correct detections	*False positives*	*Correct detections*	*False positives*	*Correct detections*	*False positives*	*Correct detections*	*False positives*
Mean SNR (2.5-3.6 kHz)	---		17.4 dB		4.0 dB		10.7 dB	
Manual counts	1118 100%	---	366 100%	---	489 100%	---	263 100%	---
Spectrogram correlation detector (SC)	**462 41.3%**	**53 10.3%**	**218 69.6%**	**27 11.0%**	**79 16.6%**	**2 2.5%**	**165 62.7%**	**24 12.7%**

The SC detector shows comparable results for the first and third period although the SNR of those periods differs. This indicates the robustness of the SC detector against varying SNRs. On the other hand, the detection rate dropped down to 17% during the second period. Thus, the SC detector did not perform well for calls with SNR < 10 dB.

4.4.2.3 Comparison of the performance of the HMM and the SC detector

The HMM detector performed better throughout all periods as compared to the SC detector. The largest difference in the number of detected calls (+15%) was observed for the first period (see Figure 4.13). For the third period, with mean SNRs of 10.7 dB, the performance of both detectors was similar (HMM: 69%; SC: 63%). For the second period, which was characterized by low SNRs (< 10 dB) the HMM detector performed 11% better than the SC detector. Furthermore, the number of false positives of the HMM detector was around 4% lower. As a conclusion, the HMM detector (HDT) worked reliably under various SNRs and produced a low number of false positive detections. The detection performance varied depending on the SNR of the calls but detection rates of at least 60% can be expected when applied to the entire PALAOA data set.

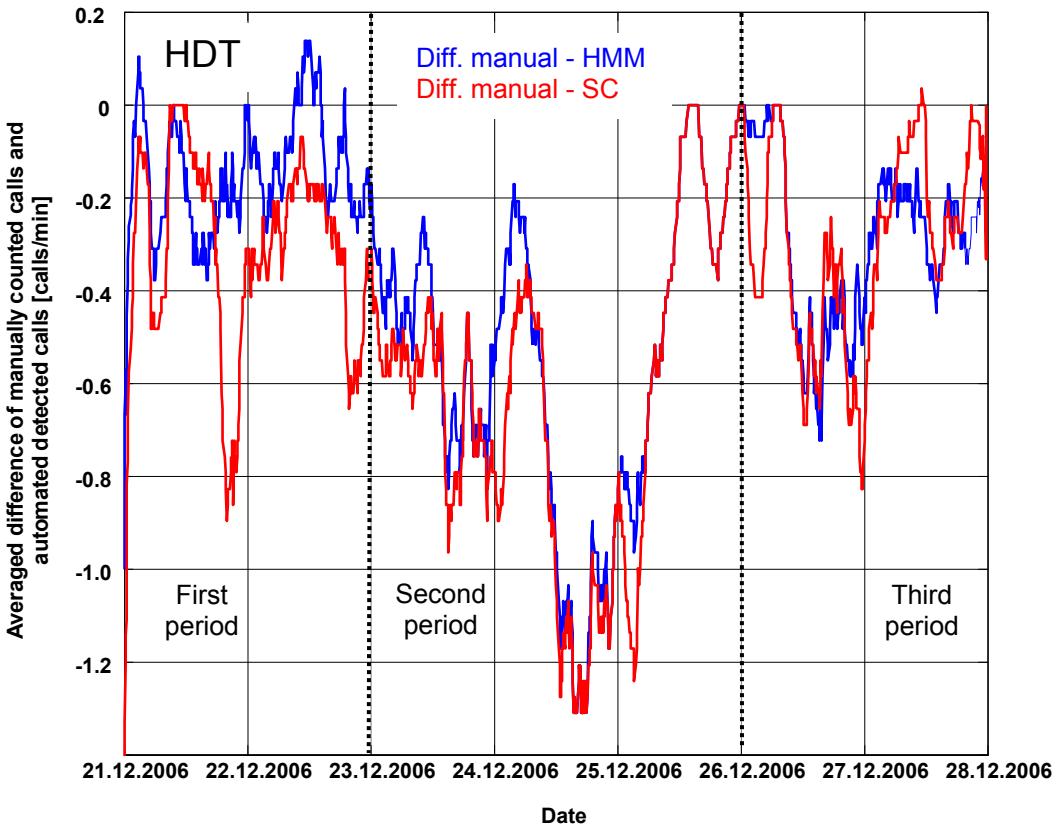


Figure 4.13: Averaged difference of manually counted calls and automated detected calls (HMM - blue; SC - red). Numbers > 0 indicate false positives; numbers < 0 indicate missed calls. The time series shown gives 3 hour running mean values - i.e. the average over 24 files of one minute duration - see chapter 3.1.1 for details). At 25 December the blue line is concealed by the red one.

4.5 Applying the detectors to large data sets

4.5.1 ZCR detector

4.5.1.1 Methods

The ZCR detector was implemented on the XBAT software system. While XBAT is quite advanced, it is not yet capable of handling thousands of files at once. For this reason a sub data set of the PALAOA data set was created, spanning the period from September 2006 to January 2007. This sub data set was further divided into data sets of 1 week duration which can be handled by XBAT. All data sets, consisting of about 10000 wav files of 1 minute duration, were resampled to 11025 Hz @ 16 bit. All data files of the respective week were imported to XBAT as a data stream and then handled by the software as a single sound file. Loading the data and scanning through the entire sub data set of around 160000 files took around one week.

4.5.1.2 Results and discussion

In total 70112 LDTs were detected. The timeline is shown in Figure 4.14.

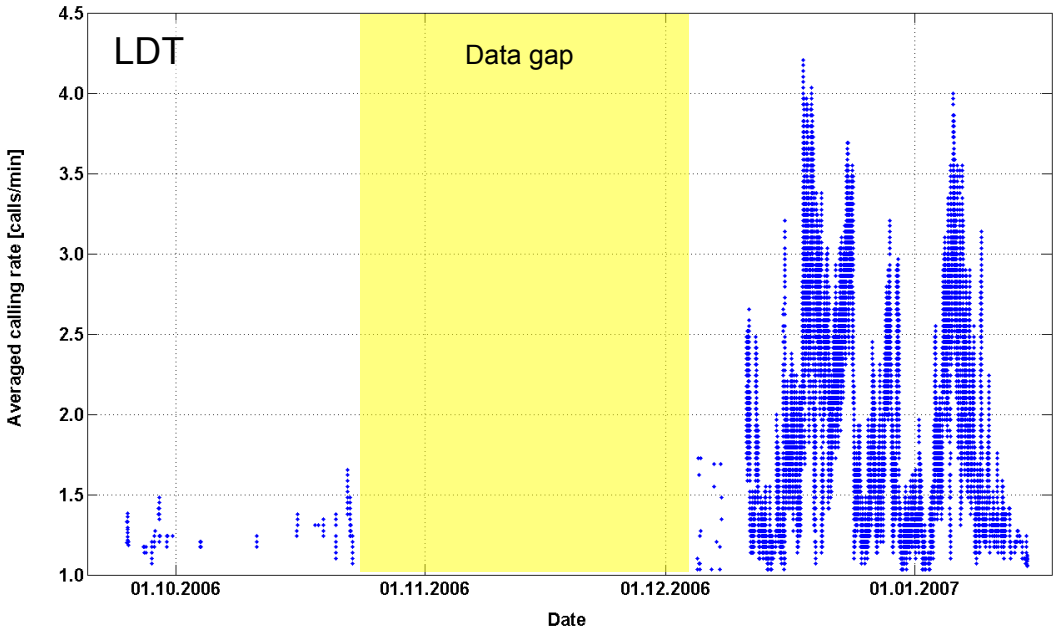


Figure 4.14: Averaged call rates of LDTs as obtained by the ZCR detector applied on the PALAOA sub data set (01.09.2006 – 31.01.2007). The time series shown gives running mean values over 60 detections.

During November 2006 no data is available due to a malfunction of PALAOA 06 (see chapter 2.4). The first LDT was detected on 24 September 2006 and the last one on 15 January 2007. The calling rate peaks between mid-December and early January.

4.5.2 HMM detector

4.5.2.1 Methods

The HMM detector was implemented into the PALAOAdb system as described in chapter 2.6. The PALAOAdB workers were installed on eight PCs in parallel and started when the PCs were not needed for work (e.g. during night time or the weekend). Screening the whole data set of around half a million files took around a month.

4.5.2.2 Results and discussion

A total of 16080 HDTs were detected. The timeline of detections is presented in Figure 4.15.

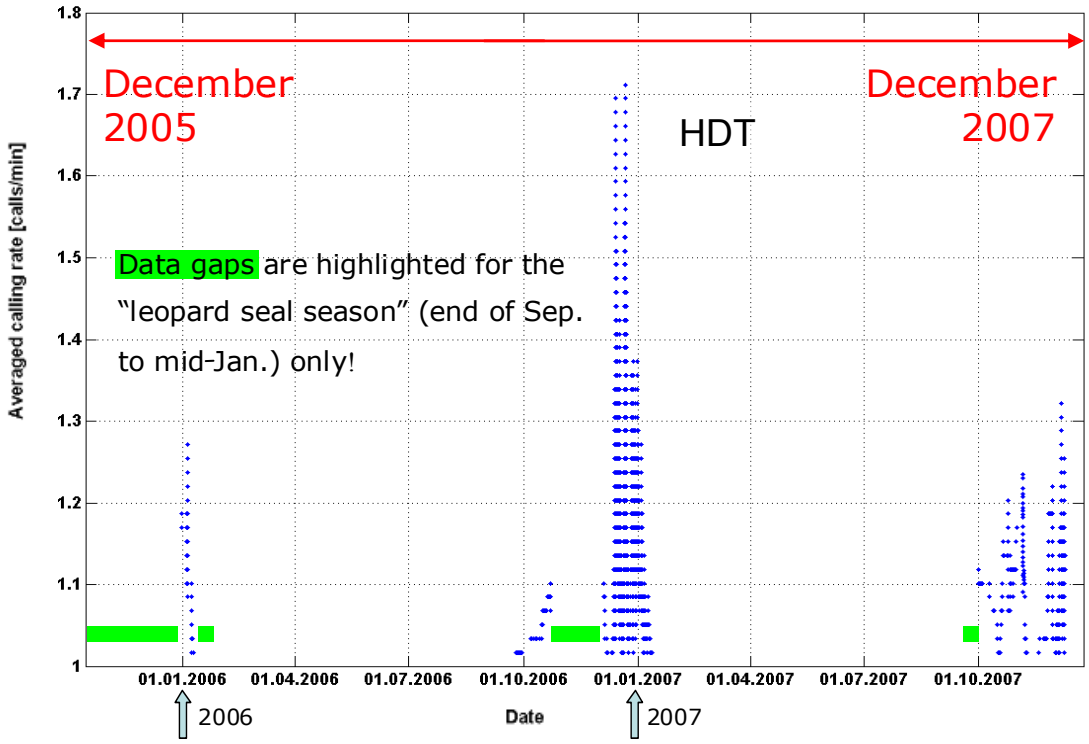


Figure 4.15: Averaged call rates of HDTs as obtained by the HMM detector applied on the PALAOA data set (01.12.2005 – 09.12.2007). The time series shown gives running mean values over 60 detections.

In interpreting these results it has to be considered that the PALAOA data set is discontinuous. In chapter 5.2 the HDT timeline of Figure 4.15 will be discussed in detail (under consideration of the available data and the acoustic behavior of the leopard seal). In Figure 4.16 the HDT detections for the season 2006/2007 are depicted. Equivalent to the LDTs highest HDT calling rates occur from mid-December to January.

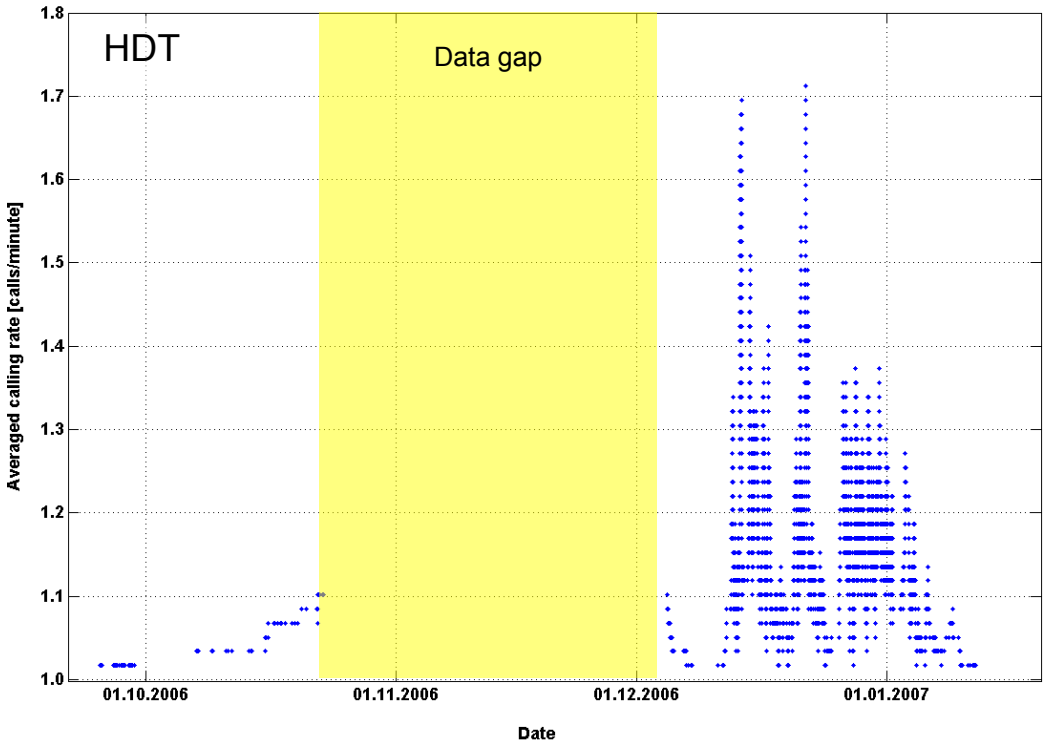


Figure 4.16: Averaged call rates of HDTs as obtained by the HMM detector applied on a PALAOA sub data set (01.09.2006 – 31.01.2007). The time series shown gives running mean values over 60 detections.

The occurrence of the first and last HDT in each ”leopard season” is given in Table 4.9. The result for the season 2006/2007 agrees with the results of the ZCR detector (first LDT occurred 24 September 2006).

Table 4.9: Date of first and last HDT in the observed seasons.

Season	***First HDT***	***Last HDT***
Season 2005/2006	---	10 January 2006
Season 2006/2007	25 September 2006	12 January 2007
Season 2007/2008	30 September 2007	---

To summarize, leopard seal vocalizations can be expected in the vicinity of PALAOA between end of September and mid of January.

5. Time series analysis to study temporal patterns of the vocal behavior of leopard seals

5.1 Seasonal patterns

5.1.1 Methods

Seasonal patterns of the acoustic behavior of leopard seals were mainly derived from the results of the automated detection runs. While the HMM detector was applied to the entire PALAOA data set (January 2006 – December 2007) the ZCR detector could only be used to screen a smaller subset of data (September 2006 – January 2007). To compensate the lack of knowledge on the presence/absence of LDTs outside the automatically screened period, manual checks of the recordings of minute 00:00 and 00:01 each day were added.

5.1.2 Results and discussion

5.1.2.1 Presence of leopard seal vocalizations at PALAOA

Presence and absence of leopard seal vocalizations (LDT and HDT) are presented in Figure 5.1. This graph shows in the first row ('Data') the availability (color coded) of acoustic data for the years 2006 and 2007. The second and third row ('LDT', 'HDT') indicate (a) which periods were analyzed either manually (M) or automatically (A) with the ZCR detector (for LDTs) or the HMM detector (for HDTs). If a single HDT or LDT was detected for a given week, the respective box was colored to indicate the presence of a leopard seal vocalization. Figure 5.1 shows that leopard seal vocalizations were detected in the vicinity of PALAOA between end of September and mid of January.

Previous studies found that free-living leopard seals were vocal only between November and mid-January (Thomas 1982; Stirling & Siniff 1979). In a study of captive animals, Rogers *et al.* (1996) investigated the vocal behavior of two individuals – a sexually immature male and a sexually mature female) over a three year period between May 1990 and December 1992. Underwater sound recordings were regularly made for 10 hours on a single day each month. The resulting data set was sub-sampled by extracting 20 min samples for every recorded hour. For each sub-set, the number of vocalizations was counted manually.

A major result of Rogers *et al.* (1996) was that only the female seal was vocalizing and only in December and early January when she had elevated oestradiol levels - which indicates that she was in oestrus. Thus the difference in the onset of vocal activity between the PALAOA data and Rogers *et al.*'s (1996) study is most likely attributed to the presence of sexually mature males in the vicinity of PALAOA. As leopard seals are solitary living pack ice seals, long distance acoustic communication between males and females is thought to play an important role during the mating season (Rogers 2003). It has been suggested that these vocalizations may play a role in mate selection and male-male competition (Rogers 2003).

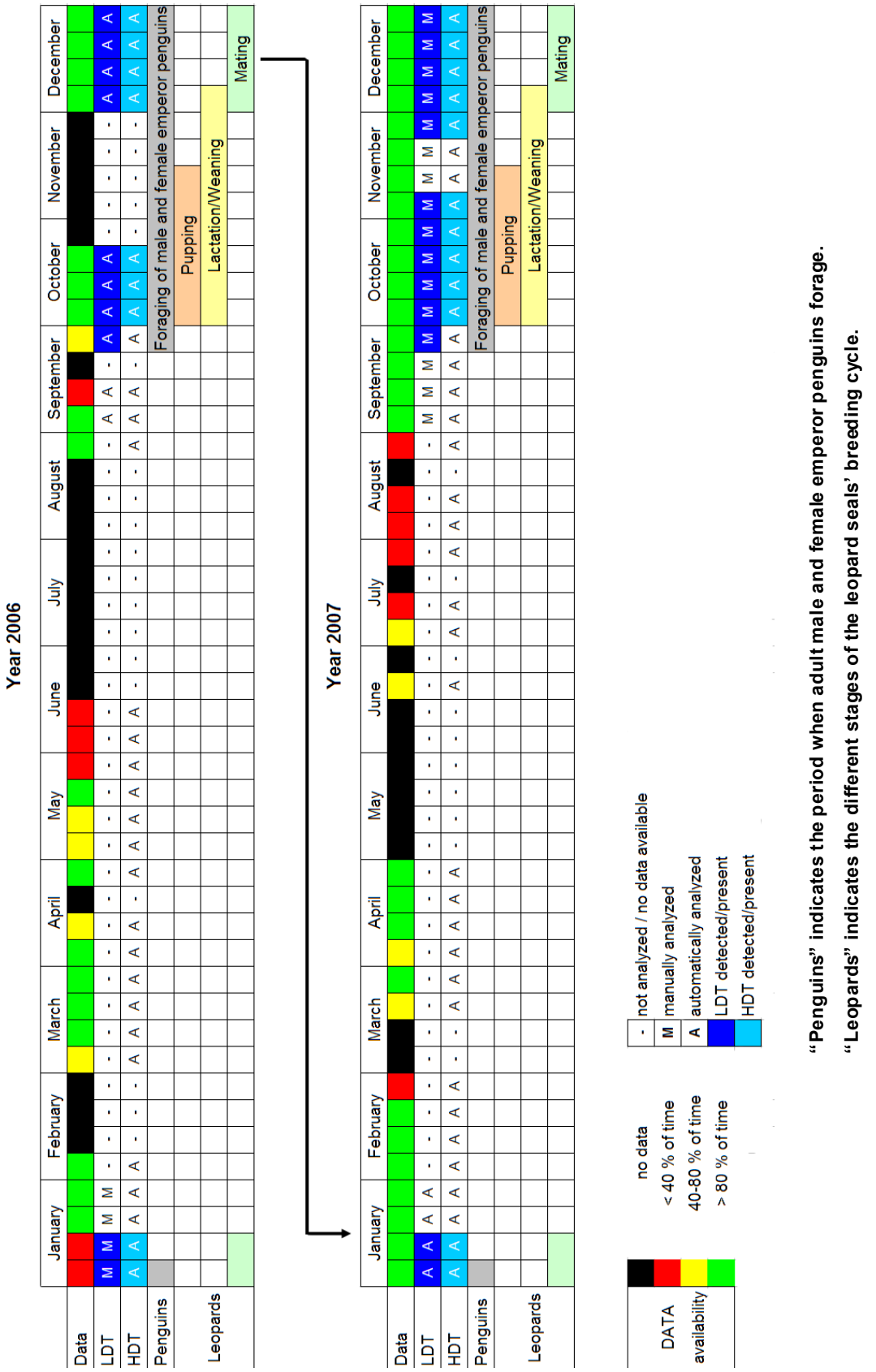


Figure 5.1: Seasonal patterns of the vocal behavior of the leopard seal (2006 and 2007).

A possible reason for the early presence of leopard seals in the vicinity of PALAOA could be the emperor penguin colony of around 4500 breeding pairs which is located in Atka Bay, around 10 km S-SE of PALAOA. The occurrence of leopard seal vocalizations from late September to early January coincides with the period when both adult female and male emperor penguins forage[5.1] (see Figure 5.2; parents foraging).

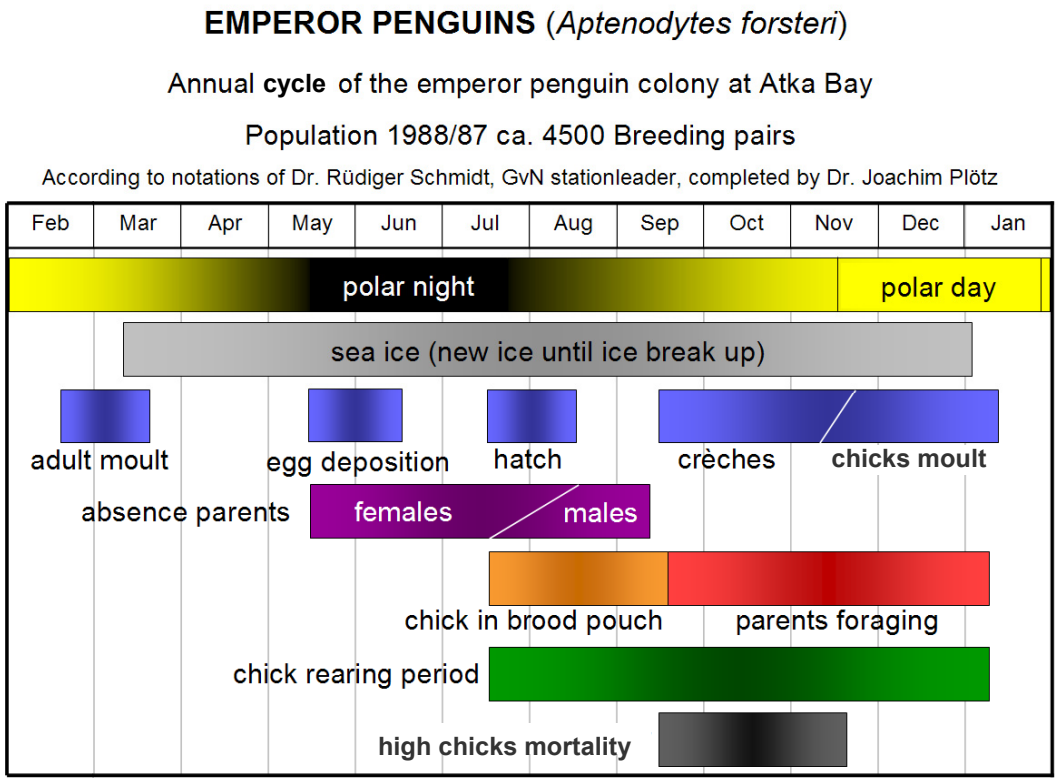


Figure 5.2: Life cycle of the emperor penguin (*Aptenodytes forsteri*).

Previous studies (on other leopard seal populations) reported that leopard seals stay in the vicinity of penguin colonies during January and February when the inexperienced chicks enter the water for the first time (Lowry *et al.* 1988; Siniff 1991). By contrast, the results of the acoustic observations at Atka Bay suggest that the leopard seals migrate towards Atka Bay as early as September, possibly to feed on adult foraging emperor penguins. The period of observed vocalizations is, at least, surprisingly well matched with the period when both parents start to undertake regular foraging trips and move between the colony and their oceanic feeding areas to feed their chick. The continuous presence of a substantial number of birds in the ocean is likely to provide an attractive feeding spot for leopard seals in the vicinity of PALAOA.

Between September and early December leopard seal vocalizations are not continuously present in the PALAOA recordings. Most likely, this reflects the movements of the animals, as leopard seals are known to migrate during austral spring between September and November (Rogers *et al.* 2005). In

[5.1] Forage means searching for food.

contrast (adult) leopard seals remain stationary between December and January in a certain area for mating (Rogers *et al.* 2005).

5.1.2.2 Seasonal calling rates at PALAOA

The calling rate of leopard seals in the vicinity of PALAOA peaks between mid-December and early January (see Figure 4.12 and 4.14). This peak in the calling rate is likely to indicate the peak of mating activity. Male leopard seals are known to increase their calling rate during this period, presumably to attract females and compete with other males (Rogers 2003). This acoustic behavior was also observed in other seal species (Van Parijs 2003). Rogers *et al.* (1996) also found high individual vocalization rates for a captive female leopard seal while sexually receptive. Females come into estrus when they re-enter the water after they have weaned their pup (see Figure 5.1). They are widely dispersed on the pack ice at this time, and, therefore, they might need to advertise where they are, as do the males (Rogers 2003).

The calling rates between the end of September and early December are lower compared to the mating season (mid-December to early January). Pupping is believed to occur from October to mid-November (Rogers & Cato, 2002), and leopard seal pups have been observed to be accompanied by adults until the end of December (Southwell *et al.* 2003). During the lactation period, which is assumed to last for about four weeks, female leopard seals attend their pup on the ice. Therefore it can be hypothesized that mainly the vocalizations of male leopard seals are heard during this period. However, further acoustic studies are necessary to support this hypothesis.

5.2 Diurnal patterns

5.2.1 Methods

To reveal possible diurnal patterns in the vocal behavior of leopard seals, the manually analyzed PALAOA and DIPS (Drescher Inlet Pilot Study) data sets were examined in detail. These data sets contain counts of all leopard seal call types and provide a high degree of accuracy (chapter 3.1.1). Both data sets were recorded during the last third of December (between 21 December and 25 December 2003 for DIPS and 21 December and 27 December 2006 for PALAOA; see 3.1.1 for details). The analysis included more than 20000 vocalizations (around 10000 vocalizations per data set). To examine potential diurnal vocalization patterns, the number of vocalizations for both data sets for each hour of a day[5.2] (00-23) was extracted and the actual calling rates were calculated.

[5.2] The hours of day refer to local time. The difference between coordinated universal time (UTC) and local time is for PALAOA -32 minutes and for DIPS -76 minutes.

5.2.2 Results and Discussion

Because the time series of vocalization rates appears to be anomalous on 24 December 2006 (see Figure 5.4), this day was excluded from further analysis of diurnal patterns to allow a comparison between DIPS and PALAOA data. The result of the analysis is shown in Figure 5.3. Leopard seal vocal activity was found to exhibit a distinct diurnal pattern for both data sets. Both diurnal cycles show a similar pattern with high calling rates during night time[5.3] and low calling rates during forenoon (see Table 5.1). The differences in the calling rates between both locations could either be caused by a different number of vocalizing animals or differences of individual calling rates.

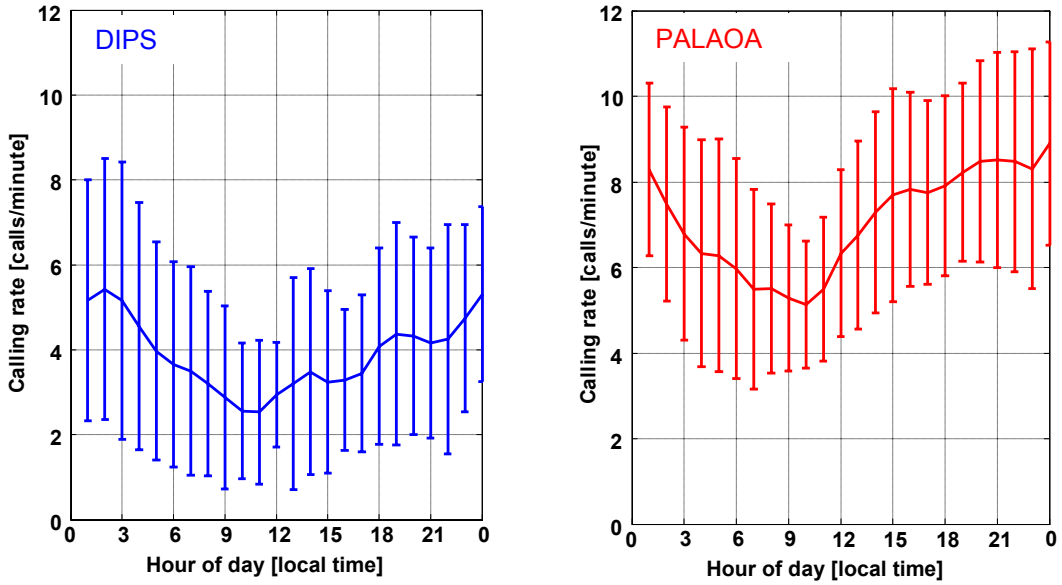


Figure 5.3: Diurnal calling pattern of leopard seals derived from the DIPS (blue, left panel) and PALAOA (red, right panel) data set. The curves are representing mean calling rates per minute (and standard deviation) in the respective hour.

Similar diurnal patterns have also been reported by other studies (Thomas & DeMaster 1982). These studies suggested that this pattern most likely reflects the haul out[5.4] behavior of the leopard seals. The number of animals resting on the ice peaks at around noon when solar radiation and outside temperature are highest (Hofmann *et al.* 1977; Gilbert & Erickson 1977; Thomas & DeMaster 1982; Rogers & Bryden 1997). Accordingly, lowest calling rates coincide with periods when most animals are on the ice (at Atka Bay and Drescher Inlet in late forenoon at around 10 am and 11 am).

[5.3] It has to be considered that also during night time a reasonable amount of sunlight is available in high latitudes during austral summer.
[5.4] Hauling out means resting on the ice.

Table 5.1: Maximum and Minimum mean calling rates [calls/min] at Drescher Inlet (DIPS) and Atka Bay (PALAOA).

	Max. mean calling rate		*Min. mean calling rate*		*Decrease*
	[hour - local time]	*[calls/min]*	*[hour - local time]*	*[calls/min]*	*[%]*
DIPS	02	5.6	11	2.6	-54
PALAOA	00	8.9	10	5.1	-43

As mentioned above, an irregularity in the daily variation of the call rate at PALAOA was observed (see Figure 5.4). On 22 December and 23 December 2006 the diurnal pattern is quite pronounced but on 24 December, the calling rate remains high throughout the day (between 7.8 and 9.0 calls per minute).

It is known from earlier studies (Gilbert & Erickson 1977; Rogers & Bryden 1997) that changing environmental conditions influence diurnal haul out patterns. Rogers & Bryden (1997) found a significant negative correlation between the windchill temperature and the number of leopard seals hauled out on the ice, while Gilbert & Erickson (1977) reported a significant negative correlation between the number of Antarctic pack ice seals hauled out and the wind speed.

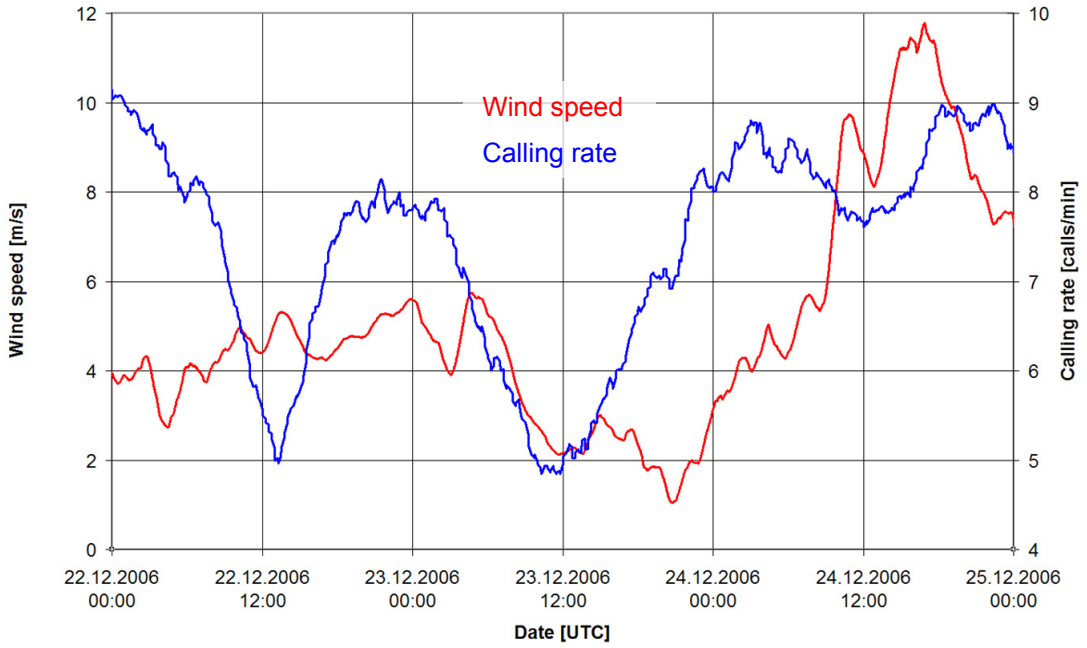


Figure 5.4: Wind speed near and calling rate at PALAOA station for the period 22 - 25 December 2006.

The analysis of the wind speed during this period showed, that on 24 December wind speed increased between 12 am and 2 pm from around 2 m s^{-1} to 12 m s^{-1} (see Figure 5.4). This increase in wind speed implied a drop in the windchill temperature. The overall outside temperature at 2 pm was around -3.5°C. Accounting for a wind speed of 12 m s^{-1}, the windchill temperature dropped down to -22°C. With such cold temperatures, individuals may have remained in the water instead of hauling out. The observation of increased calling rates

during this period supports the hypothesis that the diurnal pattern in vocal activity reflects leopard seal haul out behavior.

The direct correlation between number of seals in the water and vocalization rate could be a useful indicator for estimating the number of animals in a certain area. However, it has to be taken into account that calling rates also show some seasonal and diurnal variation over time. Rogers (2007) reported mean calling rates for male leopard seals between 2.8 (adults) and 3.8 (sub-adults) calls per minute[5.5]. Assuming these rates, a maximum of 3 leopard seals were vocalizing between 22 December and 24 December 2006 in the vicinity of PALAOA (highest vocalization rate: 8.9 calls per minute). Unfortunately Rogers (2007) did not provide the exact date for which the mean calling rates were calculated. Therefore, the derived numbers of vocalizing animals for the PALAOA sub data set mentioned above can only be a rough estimation. Further acoustic studies (e. g. by using acoustic tags) are necessary to investigate the degree of variability of calling rates among individuals, sexes and/or age classes, and over time, before calling rates can be reliably used in estimations of the number of vocalizing individuals.

[5.5] Recordings were made between November and early January between 4 pm and 3 am.

6. Data driven classification methods to study geographical and individual variations of leopard seal vocalizations

Applying standard analysis tools (for example principal component analysis (PCA), hierarchical cluster analysis or neural networks) to the data at hand, requires the definition of suitable parameters. The purpose of this chapter is to test various parameter sets for their suitability and to test the different methods for effectiveness. Finally the results are briefly discussed in the biological context.

6.1 Geographical variation

To investigate whether dialects or geographical variations exist in leopard seal vocalizations, recordings from Atka Bay (PALAOA), Drescher Inlet (DIPS) and Prydz Bay[6.1] (see Figure 6.1) were analyzed and compared.

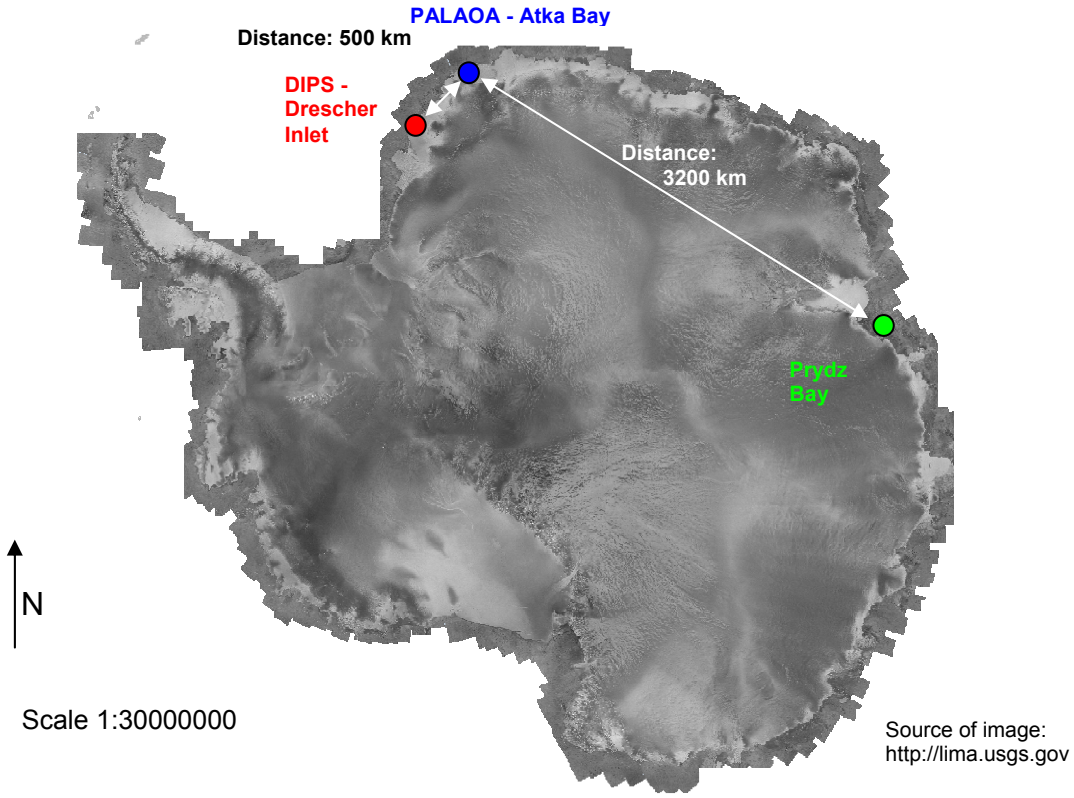


Figure 6.1: Map of Antarctica and adjacent coastal ocean with positions of the three geographic locations Atka Bay (blue), Drescher Inlet (red) and Prydz Bay (green).

[6.1] Kindly provided by Tracey Rogers

6.1.1 Principal component analysis of the data

6.1.1.1 Methods

For each of the three locations, 150 HDTs with high SNR (> 20 dB) were analyzed in detail. All chosen samples were recorded in mid-December, but during different years. To extract the following characteristic parameters (see Figure 6.2) for the 450 HDTs in a reasonable amount of time, the Matlab™ based GUI described in chapter 3.2.1 was used:

- RR1_start: Initial PRR - HDT part 1
- RR1_max: Maximum PRR - HDT part1
- RR1_end: Final PRR - HDT part 1
- RR2_start: Initial PRR - HDT part 2
- RR2_end: Final PRR - HDT part 2

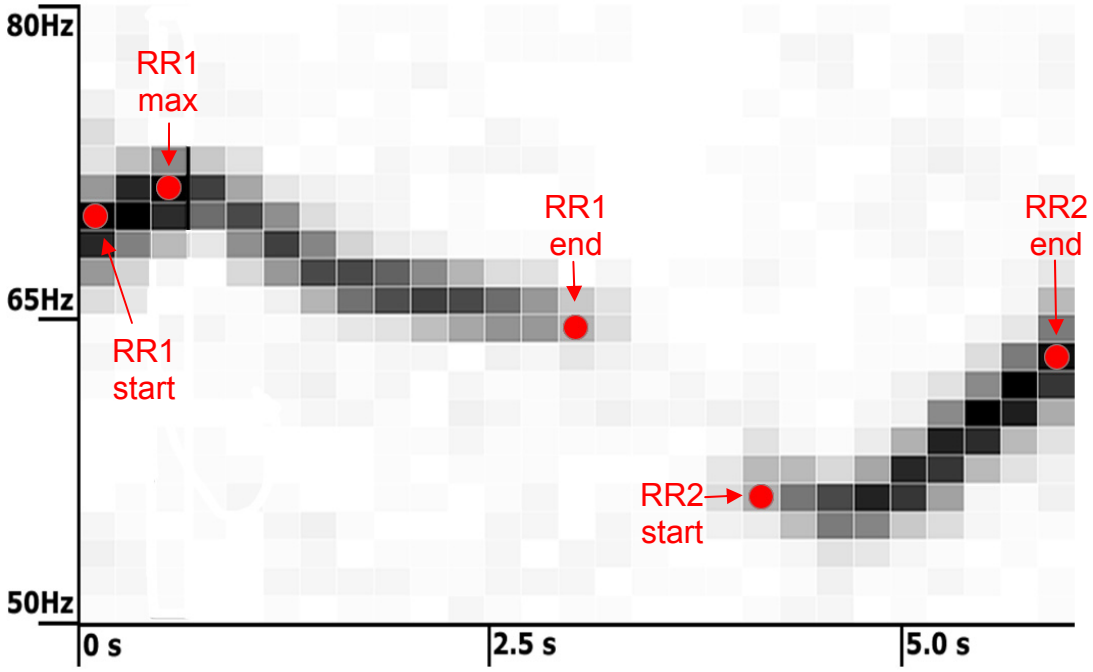


Figure 6.2: Example spectrogram of the HDT's pulse repetition rate and with parameters used in the analysis.

From these parameters, a matrix was generated which contained the 450 realizations of the five parameters. The parameter matrix was first normalized to zero mean and unity variance before applying a Principal Component Analysis (PCA) (Bahrenberg *et al.* 1992) to the data set.

The PCA transforms the possibly correlated input variables into uncorrelated variables, called principal components, and performs a coordinate rotation that aligns the transformed axes with the directions of maximum variance. The ensuing hierarchy allows a visualization of the vectors at a lower dimensionality which provides visual information on possible existing clusters.

6.1.1.2 Results and discussion

The first three principal components of the PCA (explaining 93.5% of the variance, see Table 6.1) are plotted in Figure 6.3. The clouds' respective color represents the three recording locations: Atka Bay (blue), Drescher Inlet (red) and Prydz Bay (green).

Table 6.1: Variance and cumulative variance covered by the principal components.

	% of variance	*% cumulative variance*
Component 1	79.1%	79.1%
Component 2	8.0%	87.1%
Component 3	6.4%	93.5%
Component 4	4.7%	98.2%
Component 5	1.8%	100.0%

Clearly, the characteristics of the calls recorded at Drescher Inlet (red) differ from those recorded at PALAOA and Prydz Bay. The clouds representing the PALAOA and the Prydz Bay recordings overlap, implying that they are not completely separable.

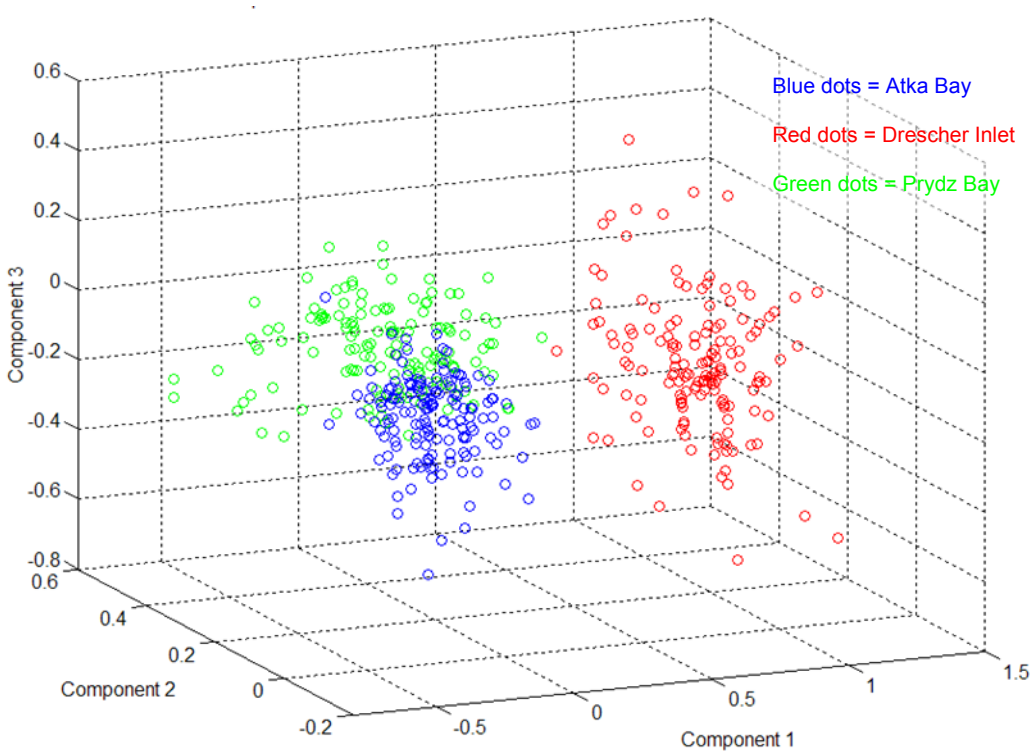


Figure 6.3: Scatterplot of the first three principal components of a PCA when applied to the PRR of 450 HDTs recorded at Atka Bay (blue), Drescher Inlet (red) and Prydz Bay (green).

6.1.2 Clustering using a hierarchical cluster analysis

6.1.2.1 Methods

To go beyond the visual impression of clustering, an unsupervised hierarchical cluster analysis (Bahrenberg *et al.* 1992) was conducted on the data set. The “unsupervised” approach disregards the known origin of the calls. The calculations followed the scheme:

- Calculating the Euclidian distance between all samples.
- Merge samples using the so-called ”Ward algorithm”. This algorithm starts out by representing each sample (450) as a single cluster. Based on a minimum variance criterion, the Ward algorithm proceeds to successively merge the samples until only one cluster (containing all 450 samples) remains. The result is a hierarchical cluster tree with 450 clusters at the top and 1 cluster at the bottom of the tree.
- Subjective selection of an appropriate level within the hierarchical cluster tree (e.g. the level of 3 clusters). The selection of an appropriate level will be guided by the scientific hypothesis to be addressed.

6.1.2.2 Results and discussion

Here the scientific question was whether the vocalizations recorded at the three locations can be clearly classified (see Table 6.2). Hence the level of three clusters was selected. As for all samples the geographic origin is known (but not used in the analysis), a sample is considered to be clustered ”correctly” if it was allocated to the cloud which contains the majority of calls from that geographic origin.

Table 6.2: Results of the cluster analysis (geographical data).

Location	*No. of samples*	*No. of correctly clustered data*	
Atka Bay (blue)	150	113	75.3%
Drescher Inlet (red)	150	150	100.0%
Prydz Bay (green)	150	71	47.3%

All samples recorded at Drescher Inlet were correctly assigned to a single cluster, while only 75% respectively 47% of the samples recorded at Atka Bay and Prydz Bay were correctly allocated to these respective clusters. Thus the cluster algorithm was not capable of separating the three locations unequivocally. For this reason a supervised classifier (taking the known origin of the calls into account) was tested on the data as this method could provide a higher discriminatory power.

6.1.3 Supervised classification using a neural network

6.1.3.1 Methods

Because the origin of each call is known, a supervised classification (taking the known origin of the calls into account) method can be applied to the data. A standard multilayer perceptron type neural network with 5 inputs, 10 hidden neurons and three outputs (see Figure 6.4), was trained to classify the calls according to the location.

At first, a randomly selected sample containing 75% of the data was extracted as the *training set*. The five parameters (as defined in Figure 6.2) were used as the input vector, and the three possible locations were coded as a three dimensional binary vector and used as the target outputs.

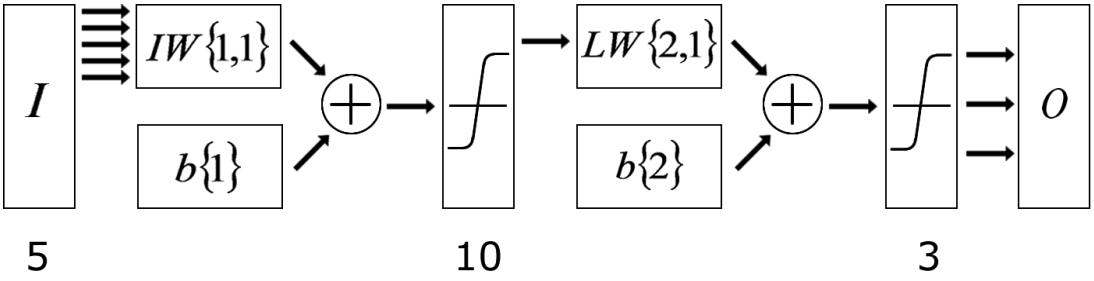


Figure 6.4: Schematic of the used neural network, as displayed by the MATLAB Neural Network Toolbox (Demuth and Beale, 2004). 5 inputs ***I*** (the call parameters) are mapped trough a hidden layer of 10 neurons to 3 outputs ***O*** (the locations). The weights (***IW, LW)*** and biases (***b***), which determine the functionality of the network are adjusted automatically by the training process.

The network was trained using the Levenberg Marquard method (Hagan and Menhaj 1994). In short, the algorithm adapts the weights of the network to minimize the error on the training set, i.e. the squared sum of the differences between the targets and the actual network outputs. Figure 6.5 shows the learning curve of the neural network. After 70 training epochs (iterated presentations of the data set for weight adaption) the error in the training set reached a plateau of minimum error. The representative parameters of the used neural network are given in the appendix.

The remaining 25% of the data were then used as the test set to evaluate the performance of the net on data not used for the previous training exercise (cross validation method). As expected, the network error on these data is slightly higher than on the training set because these data were unknown to the net, thus representing generalization performance instead of memorizing capability. As each location was represented by an output neuron which could generate values between zero and one, the location with the maximum activation was selected (“winner takes all” rule).

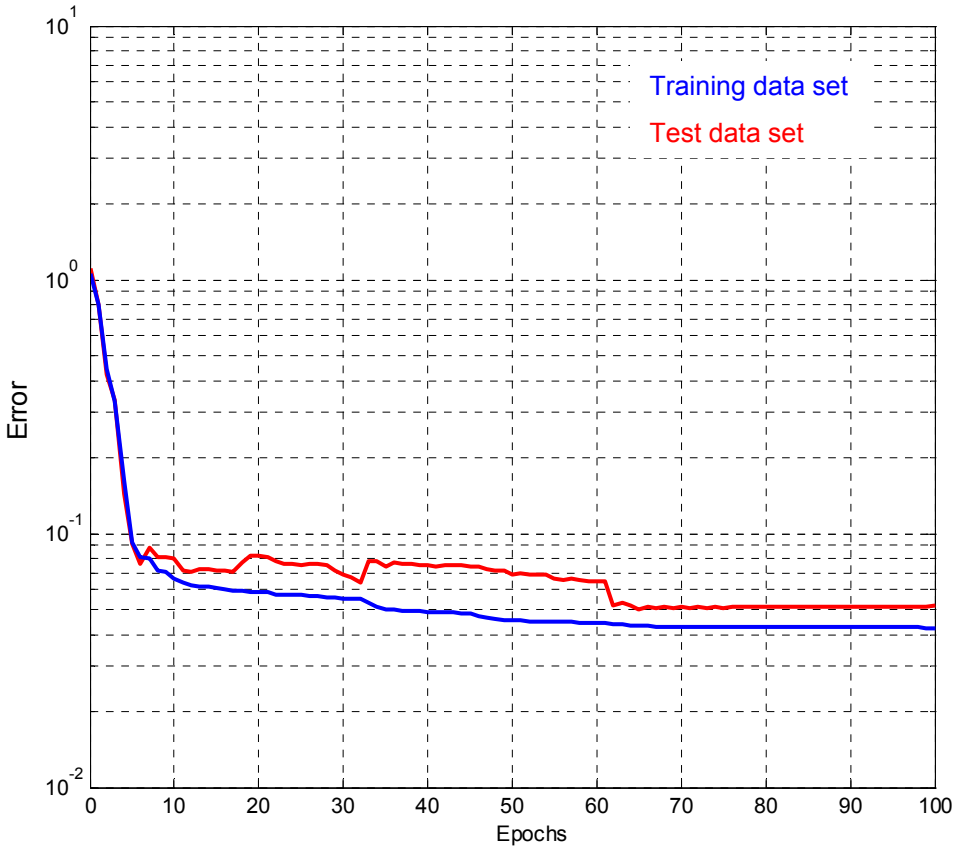


Figure 6.5: Evolution of the error of the neural network as a function of training epochs. Blue: Learning curve of the training data set (75% of the complete data set); Red: performance on test data set, which was excluded from training process (25% of the complete data set).

6.1.3.2 Results and discussion

The number of correct classifications when applying the neural network as given in the appendix is presented in Table 6.3.

Again, the network classified all calls from Drescher Inlet correctly, while the ambiguity between Atka Bay and Prydz Bay prevails. However, in comparison with the cluster analysis, a significantly higher portion of calls was recognized correctly from these regions, 91% (Atka Bay) and 86% (Prydz Bay). Hence, the network is capable of identifying the origin of the HDTs with 92% probability.

Table 6.3: Results of the neural network classification (geographical data).

Location	***No. of training samples***	***No. of test samples***	***No. of correctly classified data***	
Atka Bay (blue)	105	45	41	91%
Drescher Inlet (red)	118	32	32	100%
Prydz Bay (green)	114	36	31	86%
Total	337	113	104	92%

6.1.4 Overall results and discussion

Using the neural network 92% of the HDTs were correctly classified. While the data recorded at Drescher Inlet form an independent cluster, the data recorded at Atka Bay and Prydz Bay cannot be entirely separated, Surprisingly, this implies that the vocalizations between the two neighbored sampling sites, PALAOA and Drescher (distance 500 km), differ more than the vocalizations recorded at PALAOA and Prydz Bay (distance 3200 km).

These results coincide with recent finding in Weddell seals. Terhune *et al.* (2008) investigated male Weddell seal vocalizations recorded at 6 different locations in Antarctica including Atka Bay and Drescher Inlet. The analysis of the data revealed that the calls recorded at Drescher Inlet and PALAOA (distance 500 km) were more distinct from each other than they were to any other locations (distance up to 9000 km) included in the analysis.

A possible mechanism that could produce such a pattern has been discussed in the literature, suggesting that differences among similar species whose distributions overlap geographically are accentuated in regions where the species co-occur but are minimized or lost where the species' distributions do not overlap. This pattern results from evolutionary change driven by competition among species for a limited resource (e.g. food) and is called character displacement (Brown & Wilson 1956).

On a cautionary note however, it has to be mentioned that the number of animals which produce the analyzed vocalizations are unknown for all three sites. This is a problem as the clusters could represent individual variation rather than geographical variation. However, the observed calling rates at PALAOA and DIPS (see chapter 5.2.1) showed evidence that the data sets contain vocalizations of at least 2-3 animals. Furthermore, leopard seal vocalizations are continuously audible in the recordings investigated. Considering the temporal coverage of the recordings (1 week) and the fact that animals most likely stay on the ice for a certain time per day (haul out) it can be assumed that more than one animal was vocalizing. Nevertheless, further analyses on individual variation within leopard seal vocalizations are necessary. This issue is addressed to the next section.

6.2 Individual variation

The identification of individuals is interesting for two reasons:

- To derive the number of vocalizing animals in mono sound files and hence to distinguish between individual and geographic variation.

- To identify migrating animals when they return to the vicinity of the stations, or at different locations, possibly providing migratory information.

So far, individual leopard seals could not be identified by their characteristic acoustic features. Rogers & Cato (2002) analyzed field recordings of 5 seals for

individual variation. They concluded that the acoustic features used in their study (i.e. min. frequency, max. frequency, duration, amongst others) are not suitable to identify individuals. Here, we use the dominant pulse repetition rate of the first pulsed element (RR1_dom) in combination with spectral parameters of the HDT to address the same question but based on a different parameter set.

The above analysis of the geographical variation within leopard seal vocalizations showed that even though it was likely that vocalizations of at least two animals were recorded each at PALAOA and DIPS, it was not possible to (visually) identify sub classes in the PCA scatter plot. Furthermore, it was suspected that the differences within the PRR parameters between individual animals are small. For this reason further analysis concentrated on the RR1_dom. The following analysis aims to extract the RR1_dom at the highest possible resolution rather than deriving the temporal course of the whole PRR.

In human speech the so-called formants play an important role (formants are the distinguishing or meaningful frequency components of human speech) in identifying individuals. As the seals' vocal tract is similar to humans it appears appropriate to also include the HDTs main spectral characteristics (see Figure 6.6) in the analysis.

6.2.1 Analyzing a PALAOA subset of data

6.2.1.1 Methods

A subset of PALAOA data containing vocalizations of an unknown number of animals was investigated. The data set covered a 2 hour period recorded on 21 December 2006 between 12 am and 2 am. A set of 100 HDTs with high SNR (> 20 dB) was extracted. For each call, the following two parameters were determined:

- The RR1_dom frequency and
- the frequency range of the first pulsed element, as defined by F25 and F75 (see Figure 6.6).

To calculate the RR1_dom value automatically and with highest accuracy a Matlab™ routine was programmed. The routine calculated the power spectral density of the PRR and extracted the peak frequency. To get a maximum frequency resolution the length of the window (Hamming) used for the calculations was set to the total duration of the investigated first pulsed element. F25 and F75 describe the acoustic frequencies where 25% and 75% of the calls' total energy occur (see Figure 6.6). The F25 and F75 of the first pulsed element of the HDTs were automatically calculated using a built-in function of XBAT. The advantage of this method is the high degree of "objectiveness" of the measurement as the absolute maximum and minimum frequency of a signal are less affected by the SNR.

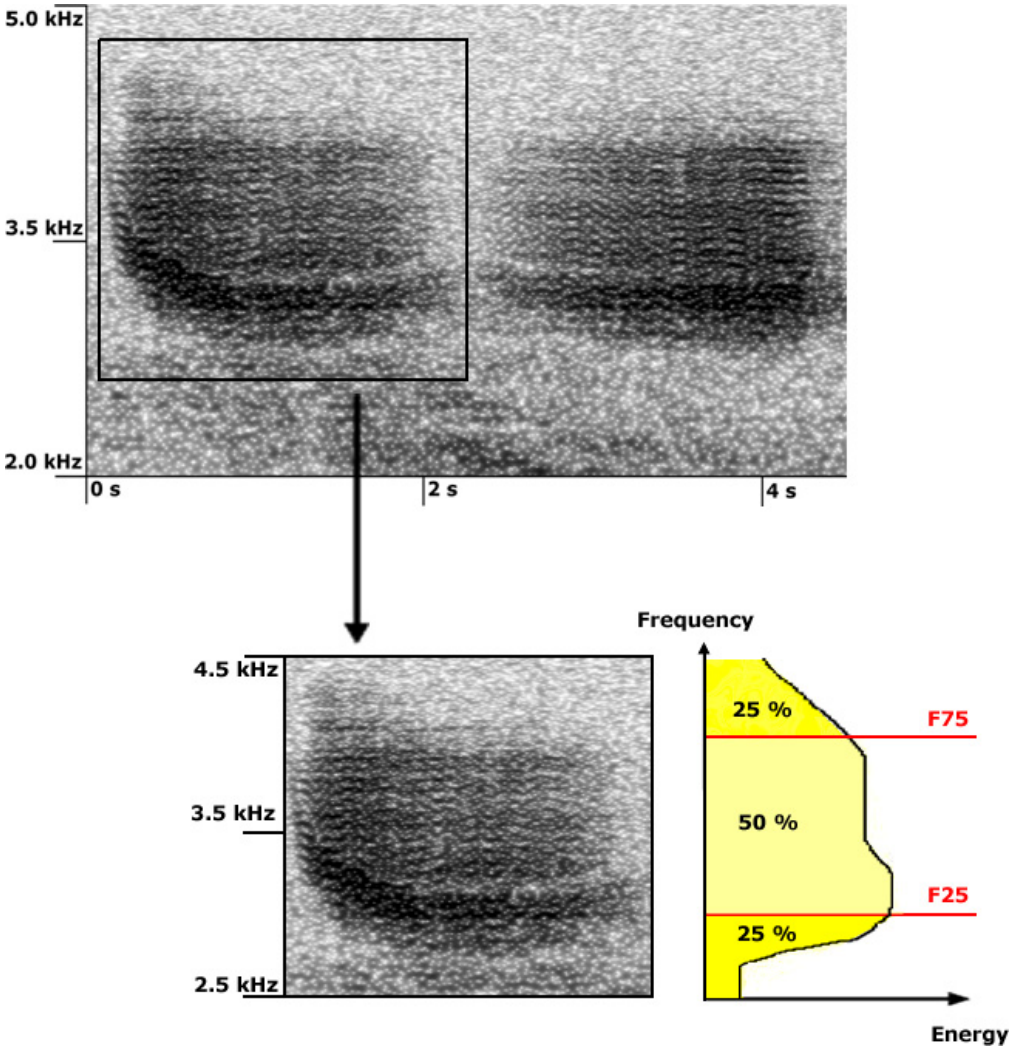


Figure 6.6: Spectrogram of a HDT (upper panel), first pulsed element of the HDT (lower left panel) and frequency-energy plot of the selected first pulsed element (lower right panel).

6.2.1.2 Results and discussion

A matrixplot of the analyzed parameters F25, F75 and RR1_max is given in Figure 6.7.

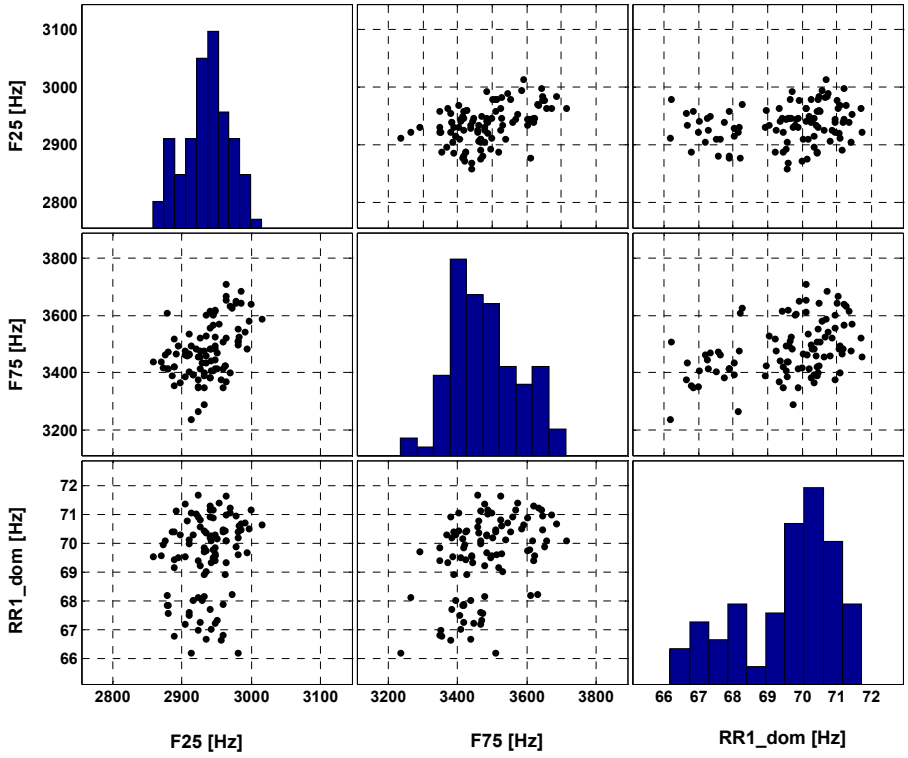


Figure 6.7: Matrixplot of the analyzed parameters (F25, F75 and RR1_dom) of HDTs of an unknown number of vocalizing animals. The boxes along the diagonal give the frequency distribution of the respective parameter.

The Figure depicts a bimodality of the parameter RR1_dom. Furthermore, the parameter F75 shows some evidence for bimodality. A quasi 3D view of parameters is given in Figure 6.8.

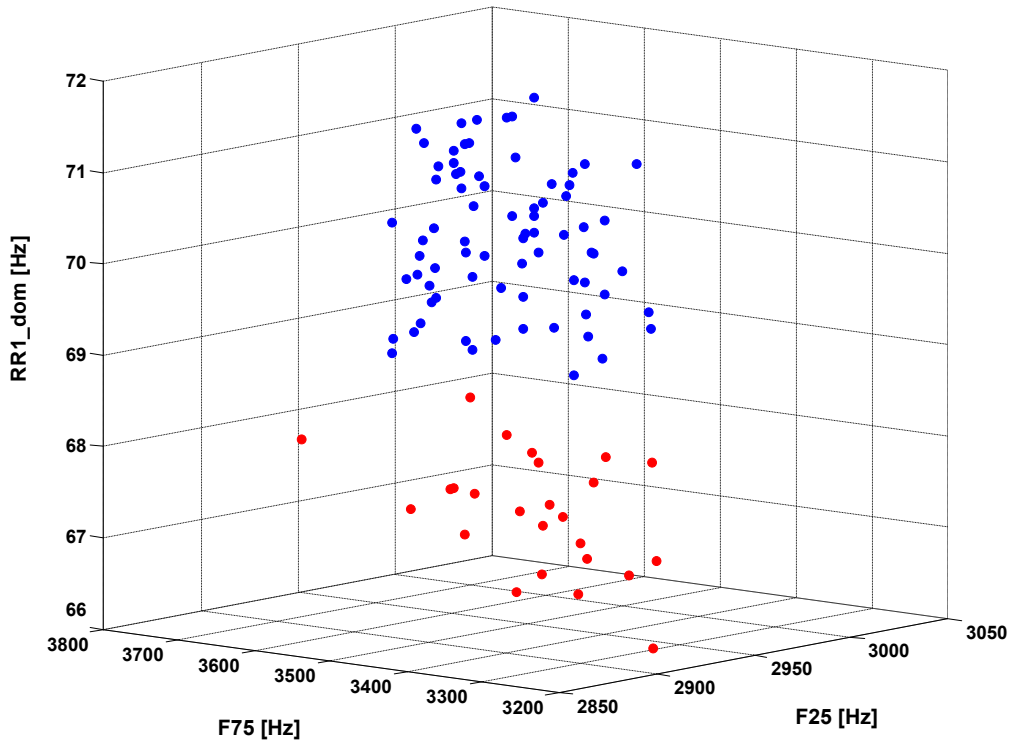


Figure 6.8: Scatterplot of the analyzed parameters (F25, F75 and RR1_dom) of HDTs of an unknown number of vocalizing animals. Red and blue data clouds subjectively colored by the author.

Two data clouds (subjectively colored blue and red) are observable in the scatterplot. This clustering gives a first hint that during the 2 hour period of observation at least 2 animals might have been vocalizing. The overall derived calling rate in the analyzed period was 7 calls per minute. Rogers (2007) reported mean calling rates for male leopard seals between 2.8 (adults) and 3.8 (sub-adults) calls per minute (see chapter 5.1.2). This implies that 2-3 animals were simultaneously vocalizing - a value that nicely fits the result of the cluster analysis and provides first evidence that individual leopard seals can be identified by the selected parameters.

Unfortunately the PALAOA 06 data are single-channel only, which prohibits validating the result by calculating the respective sources' positions. Future studies will attempt to extract the number of vocalizing individuals by analyzing the multi-channel data of PALAOA/PALAOA-S.

6.2.2 Analyzing a data set of known individual leopard seals

6.2.2.1 Methods

To test if individual leopard seals vocalizations can be separated by this method, acoustic data from known individuals are indispensable. For this reason Tracey Rogers was contacted as she had tried to identify individual leopard seals by their vocalizations in 2002 (Rogers & Cato 2002). She kindly provided data of individuals which were used for the analysis. The data set contained in total 105 HDTs: 41 HDTs of seal A, 32 HDTs of seal B and 32 HDTs of seal C.

6.2.2.2 Results and discussion

In Figure 6.9 a matrixplot of the parameters RR1_dom, F25 and F75 of the first pulsed element (Part 1) of the HDT is given.

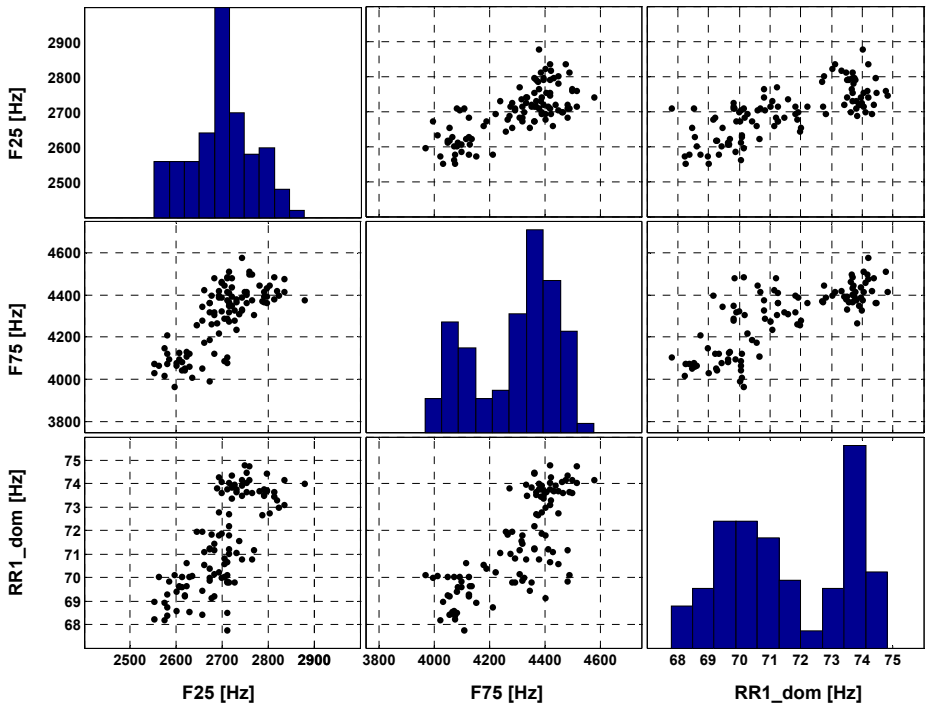


Figure 6.9: Matrixplot of the analyzed parameters (F25, F75 and RR1_dom) of HDTs of seal A, B and C. The diagonal gives the frequency distribution of the respective parameter.

While showing similar bimodalities as visible in Figure 6.7, the matrixplot also indicates that the three seals can not uniquely be identified by the RR1_dom parameter only. A combination of the three parameters is necessary to increase the discriminatory power. The scatterplot (see Figure 6.10) shows that the three seals can be distinguished quite well as three individual clusters when using the analyzed parameters F25, F75 and RR1_dom.

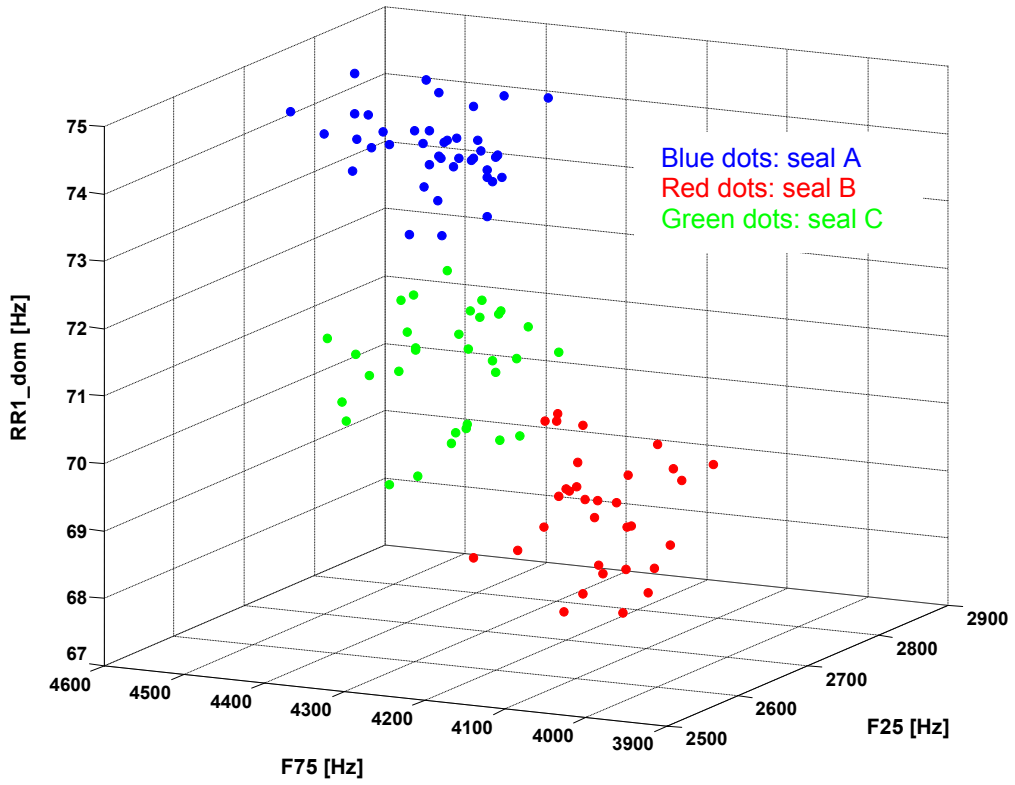


Figure 6.10: Scatterplot of the analyzed parameters (F25, F75 and RR1_dom) of HDTs of seal A, B and C.

However, to distinguish the individuals a cluster analysis (as described in chapter 6.1.2) was conducted.

The results of the cluster analysis, which are compiled in Table 6.4 show that all samples were correctly clustered. Consequently the seals can be clearly identified by the acoustic features of the HDTs.

Table 6.4: Results of the cluster analysis (individual data).

Location	*No. of samples*	*No. of correct clustered data*	
Seal A	41	41	100%
Seal B	32	32	100%
Seal C	32	32	100%

In summary, the combination of PRR and acoustic parameters allows distinguishing individual leopard seal vocalizations. This method significantly furthers investigations of leopard seals at an individual level and will therefore broaden the possibilities of passive bio-acoustic studies concerning leopard seal ecology considerably.

7. Résumé and outlook

7.1 Development and installation of PALAOA

Developing and installing PALAOA on the Antarctic ice shelf was successful. The observatory provides unique data from the coastal region of the Southern Ocean, featuring various biotic and abiotic sounds. To date (December 2007), more than 9400 hours (4.5 TByte) of acoustic data were collected.

An outstanding feature of PALAOA is the real-time data connection between PALAOA and the AWI in Bremerhaven. This data connection allows (a) analyzing a reduced-quality stream of the acoustic data in real-time and (b) to control and monitor PALAOA remotely from Germany.

The overall annual coverage of the acoustic recordings (as shown in Figure 7.1) however clearly indicates some remaining gaps for austral winter (June to August), which are caused by a lack of energy during this period.

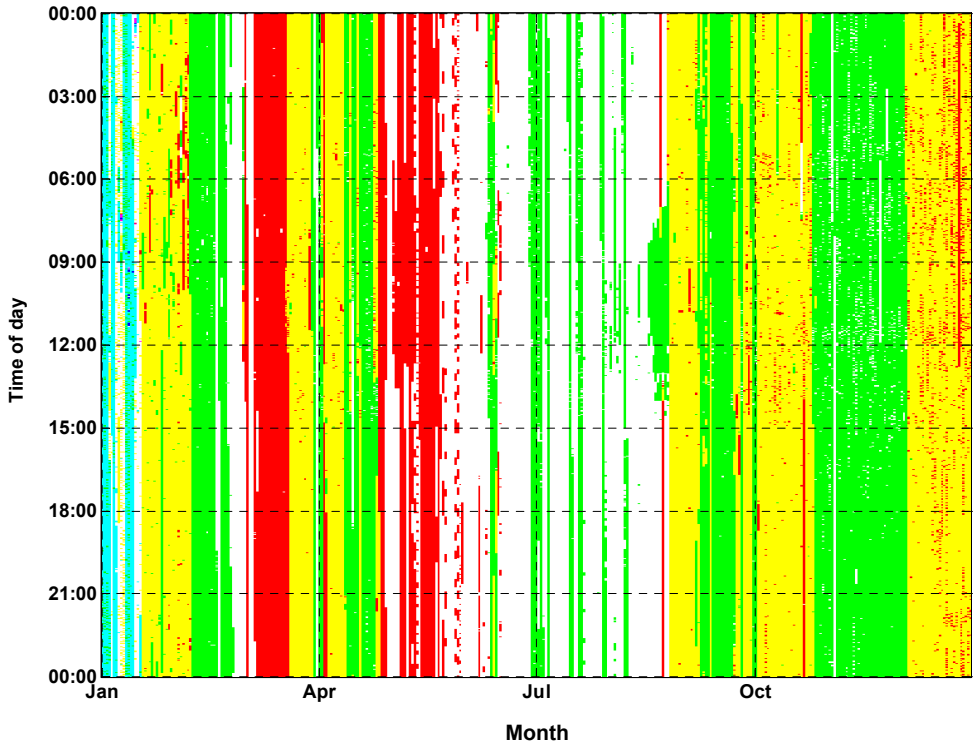


Figure 7.1: Overall annual coverage of PALAOA acoustic recordings as at 17 January 2008. For each day, the coverage is plotted vertically as colored dots representing a record of one minute duration. Red and green dots represent recordings made in 06 or 07 (exclusively), yellow dots represent recordings made during both 06 and 07, and cyan dots indicate recordings in 06, 07 and 08.

PALAOA was developed iteratively over a period of 3 years (between 2004 and 2007) and was continuously improved and optimized with respect to saving power. In addition to the solar panels and wind generators, a third independent back-up power supply appears necessary for calm periods during austral winter. Attempts to use a fuel cell for this purpose fell short of success but latest

improvements in this technology appear promising and will be tested in the next season.

After the successful deployment of the hydrophones by means of a hot water drilling operation, the capability of the acoustic array of PALAOA to locate acoustic sources was impaired by the loss of two hydrophones in the first year of operation. To overcome this drawback, a satellite station (PALAOA-S) was developed and successfully operated during the field season 2006/2007. PALAOA-S is synchronized with PALAOA using GPS time signals to provide an alternative third data channel for localization. The next development step, comprising several satellite listening stations, is planned for deployment during the 2008/2009 field season. The stations will be deployed on the sea ice during austral summer during the peak of vocal activity of marine mammals, as revealed by the current PALAOA dataset. The distance between the satellite stations will amount to several kilometers. This long base-line array will allow localization of even distant sound sources with high precision. It can be used to obtained source levels of biotic and abiotic signals as well as spatially resolved mammalian track data for behavioral studies.

7.2 Automated detection and diver safety

For the first time ever, automated detectors were developed in this study for the detection of pinniped vocalizations. The two complementary algorithms focus on detection of the predominant leopard seal vocalizations: the low (LDT) and high (HDT) double trills. Both algorithms are based on techniques commonly used in human speech recognition. The LDT detector is based on the signals' zero crossing rate (ZCR) while the detector for the HDT utilizes the temporal changes in the pulse repetition rate (PRR) of the amplitude modulated signal in combination with a Hidden Markov Model (HMM) for the detection. For both detectors a performance test was conducted on a reference data set of one week duration. The performance test revealed that each detector allows a fast and reliable detection of its respective leopard seal vocalization type. The detection rates of the detectors were estimated at 85% for HDTs (SNR > 14 dB) and 74% for LDTs (SNR > 13 dB).

Once constructed and validated, the detectors were applied to large data sets to extract information on seasonal and temporal patterns of the vocal behavior of leopard seals. The analysis revealed that leopard seal vocalizations are present in the recordings from the end of September to mid-January. This finding provides new information, as leopard seal calls have so far been reported to occur between November and mid-January only. With regard to diving operations, this new finding suggests that the presence of leopard seals can be acoustically determined during the 4 months spanning austral summer.

With LDTs and HDTs circumpolar distribution, the developed detectors allow their use in areas beyond the Weddell Sea. On a cautionary note however, it has to be mentioned that the reliability of the detectors is dependent on the calls' signal to noise ratio. Thus, the detection of leopard seal vocalizations in the vicinity of ships and their increased noise level will be impaired, if not

impossible in real time. To avoid this difficulty the detectors may be used to screen acoustic recordings obtained prior to arrival of the ship at the sampling site in a relatively short time.

The developed techniques will be used to further analyze the continuously growing PALAOA data set for leopard seal vocalizations. Furthermore, both detectors seem to be suitable to be fine tuned to other marine mammal vocalizations. Consequently future work on the detectors will focus on broadening their application.

7.3 Identification of individual leopard seals

This study presented for the first time a method which permits the identification of individual leopard seals by the pulse repetition rate (PRR) feature derived from the HDT. The performance of the classification system was tested by analyzing data sets from three known individuals. The system correctly allocated all 105 seal vocalizations of a test set to the corresponding seal.

This result will broaden the potential of acoustic studies of the leopard seal. The identification of individuals allows estimating the number of vocalizing animals within an area from mono sound files. Multi-channel recordings and a complex calculation of the sound sources' positions might - at least for the purpose of identification - even become obsolete. Furthermore, the method enables - in combination with long-term recordings - the study of migration patterns of individuals. Future studies could, for example, deal with the question whether the same individuals return to Atka Bay every year or if different animals migrate to this area from year to year.

7.4 Leopard seal ecology

The analysis of the PALAOA data set using the automated detection algorithms revealed seasonal and diurnal patterns of the vocal behavior of the leopard seal. The results of the analysis indicate that the calling rate is affected by two factors: (1) the number of animals in the water and (2) the season.

Leopard seals show - as do other seal species - a diurnal pattern in the calling rate which is most likely correlated with the haul-out behavior of the animals. During late forenoon, when most of the animals rest on the ice, calling rates are low. In contrast calling rates peak in night time. The observed diurnal pattern in calling rates is, however, influenced by the wind chill temperature as the seals re-enter the water (possibly for energetic reasons) when the wind chill temperature is low.

The observed calling rates vary widely on a seasonal time scale. Highest calling rates were observed during mating season from early December to mid-January. However, leopard seals also vocalize at lower calling rates outside the actual mating season and can thus be investigated using passive acoustic techniques for extended periods.
In 2003, acoustic recordings conducted in the Drescher Inlet contained many leopard seal vocalizations but not a single individual was sighted by the

experienced seal biologists who spent several hours each day on the sea ice. Thus it is not surprising that airborne and ship-based visual surveys of leopard seals aimed at deriving population densities are subject to high uncertainties. Because of the high vocalization rates during the mating season, this period seems to be quite suitable to conduct acoustic surveys on leopard seals. In this context, the results of this thesis provide information which is useful in several regards:

(1) to plan acoustic surveys (considering the seasonal and diurnal calling patterns),

(2) to screen/analyze the recorded data in a reasonable amount of time (using the developed detectors) and

(3) to derive the number of vocalizing animals by identifying individuals (using the developed identification algorithm).

Overall, this study successfully built a reliable tool to obtain acoustic underwater recordings for one of the most remote regions on this planet and developed automated detection and identification tools for systematic analysis of the resulting, continuously growing, large acoustic data sets.

8. Acknowledgements

First of all I want to thank my wife Karolin. She was very patient with me and ensured that I was not troubled with other problems during the critical state of my thesis. Without her support I wouldn't have finished the work in the predetermined time-frame.

Special thanks to Olaf Boebel for his support, guidance and particularly in pushing me to be efficient. I really appreciated the discussions (in particular on Sunday mornings) on earlier versions of the manuscript and learned a lot about scientific writing and structuring work from him.

Lars Kindermann and I developed together PALAOA and he also provided the PALAOAdB database system. I have never met a person with such a broad and consolidated knowledge before. Working together with him in the lab or the field was always a pleasure. I really learned a lot from him - especially in engineering, programming and developing unorthodox ideas and solutions.

For supporting the manual analysis of the PALAOA data set and spending weeks listening to the PALAOA data stream I am indebted to Cornelia Kreiß. She also conducted the literature research on the leopard seals' biology.

Horst Bornemann, Elke Burkhardt, Ilse van Opzeeland and Jochen Plötz – I appreciate their helpful comments on prior manuscripts of the thesis, the pleasant atmosphere in the working group and the relaxing coffee breaks. I will definitely miss this in the US.

Harold Figueroa helped me with the development and implementation of the ZCR detector into XBAT. It was really great to work with Harold and his awesome software.

I want to thank Marie Roch for her suggestions on the HMM detector. Marie is always helpful and I'm really looking forward to collaborating with her on the detection and classification of odontocete vocalizations.

I thank my official supervisors Prof. Wolfhard Symader and Prof. Peter Lemke for supervising my thesis and writing the expertise.

The hot water drilling operation was supervised by Albert Ziffer and Holger Schubert. Without their help the deployment of the acoustic instruments wouldn't have been possible.

The seismic study to determine the ice shelf thickness was supervised by Daniel Steinhage and Sven Riedel. Thanks for your help.

Field work and maintenance of PALAOA would not have been possible without the help of the AWI logistics department and the overwinterer teams of Neumayer Base 2005, 2006 and 2007. Special thanks to Peter Hennig, Marc Brüggemann, Eric-Roger Brücklmeier and Mirko Denecke.

My thanks to Roger Verhoeven, Christian Müller and Cord Drücker for their help in setting up PALAOA 05 and PALAOA 06 in Bremerhaven.

Apart from that I want to express my gratefulness to the BIS – Bremerhavener Gesellschaft für Innovationsförderung und Stadtentwicklung for funding the development of the PALAOA energy and control module. Furthermore thanks to Saad El Naggar for encouraging me to apply for the money and Edith Davidis for administrating the project.

Sharon Nieukirk, Elke Klinck and Steffi Meyer thanks a lot for proofreading the entire thesis.

Last but not least I owe gratitude to my family especially Bernd, Elke, Isabelle, Peter and Rena and parents-in-law Christof and Roswitha for supporting me and my studies.

9. Bibliography

- A -

Ainslie, M. A. & McColm, J. G. (1998): A simplified formula for viscous and chemical absorption in sea water. In: Journal of the Acoustical Society of America, 103(3), pp. 1671 - 1672.

- B -

Bahrenberg, G., Giese, E. & Nipper, J. (1992): Statistische Methoden in der Geographie, Band 2, Teubner Verlag,Stuttgart, 416 pp.

Bester, M. N., Ferguson, J. W. H. & Jonker, F. C. (2002): Population densities of pack ice seals in the Lazarev Sea, Antarctica. In: Antarctic Science, 12(2), pp. 123 - 127.

Bonner, W. N. (2004): Seals and seal lions of the world. Facts on File Inc, New York, 224 pp.

Borsa, P. (1990): Seasonal occurrence of the leopard seal, *Hydrurga leptonyx*, in the Kerguelen Islands. In: Canadian Journal of Zoology, 68, pp. 405 - 408.

Brown, K. G. (1957): The leopard seal at Heard Island 1951- 54. In: Australian National Antarctic Research Expedition Interim Report, 16, pp. 1 - 34.

Brown, W. L. & Wilson, E. O. (1956): Character displacement. In: Systematic Zoology, 5, pp. 49 - 65.

Buxhoeveden, C., Fahrbach, E., Plugge, R. & Schütt, E. (1992): Water level measurements. In: Bathmann, U., Schulz-Baldes, M., Fahrbach, E., Smetacek, V. & Hubberten, H.-M. (Eds.): Reports on Polar Research, 100, pp. 56 - 57.

Buness, A. H., Druivenga, G. & Wiederhold, H. (2000): SISSY - eine tragbare und leistungsstarke seismische Energiequelle. In: Geologisches Jahrbuch, Serie E (52), pp. 63 - 88.

- C -

Catchpole, C. K. & Slater, P. J. B. (1995): Themes and Variations. In: Catchpole, C. K. & Slater, P. J. B (Eds.): Bird Song: Biological Themes and Variation, Cambridge University Press, Cambridge, pp. 164 - 173.

- D -

Deller, J. R., Hansen, F. H. L. & Proakis, J. G. (2000): Discrete-Time Processing of Speech signals. In: Wiley-IEEE Press, Chapter 12, pp. 677 - 744.

Demuth, H. & Beale, M. (2004): Neural Network Toolbox User's Guide, The MathWorks Inc, Natick, 447 pp.

- E -

El-Sayed, S. Z. (1971): Biological aspects of the pack ice ecosystem. In: Deacon, G. (Ed.): Symposium of Antarctic ice and water masses, Heller, Cambridge, pp. 35 - 54.

Erickson, A. W., Siniff, D. B., Cline, D. R. & Hofmann, R. J. (1971): Distributional ecology of Antarctic seals. In: Deacon, G. (Ed.): Symposium of Antarctic ice and water masses, Heller, Cambridge, pp. 55 - 76.

- F -

Fahrbach, E. & Rohardt, G. (2008): Die Zirkulation im Weddellmeer - ein Fenster zur Tiefsee. In: Fütterer, D. K. & Fahrbach, E. (Eds.): Polarstern - 25 Jahre Forschung in Arktis und Antarktis. Delius Klasing Verlag, Bielefeld, pp. 137 - 147.

- G -

Geraci, J. R. (1975): Pinniped nutrition. In: Ronald K. & Mansfield A.W. (Eds.): Symposium on the biology of the seal (Guelph, Ontario, August, 1975). Publications of the International Council for the Exploration of the Sea, 169, pp. 312 - 32.

Gilbert, J. R. & Erickson, A.W. (1977): Distribution and abundance of seals in the pack ice of the Pacific Sector of the Southern Ocean. In: Llano, G. (Ed.): Adaptions within Antarctic Ecosystems, 3rd SCAR Symposium on Antarctic Biology, Smithonian Institute, Washington, D.C., SCAR, pp. 703 - 740.

Gwynn, A. M. (1953): The status of leopard seals at Heard Island and Maquarie Island, 1948-1950. In: Australian National Antarctic Research Expedition Interim Report, 3, pp. 1 - 33.

- H -

Hagan, M. T. & Menhaj, M. (1994): Training feedforward networks with the Marquardt algorithm. In: IEEE Transactions on Neural Networks, 5(6), pp. 989 - 993.

Hall-Aspland, S. A., Rogers, T. & Canfield, R.B. (2005): Stable carbon and nitrogen isotope analysis reveals seasonal variation in the diet of leopard seals. In: Marine Ecology Progress Series, 305, pp. 249 - 259.

Hall-Aspland, S. A. & Rogers, T. L. (2004): Summer diet of leopard seals (*Hydrurga leptonyx*) in Prydz Bay, Eastern Antarctica. In: Polar Biology, 27(12), pp. 729 - 734.

Heijmans, H. J. A. M. (1997): Composing morphological filters. In: IEEE Transactions on Image Processing, 6(5), pp. 713 - 723.

Hempel, G. (1987): The krill-dominated pelagic system of the Southern Ocean. In: Environment International, 13, pp. 33 - 36.

Hiruki, L. M., Schwartz, M. K. & Bovenp, P. L. (1999): Hunting and social behaviour of leopard seals (*Hydrurga leptonyx*) at Seal Island, South Shetland Islands, Antarctica. In: Journal of Zoology, 249(1), pp. 97 - 109.

Hofman, R. J., *et al*. (1977): The leopard seal (*Hydrurga leptonyx*) at Palmer Station, Antarctica. In: Symposium on Antarctic biology, pp. 769 - 782.

- I -

Insley, S. J. (2001): Mother–offspring vocal recognition in northern fur seals is mutual but asymmetrical. In: Animal Behaviour, 61, pp. 129 - 137.

- J -

Jefferson, T. A., Leatherwood S. & Webber, M. A. (1993): Marine mammals of the world. FAO species identification guide, Rome, 320 pp.

Jessopp, M. J., Forcada, J., Reid, K., Trathan P. N. & Murphy E. J. (2004): Winter dispersal of leopard seals (*Hydrurga leptonyx*): environmental factors influencing demographics and seasonal abundance. In: Journal of Zoology, 263, pp. 251 - 258.

- K -

Kaiser, M., Attrill, M., Jennings, S., Thomas, D. N., Barnes, D., Brierley, A., Polunin, N., Raffaelli, D. & Williams, P. Le B. (2005): Marine Ecology - Processes, Systems, and Impacts, Oxford University Press, New York, 584 pp.

King, J.E. (1983): Seals of the world, 2nd edition. In: British Museum (Natural History), London and Oxford University Press, Oxford, 240 pp.

Klinck, H. & Burkhardt, E. (2008): Marine Mammal Automated Perimeter Surveillance - MAPS - cruise report of the ANT-XXIII-6 expedition. In: Reports on Polar and Marine Research, 580, pp. 114 - 121.

Kohnen, H. (1974): The temperature dependence of seismic waves in ice. In: Journal of Glaciology, 67(13), pp. 144 - 147.

Kooyman, G. L. (1981): Leopard Seal *Hydrurga leptonyx*, Blainville, 1820. In: Ridgway S. H. (Eds.): Handbook of marine mammals, Academic Press, London & San Diego, 2, pp. 261 - 274.

Kovesi, M. (2000): MATLAB and Octave Functions for Computer Vision and Image Processing. In: School of Computer Science & Software Engineering, University of Western Australia. Available from: http://www.csse.uwa.edu.au/~pk/research/matlabfns.

Kovesi, M. (1999): Phase Preserving Denoising of Images. In: The Australian Pattern Recognition Society Conference (DICTA '99), December 1999, Perth, pp. 212 - 217.

Kuhn, C. E., McDonald, B. I., Shaffer, S. A., Barnes, J., Crocker, D. E., Burns, J. & Costa D. P. (2006): Diving physiology and winter foraging behavior of a juvenile leopard seal (*Hydrurga leptonyx*). In: Polar Biology, 29(4), pp. 303 - 307.

- L -

Laws, R. M. (1993): Antarctic seals: research methods and techniques, Cambridge University Press, Cambridge, 412 pp.

Laws, R. M. (1984): Antarctic Ecology, London: Academic Press, 344 pp.

Lesinski, G. (1993): Monitoring of birds and pinnipedians on King George Island (South Shetland Islands) in 1989/1990. In: Polar Research, 14, pp. 75 - 89.

Lowry, L. F., Testa, J.W. & Calvert, W. (1988): Notes on Winter Feeding of Crabeater and Leopard Seals near the Antarctic Peninsula. In: Polar Biology, 8, pp. 475 - 478.

- M -

Mellinger, D. K. & Clark, C.W. (2000): Recognizing transient low-frequency whale sounds by spectrogram correlation. In: Journal of the Acoustic Society of America, 107(6), pp. 3518 - 3529.

Muir, S. F., Barnes, D. & Reid, K. (2006): Interactions between humans and Leopard seals. In: Antarctic Science, 18(1), pp. 61 - 74.

- N -

Nelson, B. S. & Marler, P. (1993): Innate recognition of song in whitecrowned sparrows: A role in selective vocal learning? In: Animal Behaviour, 46, pp. 806 - 808.

Nottebohm, F. (1972): The origins of vocal learning. In: American Naturalist, 106, pp. 116 - 135.

Nishimura, Clyde E. & Conlon, Dennis M. (1994): IUSS Dual Use: Monitoring whales and earthquakes using SOSUS. In: Marine Technology Society Journal, 27(4), pp. 13 - 21.

Nixdorf, U., Dunker, E., Eckstaller, A., Frenzel, A., Hannke, S., Lambrecht, A., Rohardt, G., Spindeldreher, C. & Ziffer, A. (1997): Schelfeis-Ozean-Wechselwirkung. In: Reports on Polar Research, 219, pp. 69 - 72.

Nixdorf, U., H. Mandler, Wege, C. & Ziffer, A. (1994a): Heißwasserbohrung. In: Reports on Polar Research, 152, pp. 191 - 195.

Nixdorf, U., Oerter, H. & Miller, H. (1994b): First access to the ocean beneath Ekströmisen, Antarctica, by means of hot water drilling. In: Annals of Glaciology, 20, pp. 110 - 114.

- O -

Øritsland T. (1977): Food consumption of seals in the Antarctic pack ice. In: Llano, G. (Ed.): Adaptations within Antarctic ecosystems. Proceedings of the 3rd SCAR symposium on Antarctic biology. Smithsonian Institute, Washington DC, pp. 749 - 768.

- P -

Perrin, W. F. & Bronwell, J. R. L. (2002): Minke whales Balaenoptera acutorostrata and B. bonaerensis. In: Perrin, W. F., Würsig, B. & Thewissen, J. G. M. (Eds.): Encyclopedia of Marine Mammals, San Diego, Academic Press, pp. 750 - 754.

Podos, J., Nowicki, S. & Peters, S. (1999): Permissiveness in the learning and development of song syntax in swamp sparrows. In: Animal Behaviour, 58, pp. 93 - 103.

- R -

Rabiner, L. & Juang, B. H. (1993): Fundamentals of Speech Recognition, Prentice Hall PTR, New York, 496 pp.

Ralls, K., Fiorelli, P. & Gish, S. (1985): Vocalizations and vocal mimicry in captive harbour seals, *Phoca vitulina*. In: Canadian Journal of Zoology, 63, pp. 1050 - 1056.

Ray, G. C. (1970): Population ecology of Antarctic seals, Volume 1. In: Holdgate, M. W. (Ed.): Antarctic Ecology, Academic Press, New York, pp. 398 - 414.

Reeves, R. R., Stewart, B. S., Clapham, P. J. & Powell, J. A. (2002): National Audubon Society: Guide to Marine Mammals of the World. Alfred A. Knopf, New York, 531 pp.

Richardson, W. J., Green Jr., C. R., Malme, C. I. & Thomson, D. H. (1995): Marine Mammals and Noise. Academic Press, London & San Diego, 576 pp.

Riedmann, M. L. (1990): The pinnipeds: seals, sea lions and walruses, University of California Press, Berkeley, 439 pp.

Rogers, T. L. (2007): Age-related differences in the acoustic characteristics of male Leopard seals, *Hydrurga leptonyx*. In: Journal of the Acoustic Society of America, 122(1), pp. 596 - 605.

Rogers, T.L., Hogg, C. & Irvine, A. (2005): Spatial movement of adult leopard seals (*Hydrurga leptonyx*) in Prydz Bay, Eastern Antarctica. In: Polar Biology, 28(6), pp. 456 - 463.

Rogers, T. L. (2003): Factors influencing the acoustic behaviour of male phocid seals. In: Aquatic Mammals, 29(2), pp. 247 - 260.

Rogers, T. L. & Cato, D. H. (2002): Individual variation in the acoustic behaviour of the adult male leopard seal, *Hydrurga leptonyx*. In: Behaviour, 139, pp. 1267 - 1286.

Rogers, T. L. & Bryden, M.M. (1997): Density and haul-out behaviour of leopard seals (*Hydrurga leptonyx*) in Prydz Bay, Antarctica. In: Marine Mammal Science, 13(2), pp. 293 - 302.

Rogers, T. L., Cato, D. H. & Bryden, M. M. (1996): Behavioural significance of underwater vocalizations of captive leopard seals, *Hydrurga leptonyx*. In: Marine Mammal Science, 12(3), pp. 414 - 427.

Rogers, T. L., Cato, D. H. & Bryden, M. M. (1995): Underwater vocal repertoire of leopard seals (*Hydrurga leptonyx*) in Prydz Bay, Antarctica. In: Kastelein, R. A., Thomas, J. A. & Nachtigall, P. E. (Eds.): Sensory Systems of Aquatic Mammals, De Spil Publishers, Woerden, pp. 223 - 236.

Rounsevell, D. & Eberhardt, I. (1980): Leopard seals *Hydrurga leptonyx* at Maquarie Island from 1949-1979. In: Australian Wildlife Research, 7, pp. 403 - 415.

- S -

Siniff, D. B. (1991): An overview of the Ecology of Antarctic seals. In: American Zoology, 31, pp. 143 - 149.

Siniff, D. B. & Stone, S. (1985): The role of the Leopard Seal in the Tropho-Dynamics of the Antarctic Marine Ecosystem. In: Antarctic Nutrient Cycles and Food webs, pp. 555 - 559.

Sirovic, A., Hildebrand, J. A. & Wiggins, S. M. (2007): Blue and fin whale call source levels and propagation range in the Southern Ocean. In: The Journal of the Acoustical Society of America, 122(2), pp. 1208 - 1215.

Southwell, C., Kerry, K., Ensor, P., Woehler, E.J. & Rogers, T. (2003): The timing of pupping by pack-ice seals in East Antarctica. In: Polar Biology, 26, pp. 684 - 652.

Steinhage, D., Nixdorf, U., Meyer, U. & Miller, H. (1999): New maps of the ice thickness and subglacial topography in Dronning Maud Land, Antarctica, determined by means of airborne radio-echo sounding. In: Annals of Glaciology, 29, pp. 267 - 272.

Stirling, I. & Siniff, D. B. (1979): Underwater vocalizations of leopard seals (*Hydrurga leptonyx*) and crabeater seals (*Lobodon carcinophagus*) near South Shetland Islands, Antarctica. In: Canadian Journal of Zoology, 57, pp. 1244 - 1248.

- T -

Terhune, J. M., Quin, D., Dell'Apa, A., Mirhaj, M., Plötz, J., Kindermann, L. & Bornemann, H. (2008): Geographic variation in underwater male Weddell seal Trills suggest breeding area fidelity. In: Polar Biology, 31(6), pp.671 - 680.

Thomas, J. A. & Golladay, C.L. (1995): Geographic variation in leopard seal (*Hydrurga leptonyx*) underwater vocalizations. In: Kastelein, R. A., Thomas, J. A. & Nachtigall, P. E. (Eds.): Sensory Systems of Aquatic Mammals, De Spil Publishers, Woerden, The Netherlands, pp. 201 - 222.

Thomas, J. A., Fisher, S. R., Evans, W. E. & Awbrey, F. T. (1983): Ultrasonic vocalizations of leopard seals (*Hydrurga leptonyx*). In: Antarctic Journal of the US, 17, page 186.

Thomas, J. A. & DeMaster, D. P. (1982): An acoustic technique for determining diurnal activities in leopard (*Hydrurga leptonyx*) and crabeater (*Lobodon carcinophagus*) seal. In: Canadian Journal of Zoology, 60, pp. 2028 - 2031.

Tyack, P. L. (1998): Acoustic communication under the sea. In: Hopp, S. L., Owren M. J. & Evans C. S. (Eds.): Animal Acoustic Communication, Springer, Berlin, Germany, pp. 163 - 220.

- V -

Van Parijs, S. M. (2003): Aquatic mating in pinnipeds: a review. In: Aquatic Mammals, 29(2), pp. 214 - 226.

Van Trees, H. L. (1968): Detection, Estimation and Modulation Theory. John Wiley & Sons Inc, New York, Vol. I, pp. 712.

- W -

Walker, T. R., Boyd, I. L., McCafferty, D. J., Huin, N., Taylor, I. R. & Reid, K. (1998): Seasonal occurrence and diet of leopard seals (*Hydrurga leptonyx*) at Bird Island, South Georgia. In: Antarctic Science, 10(1), pp. 75 - 81.

Watkins, N. W., Clilverd, M. A., Smith, A. J. & Yearby, K. H. (1998): A 25-year record of 10 kHz sferics noise in Antarctica: Implications for tropical lightning levels. In: Geophysical Research Letters, 25(23), pp. 4353 - 4356.

Watkins, N. W., Clilverd, M. A., Smith, A. J., Rodger, C. J., Bharmal, N. A. & Yearby, K. H. (2001): Lightning atmospherics count rates observed at Halley, Antarctica. In: Journal of Atmospheric and Solar-Terrestrial Physics, 63, pp. 993 - 1003.

10. Appendix

Parameters of the HMM model used in this study:

$$cb = \begin{Bmatrix} 0.1353 & 0.1371 & 0.1328 & 0.1340 & 0.1386 & 0.1394 & 0.1348 & 0.1291 & 0.1291 & 0.1317 \\ 0.1277 & 0.1277 & 0.1199 & 0.1225 & 0.1271 & 0.1159 & 0.1249 & 0.1151 & 0.1169 & 0.1226 \\ 0.1093 & 0.1128 & 0.1004 & 0.0987 & 0.0674 & 0.0030 & 0.0910 & 0.0940 & 0.0945 & 0.1025 \\ 0.0858 & 0.1030 & 0.0718 & 0.0327 & 0.1697 & 0.1824 & 0.0428 & 0.0688 & 0.0696 & 0.0769 \\ 0.0782 & 0.0940 & 0.0189 & 0.2442 & 0.5596 & 0.1854 & 0.0020 & 0.0363 & 0.0431 & 0.0569 \\ 0.0734 & 0.0860 & 0.0362 & 0.7946 & 0.4794 & 0.0309 & 0.0681 & 0.0175 & 0.1121 & 0.0118 \\ 0.0813 & 0.0688 & 0.1864 & 0.7119 & 0.0836 & 0.0175 & 0.1202 & 0.1905 & 0.6706 & 0.0388 \\ 0.0259 & 0.0008 & 0.6363 & 0.2565 & 0.0408 & 0.0375 & 0.0595 & 0.2268 & 0.2290 & 0.3428 \\ 0.2587 & 0.0225 & 1.0000 & 0.0738 & 0.0824 & 0.0595 & 0.0581 & 0.2268 & 0.2290 & 0.3428 \\ 0.4288 & 0.0373 & 0.4358 & 0.0234 & 0.0700 & 0.0799 & 0.0674 & 0.0126 & 0.0706 & 0.9301 \\ 0.1878 & 0.0879 & 0.0200 & 0.0521 & 0.0695 & 0.0787 & 0.0716 & 0.0774 & 0.0531 & 0.8576 \\ 0.0050 & 0.0990 & 0.0679 & 0.0745 & 0.0797 & 0.0824 & 0.0816 & 0.0890 & 0.0181 & 0.2375 \\ 0.0808 & 0.1143 & 0.0979 & 0.0894 & 0.0931 & 0.0956 & 0.0949 & 0.0913 & 0.0608 & 0.0144 \\ 0.1062 & 0.1169 & 0.1081 & 0.1066 & 0.1086 & 0.1101 & 0.1073 & 0.1023 & 0.0928 & 0.0910 \\ 0.1211 & 0.1252 & 0.1229 & 0.1212 & 0.1204 & 0.1202 & 0.1193 & 0.1175 & 0.1157 & 0.1162 \\ 0.1334 & 0.1356 & 0.1326 & 0.1293 & 0.1294 & 0.1298 & 0.1281 & 0.1267 & 0.1276 & 0.1296 \end{Bmatrix}$$

$$A = \begin{Bmatrix} 0.7062 & 0.0246 & 0.2498 & 0.2403 & 0.3491 \\ 0.0065 & 0.5582 & 0.2910 & 0.3377 & 0.1577 \\ 0.1001 & 0.0760 & 0.2306 & 0.0855 & 0.2049 \\ 0.0286 & 0.3225 & 0.1745 & 0.2540 & 0.2775 \\ 0.1586 & 0.0187 & 0.0540 & 0.0825 & 0.0108 \end{Bmatrix}$$

$$B = \begin{Bmatrix} 0 & 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 & 0 \\ 0.5125 & 0.0071 & 0.5441 & 0.2204 & 0.3991 \\ 0.0442 & 0.6583 & 0.1059 & 0.5518 & 0.1553 \\ 0.0008 & 0.2441 & 0.0769 & 0.0667 & 0.0043 \\ 0 & 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 & 0 \\ 0.0113 & 0.0632 & 0.2646 & 0.0761 & 0.0885 \\ 0.4313 & 0.0272 & 0.0086 & 0.0851 & 0.3528 \end{Bmatrix}$$

$$\pi(1) = \begin{Bmatrix} 0.8588 \\ 0.0001 \\ 0.0465 \\ 0.0288 \\ 0.0659 \end{Bmatrix}$$

Parameters of the Neural Network used in this study:

(a) Input weights and bias

$$IW = \begin{Bmatrix} +2.1194 & -1.1170 & -4.0348 & -1.6047 & +47739 \\ -0.0340 & -0.2151 & -0.2349 & +0.0914 & +0.1294 \\ +0.0122 & +0.2599 & +0.2901 & -0.0839 & -0.2720 \\ +0.0577 & +0.1106 & -0.0052 & +0.1307 & +0.3009 \\ +0.0592 & +0.1101 & -0.0264 & +0.1746 & +0.2996 \\ -5.8607 & +4.9304 & -3.9845 & -0.5050 & +3.5589 \\ +0.0065 & -0.0035 & -0.0235 & +0.0014 & +0.0244 \\ -0.0703 & -0.6206 & -0.0929 & -0.6700 & +1.4467 \\ -0.4069 & -0.5890 & -0.2088 & -0.0171 & +0.0225 \\ -0.1949 & -0.6615 & +0.5355 & -0.1800 & +0.7955 \end{Bmatrix}$$

$$b = \begin{Bmatrix} -21.5179 \\ +18.6524 \\ -16.0132 \\ -35.5309 \\ -36.4885 \\ -3.0347 \\ -0.2754 \\ -9.4646 \\ +71.2826 \\ -1.9331 \end{Bmatrix}$$

(b) Layer weights and bias

$$LW = \begin{Bmatrix} -0.1345 & +2.4667 & +2.1685 & +7.5245 & -6.8483 & -0.0376 & +1.8678 & -0.1589 & -0.1850 & -1.8737 \\ +0.0067 & 0.00000 & -0.0135 & +0.0611 & -0.0414 & +0.7837 & -0.2360 & +0.0112 & +1.2802 & +0.4338 \\ +0.0719 & -2.3709 & -1.4289 & -5.9347 & +5.2860 & -0.0492 & -0.6740 & +0.2157 & -0.4422 & +0.8918 \end{Bmatrix}$$

$$b = \begin{Bmatrix} +2.1656 \\ +1.6306 \\ 0.0072 \end{Bmatrix}$$

Die "Berichte zur Polar- und Meeresforschung"
(ISSN 1866-3192) werden beginnend mit dem Heft Nr. 377 (2000) in Fortsetzung der früheren **"Berichte zur Polarforschung** (Heft 1-376, von 1982 bis 2000; ISSN 0176 - 5027) herausgegeben. Ein Verzeichnis aller Hefte beider Reihen befindet sich im Internet in der Ablage des electronic Information Center des AWI (**ePIC**) unter der Adresse http://epic.awi.de. Man wähle auf der rechten Seite des Fensters "Reports on Polar- and Marine Research". Dann kommt eine Liste der Publikationen und ihrer online-Verfügbarkeit in alphabetischer Reihenfolge (nach Autoren) innerhalb der absteigenden chronologischen Reihenfolge der Jahrgänge.

To generate a list of all 'Reports' past issues, use the following URL: http://epic.awi.de and select the right frame: Browse. Click on "Reports on Polar and Marine Research". A chronological list in declining order, author names alphabetical, will be produced. If available, pdf files will be shown for open access download.

Verzeichnis der zuletzt erschienenen Hefte:

Heft-Nr. 569/2008 — "The Expedition ANTARKTIS-XXIII/8 of the Research Vessel 'Polarstern' in 2006/20067", edited by Julian Gutt.

Heft-Nr. 570/2008 — "The Expedition ARKTIS-XXI/1 a and b of the Research Vessel 'Polarstern' in 2005", edited by Gereon Budéus, Eberhard Fahrbach and Peter Lemke.

Heft-Nr. 571/2008 — "The Antarctic ecosystem of Potter Cove, King-George Island (Isla 25 de Mayo). Synopsis of research performed 1999-2006 at the Dallmann Laboratory and Jubany Station", edited by Christian Wiencke, Gustavo A. Ferreyra, Doris Abele and Sergio Marenssi.

Heft-Nr. 572/2008 — "Climatic and hydrographic variability in the late Holocene Skagerrak as deduced from benthic foraminiferal proxies", by Sylvia Brückner.

Heft-Nr. 573/2008 — "Reactions on surfaces of frozen water: Importance of surface reactions for the distribution of reactive compounds in the atmosphere", by Hans-Werner Jacobi.

Heft-Nr. 574/2008 — "The South Atlantic Expedition ANT-XXIII/5 of the Research Vessel 'Polarstern' in 2006", edited by Wilfried Jokat.

Heft-Nr. 575/2008 — "The Expedition ANTARKTIS-XXIII/10 of the Research Vessel 'Polarstern' in 2007", edited by Andreas Macke.

Heft-Nr. 576/2008 — "The 6th Annual Arctic Coastal Dynamics (ACD) Workshop, October 22-26, 2006, Groningen, Netherlands", edited by Pier Paul Overduin and Nicole Couture.

Heft-Nr. 577/2008 — "Korrelation von Gravimetrie und Bathymetrie zur geologischen Interpretation der Eltanin-Impaktstruktur im Südpazifik", von Ralf Krocker.

Heft-Nr. 578/2008 — "Benthic organic carbon fluxes in the Southern Ocean: regional differences and links to surface primary production and carbon export", by Oliver Sachs.

Heft-Nr. 579/2008 — "The Expedition ARKTIS-XXII/2 of the Research Vessel 'Polarstern' in 2007", edited by Ursula Schauer.

Heft-Nr. 580/2008 — "The Expedition ANTARKTIS-XXIII/6 of the Research Vessel 'Polarstern' in 2006", edited by Ulrich Bathmann

Heft-Nr. 581/2008 — "The Expedition of the Research Vessel 'Polarstern' to the Antarctic in 2003 (ANT-XX/3)", edited by Otto Schrems

Heft-Nr. 582/2008 — "Automated passive acoustic detection, localization and identification of leopard seals: from hydro-acoustic technology to leopard seal ecology", by Holger Klinck