



Practical implementation of real-time fish classification from acoustic broadband echo sounder data – RealFishEcho progress report

Authors: B. Berges, S. Sakinan, E. van Helmond

Wageningen University &
Research rapport C062/18

Practical implementation of real-time fish classification from acoustic broadband echo sounder data - RealFishEcho progress report

Year 2 – June 2018

Author(s): B. Berges, S. Sakinan, E. van Helmond

Publication date: September

Wageningen Marine Research
IJmuiden, September 2018

CONFIDENTIAL NO

Wageningen Marine Research report C062/18

B. Berges, S. Sakinan & E. van Helmond, 2018. Practical implementation of real-time fish classification from acoustic broadband echo sounder data – RealFishEcho progress report. Wageningen, Wageningen Marine Research (University & Research centre), Wageningen Marine Research report C062/18. 42 pp.

Client: RVZ
Attn.: Martin Pastoors
Louis Braillelaan 80
2719 EK ZOETERMEER
The Netherlands

This report can be downloaded for free from <https://doi.org/10.18174/458935>
Wageningen Marine Research provides no printed copies of reports

Wageningen Marine Research is ISO 9001:2008 certified.

© 2016 Wageningen Marine Research Wageningen UR

Wageningen Marine Research
institute of Stichting Wageningen
Research is registered in the Dutch
traderecord nr. 09098104,
BTW nr. NL 806511618

The Management of Wageningen Marine Research is not responsible for resulting damage, as well as for damage resulting from the application of results or research obtained by Wageningen Marine Research, its clients or any claims related to the application of information found within its research. This report has been made on the request of the client and is wholly the client's property. This report may not be reproduced and/or published partially or in its entirety without the express written consent of the client.

A_4_3_2 V27draft

Contents

Summary	5
1 Introduction	6
1.1 Key technical points	8
1.1.1 EK80 configuration	8
1.1.2 ECHO software interfacing	9
2 WP1: System calibration	11
2.1 Introduction	11
2.2 Work undertaken and current state of WP1	12
2.2.1 Alida (SCH06)	12
2.2.2 Afrika (SCH24)	14
2.2.3 Willem van der Zwan (SCH302)	15
2.2.4 Tridens	15
2.3 Work planned for the following reporting period	15
3 WP2: Data collection	16
3.1 Introduction	16
3.2 Work undertaken and current state of WP2	17
3.3 Work planned for the following reporting period	18
4 WP3: Data processing & analysis	19
4.1 Introduction	19
4.2 Work undertaken and current state of WP3	19
4.2.1 Automatic processing and data base generation	19
4.2.2 Analysis	20
4.3 Work planned for the following reporting period	24
5 WP4: Demonstrator software development	25
5.1 Introduction	25
5.2 Work undertaken and current state of WP4	25
5.3 Work planned for the following reporting period	27
6 WP5: Management, communication & reporting	28
6.1 Meetings	28
6.2 Miscellaneous	28
7 Quality Assurance	29
Annex I: Automatic reporting for fishing trip overview	30
Annex II	34
Introduction	34
Method	34
Description of the workflow	35
Pre-processing	35
Dynamic Time Warping (DTW)	35
Visualization 1- Examination within each schools (Figure II.2)	36

Visualization 2- Examination between schools (Figure II.6)	36
Visualization 3- Spectra averages (Figures II.3-II.5)	36
Results and discussion	39
References	40
Justification	41

Summary

The EU has by means of new policy restricted the discarding of fish at sea. As a result, the fishing industry now requires improved methods to identify fish species and size before the catch process begins. With the introduction of robust broadband echosounders (i.e. downward looking sonar) on the market, active acoustic data now have the potential for improved target discrimination for both type and size (i.e. determine fish species and size).

The aim of this project is to develop methods for fish classification and size estimation based on data collected on board commercial vessel equipped with broadband echosounders. These methods will then be implemented into a software package for near-real time viewing. Such a software package would help skippers to take better-informed decisions while fishing.

This project runs over three years (June 2016 to June 2019) in collaboration between Wageningen Marine Research (WMR), Redersvereniging voor de Zeevisserij (RVZ) and TNO. Progress made in the second year of the project is presented in this report.

Through the course of the reporting period, substantial amount of data have been collected and analysis show promising results. A test version of the classification software is planned to be deployed in the next phase of the project.



Europese Unie, Europees Fonds voor Maritieme Zaken en Visserij

1 Introduction

The EU has by means of new policy restricted the discarding of fish at sea. As a result, the fishing industry now requires improved methods to identify fish species and size before the catch process begins. With the introduction of robust broadband echosounder (i.e. downward looking sonar) on the market, active acoustic data now have the potential for improved target discrimination for both type and size (i.e. determine fish species and size).

The aim of this project is to develop methods for fish classification and size estimation based on data collected on board commercial vessels equipped with broadband echosounders. These methods will then be implemented into a software package for near-real time viewing. Such a software package would help skippers to take better-informed decisions while fishing.

This project runs over three years (June 2016 to June 2019) and is divided into 5 work packages (WP) which are then divided into tasks (T) (Table 1). The breakdown of these tasks is given in Table 1. The associated timeline and milestones for the different tasks is given in Table 2, including reporting milestones.

The project runs in collaboration between Wageningen Marine Research (WMR), Redersvereniging voor de Zeevisserij (RVZ) and TNO. The development of the different tasks within this project is based on data collected by three commercial freezer trawlers:

- Afrika (SCH24)
- Alida (SCH06)
- Willem van der Zwan (SCH302)

In addition, research vessel "R/V Tridens" provides data of opportunity collected during acoustic surveys conducted by WMR.

Each vessel in this study is using an EK80 system running in Frequency Modulated mode (FM, named "EK80 FM" further in this report), i.e. broadband. This is different from the EK80 system running in Continuous Wave mode (CW, named "EK80 CW" further in this report), i.e. narrowband. The EK80 system is provided by SIMRAD. Three frequency channels are used in the project: 70 kHz, 120 kHz and 200 kHz.

This document is a progress report for the second year of the project (M13, Table 2) and gives an update for each WP. In the following sections, project tasks are described with occasional text highlighted in grey (■). This denotes the project participant responsible for the task. Results for the first year of the project was reported in the first progress report (M6, Table 2) [1].

Table 1: breakdown of the WPs and tasks within the realFishEcho project.

Item #	Tasks
WP1	System calibration
T1.1	Calibration in the field
WP2	Data collection
T2.1	fishing trips (mackerel, horse mackerel, herring, sprat)
T2.2	research trips (mackerel, herring, sprat)
WP3	Data processing & analysis
T3.1	fishing trips (mackerel, horse mackerel, herring, sprat)
T3.2	research trip (mackerel, herring, sprat)
T3.3	database setup
T3.4	species identification algorithm improvement
T3.5	fish length estimation method development
WP4	Demonstrator software development
T4.1	first demonstrator software version (GUI 1): pre-processing, imaging, noise reduction, school detection, species classification
T4.2	second demonstrator software version (GUI 2): implementation of updated species classification and length estimation
T4.3	installation, test & evaluation GUI 1
T4.4	installation, test & evaluation GUI 2
WP5	Management, communication & reporting
T5.1	Project progress monitoring & reporting
T5.2	Project management and communication
T5.3	Report and presentation of results

Table 2: Milestones associated with the different tasks (Table 1) of the realFishEcho project.

Year	Date	WP	Task	Milestone #	Milestone
2016	Dec 2016	WP4	T4.1	M1	implementation first GUI version
2017	Jan 2017	WP3	T3.3	M2	Development of database
	Feb 2017	WP1	T1.1	M3	2 vessels calibrated
	Apr 2017	WP2	T2.1/2.2	M4	Dataset year 1 ready
	May 2017	WP3	T3.1/3.2	M5	Data analysis year 1
	May 2017	WP5	T5.1	M6	Interim report 1
	Jul 2017	WP4	T4.3	M7	Implementation GUI on board of vessels
	Aug 2017	WP3	T3.4	M8	Improving species classification algorithm
2018	Feb 2018	WP3	T3.5	M9	Method for length estimation
	Apr 2018	WP2	T2.1/2.2	M10	Dataset year 2 ready
	May 2018	WP3	T3.1/3.2	M11	Data analysis year 2
	May 2018	WP4	T4.2	M12	Implementation GUI version 2
	May 2018	WP5	T5.1	M13	Interim report 2
	Oct 2018	WP4	T4.4	M15	Test and evaluation GUI on board
	Aug 2018	WP5	T5.1	M14	Interim report 3
2019	Mar 2019	WP2	T2.1/2.2	M16	Dataset year 3 ready
	Mar 2019	WP4	T4.4	M17	Installation final GUI version on board
	Apr 2019	WP3	T3.1/3.2	M18	Data analysis year 3
	May 2019	WP5	T5.3	M19	Final report

1.1 Key technical points

In the first progress report [1], a description of echosounder working principles was given in order to put into context the need for calibration. In this report, a description of the operational use of the EK80 FM will be given, followed by the description of interfacing of the software development within this project (ECHO software).

1.1.1 EK80 configuration

The EK80 system is the successor of the EK60 system. The latter has been used for decades during acoustic surveys and is known for its consistency and robustness. While the EK80 works similarly to the EK60 in Continuous wave mode (CW, narrow band), it also offers broadband capabilities when operating in Frequency Modulated mode (FM, broadband). EK80 FM offers advantages in term of target discrimination (by resolving larger frequency bands) and depth resolution (through matched filtering). Whilst scientific methods for the processing of narrow band echosounders is well established [2], the use of broadband echosounders is relatively new in the field of fisheries acoustics though it has been used in the past [3].

The working principle of the EK80 is exemplified in Figure 1 for a single frequency channel. An acoustic transducer is coupled to the EK80 system (i.e. transceiver) for: (1) transmission of input signals that is further converted into acoustic waves, and (2) reception of the signals that result from the measurements of the scattered acoustic waves. The functioning of the EK80 relies on firmware updates. Then, the EK80 system interfaces with the EK80 or the ES80 software (installed on a separate computer). Both software allow:

- The visualization of data
- The configuration of the system (e.g. pulse length, ramping, frequency range)
- The logging of raw data in the *.RAW format

The data collected also uses calibration values in order for the echosounder to derive absolute acoustic intensity levels. This is necessary in order to compare results from different vessels. Further explanation on the need for calibration is explained in Section 2. Calibration is obtained after dedicated trials where measured acoustic intensity levels can be compared with those from calibration spheres (theoretically known acoustic response) [3, 4]. The EK80 software has a specific Graphical User Interface (GUI) where the state and quality of the calibration can be monitored. The result of a calibration trial is:

- A set of values that are stored within the EK80 software but also in output files.
- An update of beam at the level of the EK80 transducer.

The EK80 software is a version that offers more possibilities in term of customization (pulse length, pulse ramping, frequency range). At the contrary, the settings from the ES80 software are limited. This hampers its potential for collection of data within this project. More specifically the following can only be achieved with the EK80 software:

- Maximizing the frequency range
- Minimizing pulse ramping
- Setting pulse length

For these reasons, only data collected using the EK80 software is used in this project. However, because the ES80 is optimized for visualization, this software is preferred by skippers on board commercial vessels. It is therefore beneficial to explore the use of data collected with the ES80 for classification purposes. This will be covered by a subtask within the last year of this project.

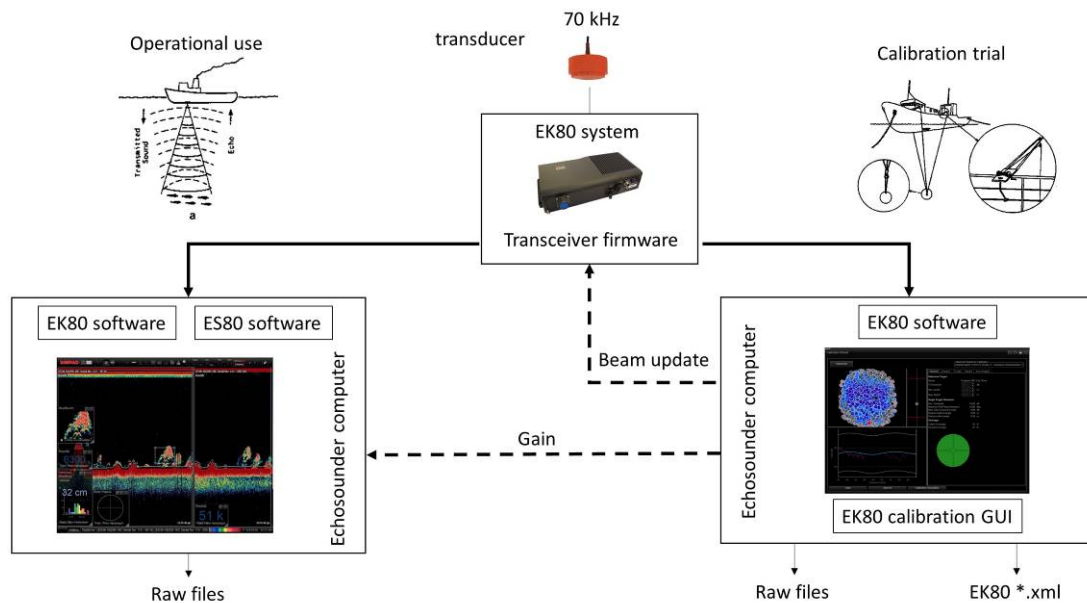


Figure 1: working principles of the EK80 system.

1.1.2 ECHO software interfacing

The ECHO software developed within this project (WP4, Table 1) is aimed at providing skippers from commercial vessel with species identification and length estimation based on the acoustic records. The ECHO software is aimed at classifying the following fish species: Herring, Horse Mackerel, Mackerel and possibly Boarfish or Sprat. Classification algorithms are in place within the ECHO software and both are using a machine learning approach (see Section 4.2.3 and [6]). The training of the algorithms relies on data collected by the commercial vessels, including calibration. Calibration is particularly important to:

- Account for the echosounder change through time
- Cross-vessel application of the software.

A schematic overview of the implementation of the ECHO software is given in Figure 2. The EK80 software operates from a specific computer and controls the different transducers. It also stores the raw data (*.RAW) on a network drive. Independently from EK80 software, the ECHO software reads the raw data and performs a near real time classification. The ECHO software also use independent files for the algorithm settings and the calibration. To date, the ECHO software can only work alongside the EK80 software and not the ES80.

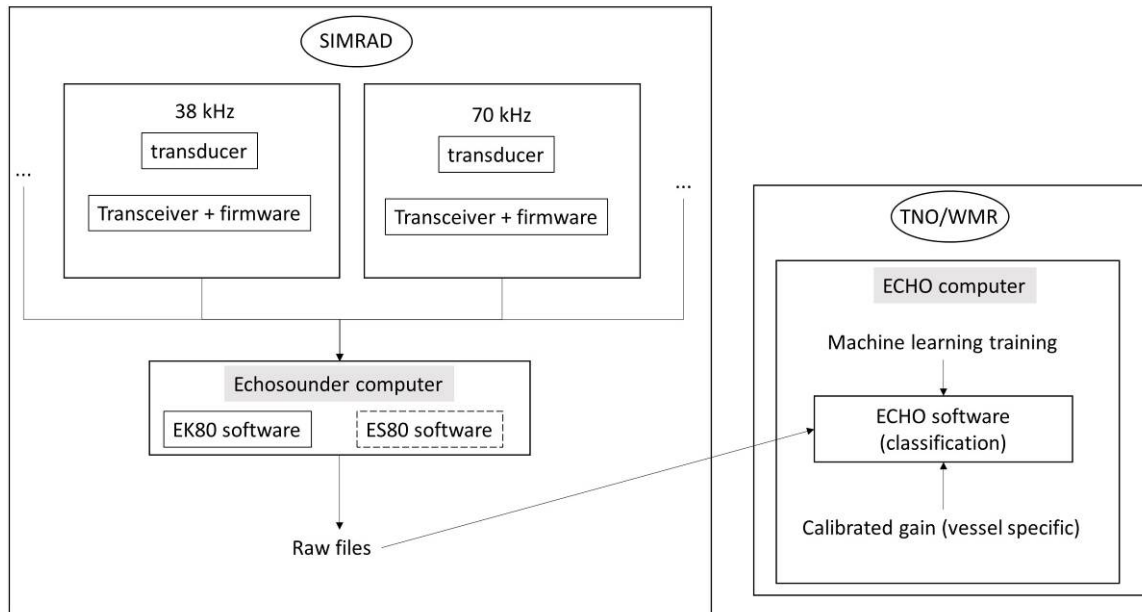


Figure 2: interfacing of the ECHO software with the EK80 system

2 WP1: System calibration

2.1 Introduction

As described in the first progress report [1], calibration is a necessary step in order to convert raw acoustic data into absolute level of acoustic intensity. The procedure employed in this project is the standard sphere calibration method [4, 5]. The calibration data are collected using the EK80 software calibration module (Section 1.1.1) that allows the operator to monitor the state of the calibration. This includes monitoring of the data quality and tracking of the sphere within the beam. The collected data consist of two types of files: raw acoustic records and an output calibration file (*.xml, Section 1.1.1). Through the course of this project, MATLAB code was developed in order to revisit the raw calibration files. This software was developed because:

- While the EK80 software produces calibration outputs, one cannot control the processing and derivation calibration corrections.
- The EK80 calibration output (*.xml files) only contains limited information on the calibration trial and therefore limits the analysis.
- The current EK80 software is unable to read past data (i.e. 2014/2015 data sets).
- It is likely that the EK80 calibration software will be updated in the future which may make it impossible to apply classifiers to data sets calibrated in different ways.

In practice, several calibration spheres are suspended in the water but the EK80 calibration software tracks the location and data acquisition of a single sphere in the water column. This is done through the processing of sphere detection in a specified depth interval. While other spheres in the water are not being focused on during this procedure, the raw data can be revisited in order to derive additional calibration gains. The final results of a calibration trial then consists of gains at frequencies computed for each sphere for each EK80 FM channel (i.e. frequency). Because the gains are computed over a large frequency band, some regions of the spectrum are not useable because of drops in the spectrum. These dips are specific to each sphere size. This is exemplified in Figure 3. For frequencies within the dip regions, gain can be calculated through: (1) interpolation; (2) combining different spheres.

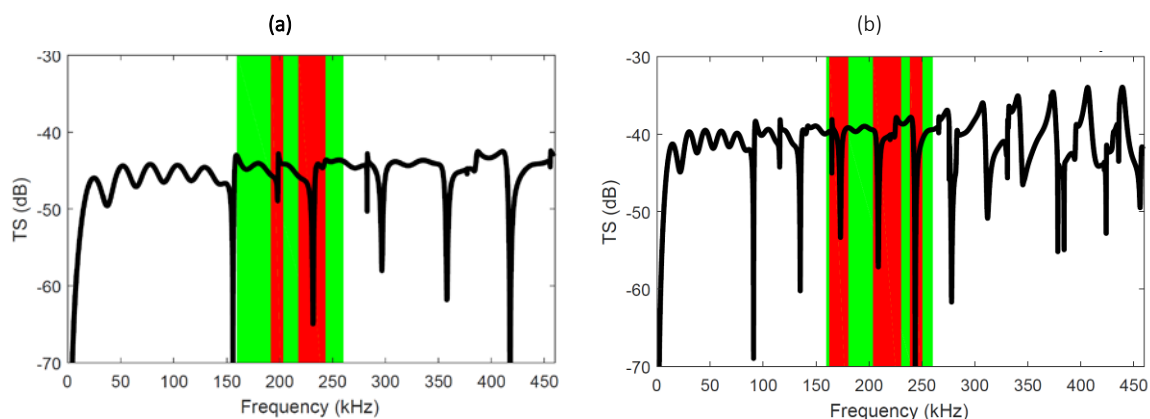


Figure 3: theoretical target strength (TS) vs frequency. This analytically inferred target strength is used to compare measurements of the EK80 FM against and in turn determine calibration gains at different frequencies. The solid black line is the theoretical target strength. The shaded area represents the working range of the EK80 FM 120 kHz channel. The subsequent green and red areas are the useable and non-useable frequency ranges for calibration respectively. (a) 22 mm WC calibration sphere. (b) 38.1 mm WC calibration sphere. Results shown are from the calibration trial performed by the Alida (SCH06) on September 2017.

A summary of the different calibration trials performed on-board the Alida (SCH06), Afrika (SCH24) and Tridens and available to this project are shown in Table 3. For each calibration trial, several size spheres could be used to derive subsequent calibration gains. No calibration was performed for the EK80 FM on-board the Willem van der Zwan (SCH302).

Table 3: summary of calibration data available to date. For each trial, different sphere types were used to derive calibration gains. This was done using the EK80 software and the self-developed software ("manual analysis" in table). While all spheres can be used to derive calibration gains, the highest frequency channels (120 kHz and 200 kHz) are best calibrated using the 25, 22 or 15 mm WC spheres.

date	data set size (GB)	analysis EK80 software	manual analysis	frequencies	Sphere type
SCH06 - Alida					
Sep/17	128	yes	yes	70 kHz 120 kHz 200 kHz	25 mm WC 38.1 mm WC
Aug/16	53	yes	yes	70 kHz 120 kHz 200 kHz	22 mm WC 25 mm WC 38.1 mm WC
Sep/15	8	no	yes	70 kHz 200 kHz 200 kHz	15 mm WC 22 mm WC 38.1 mm WC
SCH24 - Afrika					
Aug/16	24	yes	yes	38 kHz 70 kHz 120 kHz 200 kHz 333 kHz	22 mm WC 25 mm WC 38.1 mm WC
Tridens					
Jun/17	9.3	yes	yes	70 kHz 120 kHz 200 kHz	22 mm WC 25 mm WC 38.1 mm WC

2.2 Work undertaken and current state of WP1

General tasks undertaken within WP1 are:

- Development of the calibration software for the processing of calibration records (WMR). This includes automatic processing and adapting the code to account for disparities between raw data formats (e.g., EK80 FM data from 2015 have an outdated format).
- Development of workflow (WMR) for: (1) systematic analysis of the calibration results; (2) management of calibration data outputs.

2.2.1 Alida (SCH06)

- Analysis of calibration trials from September 2017, August 2016 and September 2015 (WMR).
- Comparison analysis of September 2017, August 2016 and September 2015 trials (WMR & TNO). This includes the assessment of the data quality through the consistency of the derived gains, both in terms of frequency and location in the beam. Example plots are shown for the 70 kHz channel in Figure 4 (gain and subsequent standard deviation versus frequency). Figure 5 and Figure 5 are example plots comparing the different calibration trials performed by the Alida (SCH06) (Table 3). Figure 4 shows the frequency dependent standard deviation and gains (i.e. frequency distribution of result quality) while Figure 5 shows the target detections and the standard deviation of the results in different sectors (i.e. spatial distribution of result quality). While the 70 kHz channel is not affected the issue of sphere TS dips (Section 2.1), the calibration of the 120 kHz and 200 kHz are affected for all spheres used in this study. Because discrepancies were observed between different spheres, it was decided to use the 38.1 mm sphere (sphere present across all calibration trials) to derive the final gains. The gains in the dip regions were interpolated. The final gains are shown in Figure 6. The differences between each data set can be greater than 1 dB in specific frequency regions. These discrepancies are likely due to the specificity of each data set (e.g. high noise level, change in setting, change in transducer firmware) and need to be investigated further.

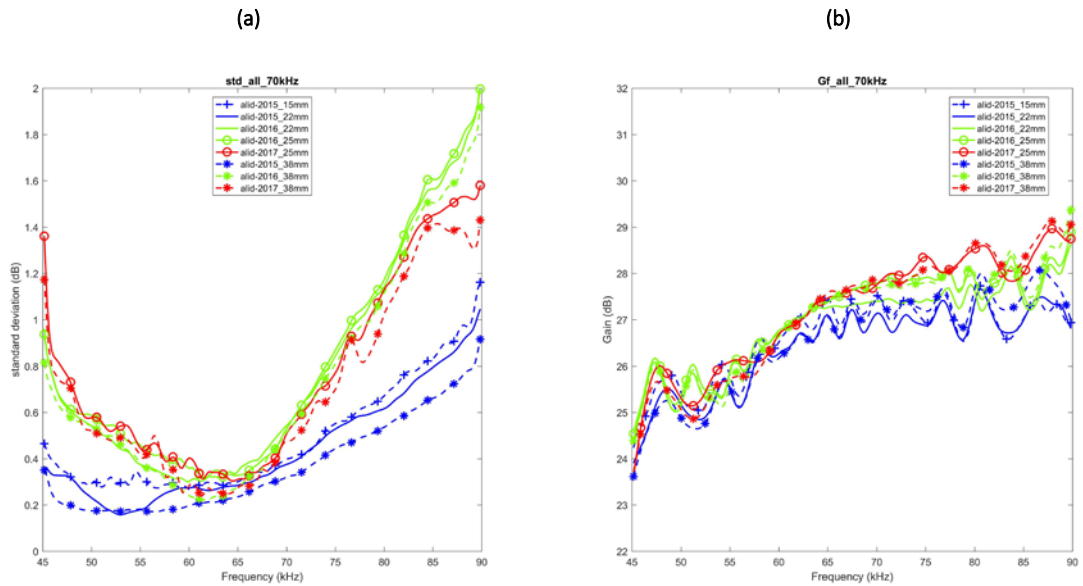


Figure 4: comparison of the results for the 70 kHz channel from different calibration trials performed by the Alida (SCH06) (Table 3). (a) Standard deviation in gain in dB of the spheres detections. (b) Resulting gain in dB. The different trials are differentiated by colour: blue (09/2015); green (08/2016) and red (09/2017). For each trial, acoustic records for different spheres were analysed. Each sphere type has a different line type and marker symbol: solid line with no marker symbols (22 mm WC); dashed line with + marker symbols (15 mm WC); solid line with O marker symbol (25 mm WC); dashed line with * marker symbol (38.1 mm WC).

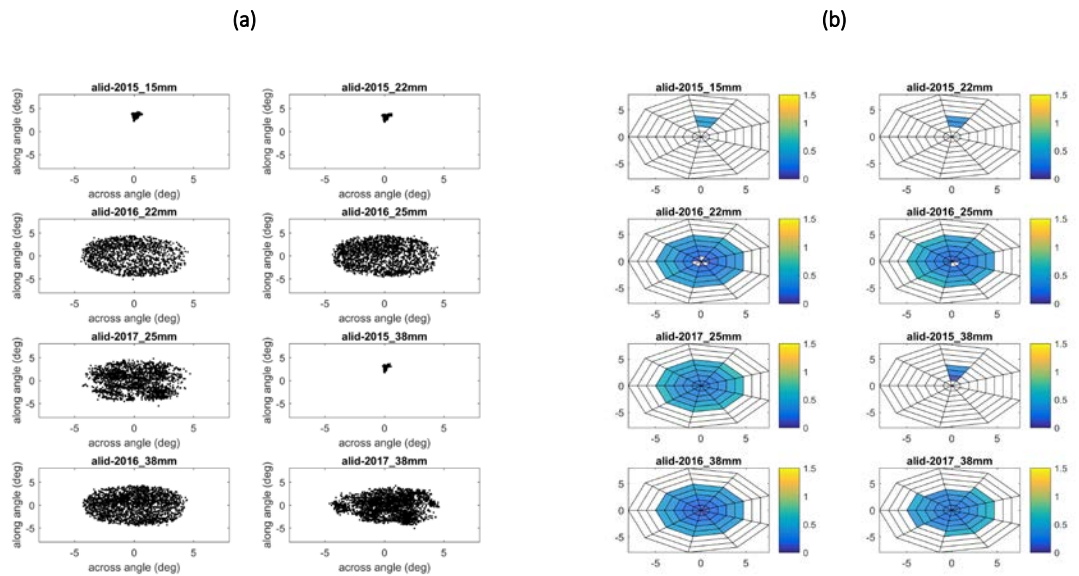


Figure 5: diagnostic of calibration results in different beam sectors for the 70 kHz channel from different calibration trials performed by the Alida (SCH06) (Table 3). Each subplot is the results from a calibration trial for a given sphere type. (a) Location of sphere detections, with each dot representing a detection. (b) Standard deviation (across the whole frequency range) in gain in dB in different sectors.

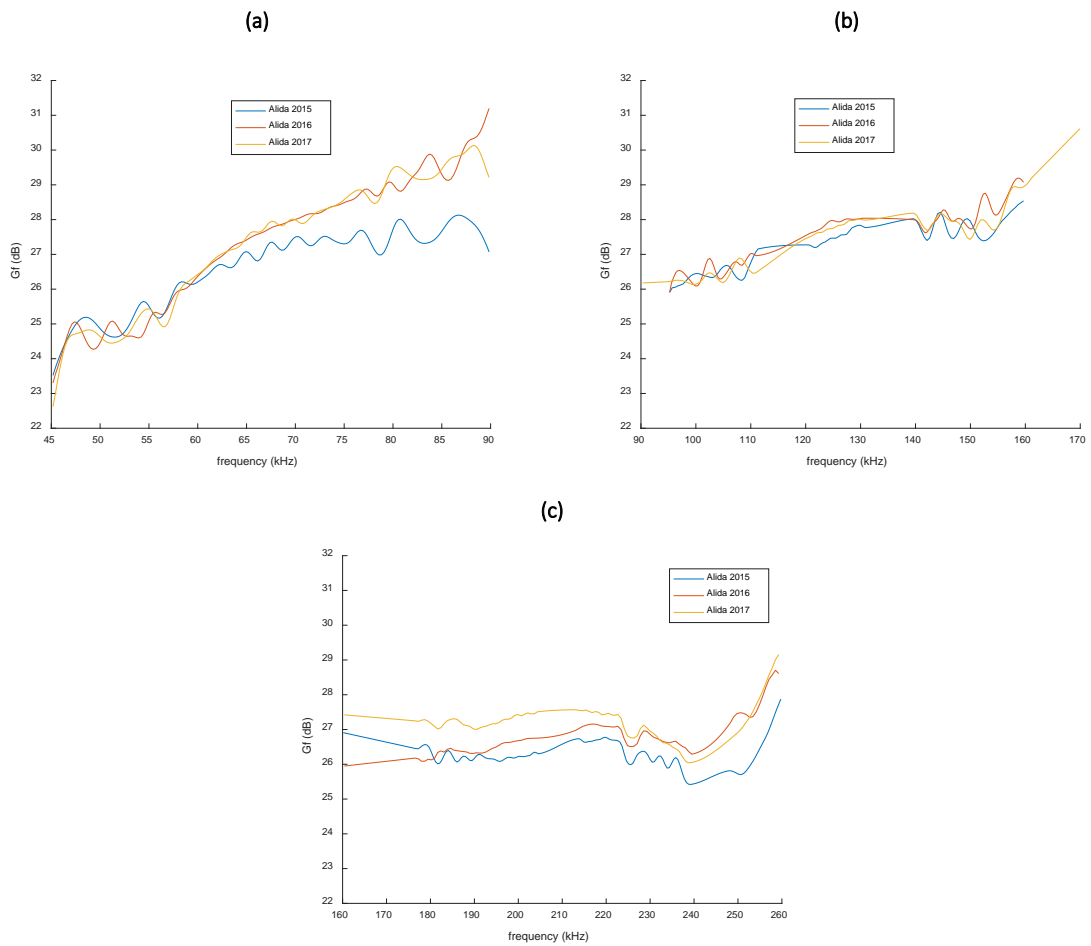


Figure 6: final gains (G_f) from the different calibration trials performed by the Alida (SCH06). The gains were computed using the 38.1 mm calibration sphere for the three calibration trials (09/2015, 08/2016, 09/2017, Table 3). (a) 70 kHz channel. (b) 120 kHz channel. (c) 200 kHz channel.

2.2.2 Afrika (SCH24)

- Analysis of calibration trials from August 2016 (WMR, Figure 7). Only one data set is available for the Afrika (SCH24) and a comparison analysis cannot be performed yet.

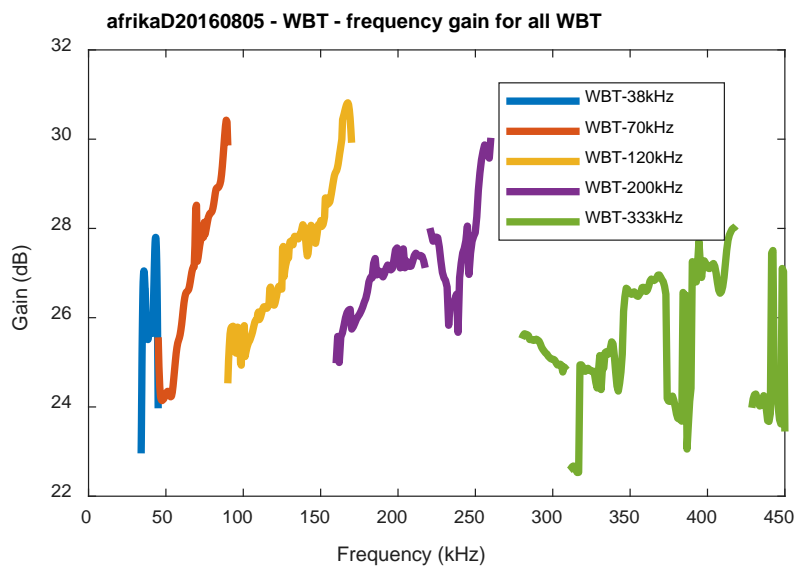


Figure 7: results from the Afrika (SCH24) calibration trial from August 2016. Gain levels versus frequency after analysis using the EK80 software. The gains are combining results from different spheres (38.1 mm WC, 22 mm WC, 25 mm WC).

2.2.3 Willem van der Zwan (SCH302)

- No broadband calibrations was performed in the reporting period. It is particularly hard for this vessel to find dedicated time to calibration because this vessel is mainly involved in the collection of narrow band data for another project (SEAT project¹)

2.2.4 Tridens

- Analysis of calibration trials from June 2017 (WMR, Figure 8).

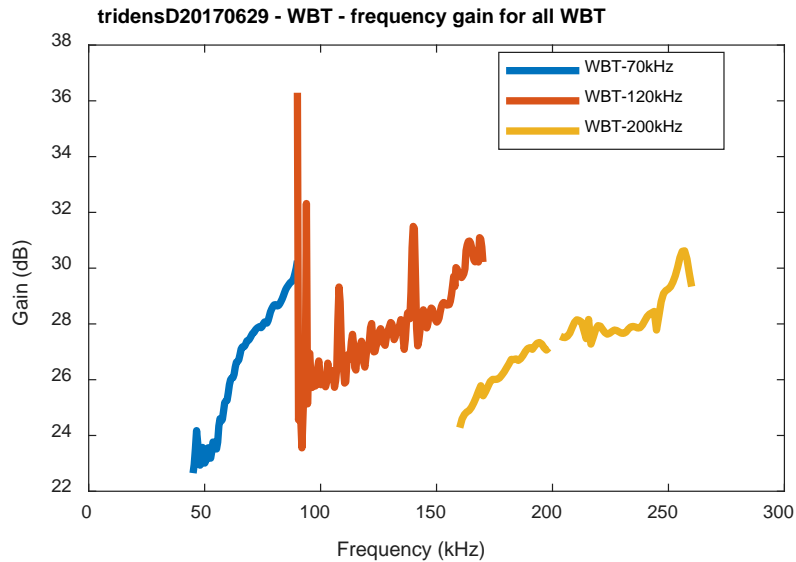


Figure 8: results from the Tridens calibration trial from June 2017. Gain levels versus frequency after analysis using the EK80 software. The gains are combining results from different spheres (38.1 mm WC, 22 mm WC, 25 mm WC).

2.3 Work planned for the following reporting period

- Investigate discrepancies between calibration results for different sphere types (WMR).
- Calibration to be performed in 2018 for the Afrika (SCH24) and Alida (SCH06).

¹ SEAT: School Exploration and Analysis Tool. See: *Sustainovate 2017 Improved Selectivity of Small Pelagics iun the North Sea & North Atlantic Using SEAT*

3 WP2: Data collection

3.1 Introduction

During the course of this project, EK80 FM data were collected on board three commercial vessels; Afrika (SCH24), Alida (SCH6) and Willem van der Zwan (SCH302). Though the focus is on commercial vessels, data from research vessel "R/V Tridens" is also available. The data from the commercial vessels are being processed and converted into a specific database format for analysis and development of the classification algorithms (Section 4).

The vessels involved in the data collection are equipped with different setups in terms of frequency channels, with a minimum of 3 frequency channels (70, 120 and 200 kHz) up to 5 channels (38 kHz, 70 kHz, 120 kHz, 200 kHz, 333 kHz). Because the collected data from all vessels contain at least data for the 70, 120 and 200 kHz channels, only these frequency channels are being used as a basis for the classification algorithms. For each data set, trawl catch are also being recorded in order to provide ground-truth for the acoustic data. Since December 2017, this also includes fish length measurements. An overview of data available to the project to date is given in Table 4.

As reported in the first progress report [1], the project experienced communication problems between the scientific team and the skippers of the commercial vessels. This led to a limited amount of data collected from commercial vessels. This issue has been solved with several large data sets collected by the Alida (SCH06) since December 2017.

The Afrika (SCH24) left European waters for ~12 months in June 2017 and was unable to collect data. Though data were collected by Willem van der Zwan (SCH302) in 09/2017 (Table 4), it was decided to leave this vessel dedicated to the collection of data for the SEAT project which has a similar context and data collection requirements. As a trade-off, the Alida (SCH6) focused on EK80 FM data for the realFishEcho project. This resulted in large data sets collected since December 2017 (see Table 4).

Table 4: summary of data available. For each data set, the summary consists of acoustic data and data from biological sampling (catch composition and fish length measurements). The number of fish schools detected using the automatic data processing (Section 4.1) is an indication of the size of the data available for analysis. To date, fish length measurements were only collected consistently by the Alida (SCH06) since December 2017. The data collection by the Afrika (SCH24) in January 2018 was done in Morocco with a WMR scientist on board. While these data cannot be used for species identification, they are aimed at comparison EK80 and ES80 (Section 1.1.1).

Acoustic data			Biological data		
date	data set size (GB)	number of detected schools	length data	# of hauls with length measurements	Species composition
SCH06 - Alida					
Mar/18	1800	72	yes	34	HOM, WBH, MAC
Jan/18	1160	98	yes	18	HOM
Dec/17	2100	77	yes	16	HOM, HER, SPR
Aug/17	44	3	no		HER
Oct/16	1210	52	no		HOM
Aug/16	1140	77	no		HER
Dec/14	40	12	no		MAC
Oct/14	67	18	no		HOM
Aug/14	77	15	no		HER
Tridens					
HERAS 2017	302	22	yes	Not processed	HER
HERAS 2016	83	39	yes	Not processed	HER
SCH24 - Afrika					
Dec/16	61	13	no		HER
Aug/16	83	26	no		HER
SCH302 - Willem van der Zwan					
Sep/17	579	11	no		HER

3.2 Work undertaken and current state of WP2

- Data has been collected consistently by the Alida (SCH06) around fishing operations since December 2017 (Table 4), including numerous fish length measurements.
- Data were collected by the Afrika (SCH24) with a scientist from WMR on board (Table 4, WMR). These data were collected in order to compare data acquired using the EK80 software and the ES80 software. As explained in Section 1.1.1, the ES80 uses fixed settings that are optimized for echograms display with a level of noise reduction for the skipper. On the contrary, the EK80 software is fully customizable and it displays the raw data without signal processing. This results in echograms that appear as noisier. The EK80 software is better suited for research and the ES80 proposes a more desirable display for the skippers. An assessment of the optimal processing parameters for an optimal processing and display should not be underestimated to allow a swift integration on the vessels.
- The project still needs more diversity in the data. The main species that are represented in the data to date are Herring and Horse Mackerel. The lack of lack of data for Mackerel and boarfish was raised in the first progress report [1]. The Alida (SCH06) collected mackerel data in March 2018 but it only consists of few hauls. As for Boarfish, no data were collected. For Mackerel, the lack of data is due to issues with data collection (Section 3.1). These were fixed by solely dedicating the Alida (SCH06) to the collection of EK80 FM data. However, the lack of data collection in 2017 corresponded to the Mackerel season.

3.3 Work planned for the following reporting period

- Further data collection. One needs to ensure that data are collected during Mackerel fishing season (WMR). Attention for collecting data for a fourth species (boarfish to be replaced by Sprat due to expected better data availability).
- Data collected during trial at sea (testing of the ECHO software).
- Collection of ES80 data for comparison with data collected with the EK80 software

4 WP3: Data processing & analysis

4.1 Introduction

Fish species classification is done by WMR and TNO separately and will result into two algorithms. Both algorithms is based on a machine learning approach and are described in [6]. In this section, further improvements of these algorithms are presented together with a trend analysis of the frequency responses.

4.2 Work undertaken and current state of WP3

4.2.1 Automatic processing and data base generation

Data sets collected by fishing vessels can be large (see Table 4) and include large sections without detection of fish schools. In order to reduce the data sets to their useable parts, automatic data processing routines (MATLAB) were developed (WMR). These routines draw on the development of the ECHO software (Section 5) for data pre-processing and has been used on the 14 data sets available to date (Table 4) and work consistently for different data settings (e.g. old data format, different transducer configuration). Reduced data sets consist of data stored in a predefined format as one file per detected fish school. The number of detection schools for each data set varies and is dependent on the extend of the data collection that was performed (Table 4).

Data format and organisation have been discussed and agreed upon (WMR & TNO). A workflow is now in place (Figure 9) which produces a standardized database in order for WMR and TNO to have a common basis for analysis and the development of classification and length estimation methods. This workflow will be used as soon as new data are provided.

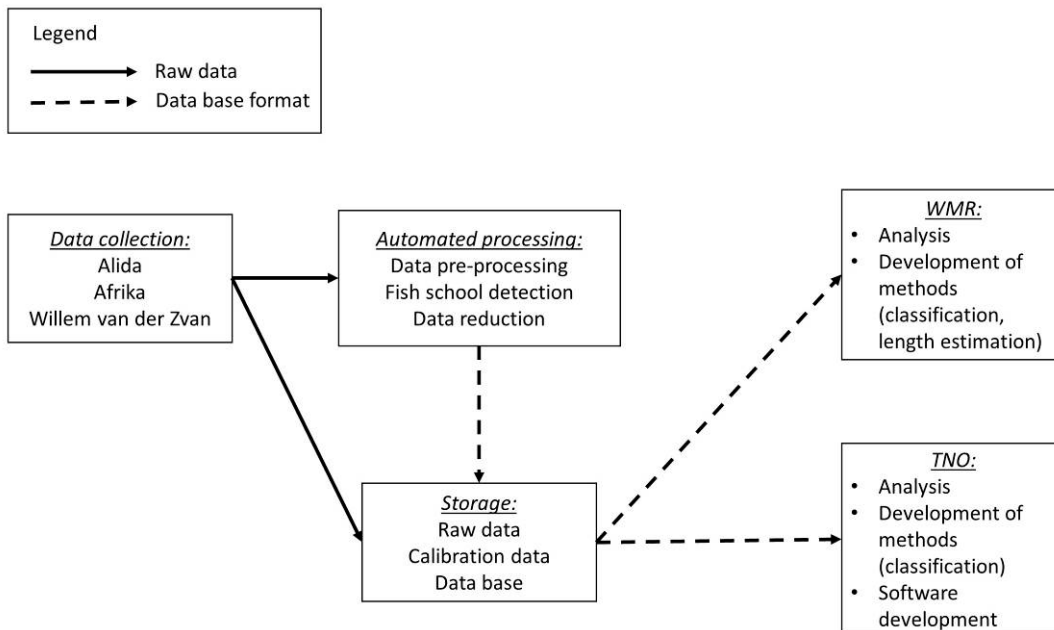


Figure 9: schematic overview of the workflow used for the management of data. Following data collection from a fishing vessel, the raw data are stored and processed using automated processing. The automated processing reduces the data to acoustic records where fish schools are detected and convert then to the data base format. The database is then used as a basis for analysis, development of methods and software.

In addition to raw data stored in a reduced form, metadata are generated. These consist of joint information on the acoustic records of fish schools (time, location, size etc...) and corresponding fishing information if any (time, catch composition, catch weight...). The metadata are used for:

- Data analysis
- Training of the classification algorithms

Drawing on the database, automatic report generation is also being developed (WMR) in order to provide prompt feedback to the skippers but also to give an overview of each data set. A description is given in Annex I.

For the analysis, each fish school is broken down into different cells (Figure 10).

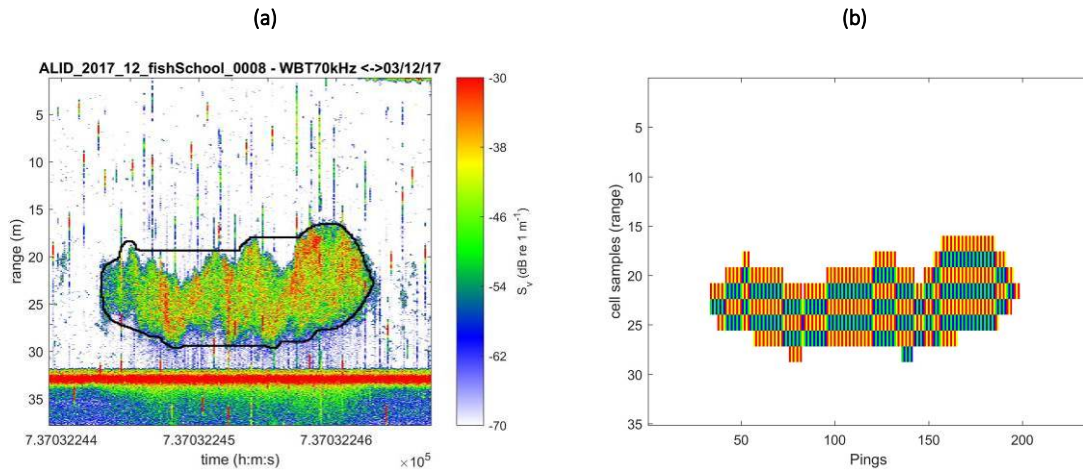


Figure 10: example of a fish school being divided into different cells. Each cell is associated with its time signal and corresponding spectrum. This constitutes the basis for the analysis and classifier training. Classification is computed for each cell. (a) Echogram of a fish school. (b) corresponding cell division,

4.2.2 Analysis

Analysis was conducted on the data collected in December by the Alida (SCH6).

4.2.2.1 Spectral trends

An analysis of the trends in frequency response for different species was carried out (WMR). A conceptual workflow of this analysis is presented in Figure 11(a). Results are shown in Figure 11(b) and (c). These early results suggest distinctive trends between Herring, Sprat and Horse Mackerel. Though not included in this analysis, the frequency response of Mackerel has been shown to be very distinct [7, 8]. This species will be included in future analysis. Further details on the analysis is presented in Annex II.

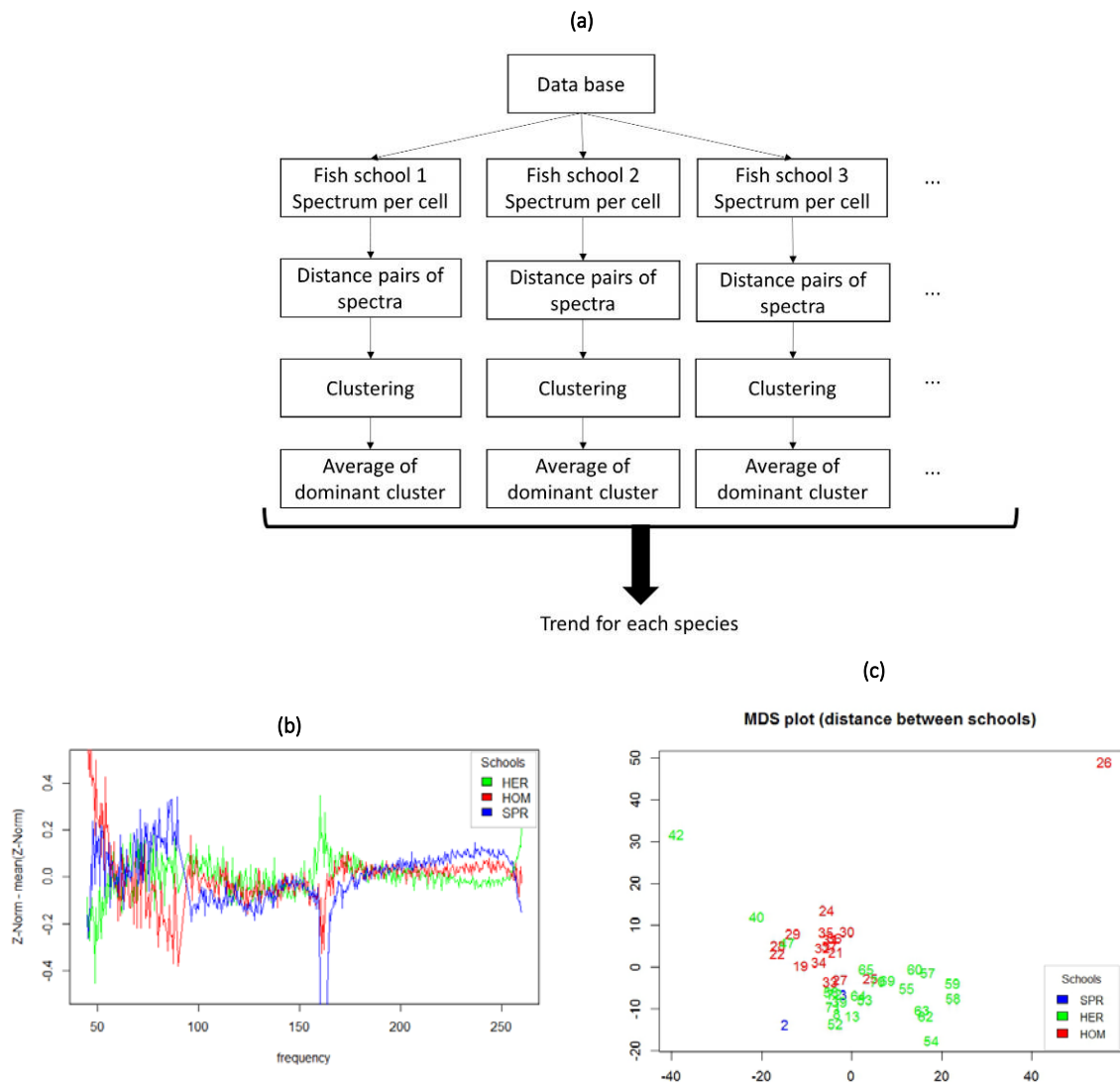


Figure 11: workflow and results of the analysis using the dynamic time warping (DTW). This analysis was conducted on the data collected by the Alida (SCH06) in December 2017 (Table 4). (a) Concept diagram exemplifying the workflow used to compare trends in spectra between different species. From the database, each fish school is divided into different cells. Each cell consist of a time series with the corresponding spectrum. For each school, the distances between the spectra of each cell are calculated using the DTW. This distance matrix is then used to create clusters, grouping of cells with similar spectra. The spectra within the dominant cluster are then averaged and the resulting spectrum is assumed representative of the school. This process then associates every school with one representative spectrum. These representative spectra are used to investigate trends between different species. (b) Comparison of de-trend spectra between Herring (HER), Horse mackerel (HOM) and Sprat (SPR) fish schools. One can observe that HOM has a distinct downward trend between 50 kHz and 90 kHz while HER and SPR are very similar. (c) Multi-dimensional scaling plot with all schools. This shows the similarities between the different schools. The different colours are the different species in the data set.

4.2.2.2 School classification

Both classifiers developed here use a mixture of features derived from: time signals, frequency signals and time-frequency signals. While auxiliary features (e.g. time of day, depth) can be employed, these can artificially enhance the performance of the classifier. This is because the range of possibilities for these quantities are not fully represented in the training data set. In short, the implementation of auxiliary features will require further data collection.

The WMR and TNO classifier have been tested (WMR & TNO). Results of the training process for the WMR classifier is shown in Figure 12. This training is done using selected data from the data set collected by the Alida (SCH06) in December 2017. The selected data are further divided into 3 data sets:

- Training: these data are used to refine the algorithm through various iterations.
- Validation: these data are used to validate the algorithm and make sure it does not become too specialised with respect to the training data.

- Testing: these data are used to assess the performance of the NN after the training is completed.

Figure 12 (a) is a plot of the convergence of the results as the training of the algorithm is optimized. The error (cross-entropy) is shown versus “epochs” (i.e. training iteration over the selected data set). The optimal training is where the performance of the validation data set is maximized (i.e. minimized error). Figure 12 (b) shows the classification results for each class (1: SPR, 2: HER, 3: HOM) for each data set (training, validation, test, all). One can observe that the overall performance is of 77.3%. Only a small part of each school is used for training and Figure 13 shows the result of applying the trained algorithm to an entire school. A summary of the results across all schools in the data set (Alida (SCH06) December 2017, Table 4) is presented in Table 5. We can conclude from the comparison shown in Table 5 that the TNO classifier performs best while both classifier exemplify relatively high success rates. The low classification rate for SPR for the WMR classifier is due to a lack of data for this species. The method employed by TNO is more robust with respect to the low number of SPR schools available for training. To date, the algorithm was only trained using the data collected by the Alida (SCH06) in December 2017 and the overall performance both algorithms will be improved further when using other data sets.

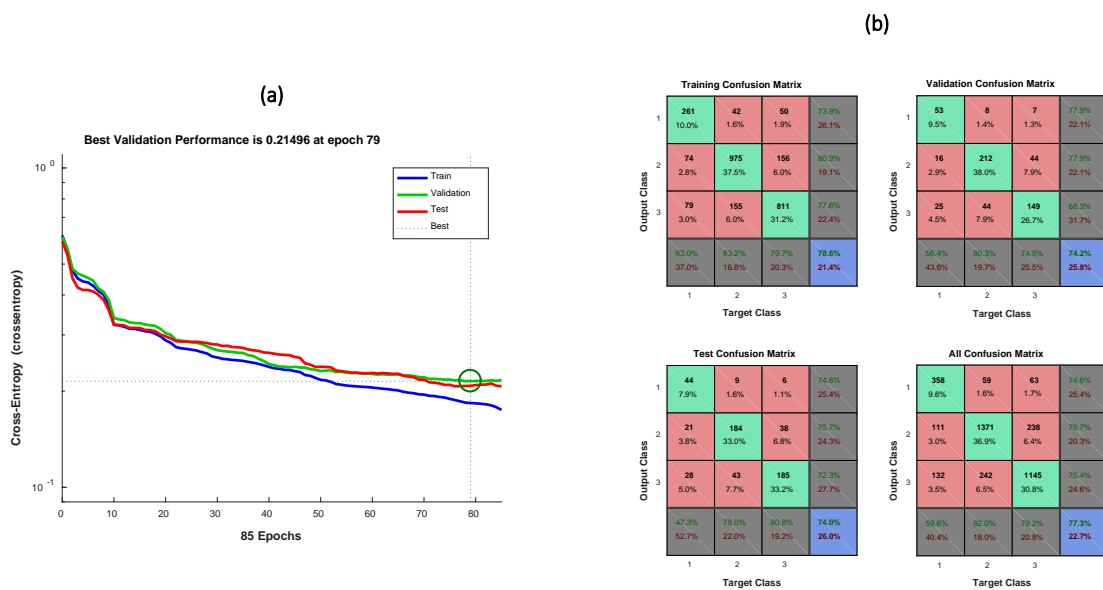


Figure 12: result of the training of the Neural Network classifier (WMR) based on data collected by the Alida (SCH06) in December 2017. (a) Performance plot for the training, validation and test sets over different iterations (“epochs”). The green circle marker shows the optimised training set. (b) Confusion plots for the training, validation and test sets. The output classes correspond to the following: 1 for Sprat (SPR); 2 for Herring (HER); 3 for Horse mackerel (HOM). Note the low number of samples for SPR. This is because the data set contained only a limited number of Sprat schools. Similar number of samples should ideally be used for the training of a NN.

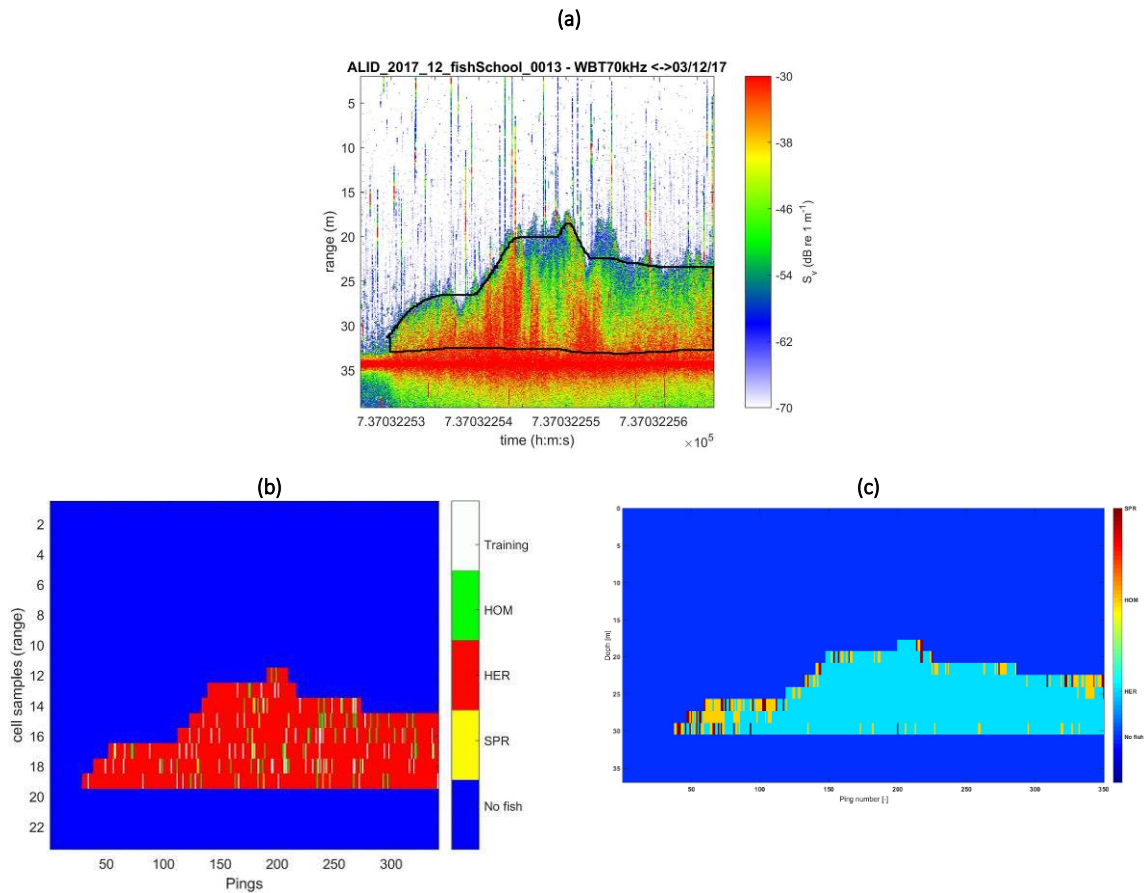


Figure 13: results from the WMR and TNO classifiers for fish school number 13 of the data collected by the Alida (SCH06) in December 2017, where HER= Herring, HOM = horse mackerel and SPR = Sprat. The catch composition of the corresponding haul consists of 100% HER and this fish school is labelled as Herring. (a) Echogram of the fish school being analysed. (b) Result from the WMR classifier (95% of cells are classified correctly). (c) Results from the TNO classifier (87% of cells are classified correctly).

Table 5: summary of results from WMR and TNO classifiers employed on the Alida (SCH06) data collected in December 2017 (Table 4). Results are shown as the percentage of correct classification by species (Herring = HER, Horse mackerel = HOM and Sprat = SPR) and classifier (TNO and WMR). It is important to note that for the data used here, the data for SPR was limited. Because the training of the NN (WMR classifier) requires data sets of similar sizes, the relatively low number of SPR data affected the overall results for this species.

	WMR classifier (%)	TNO classifier (%)
HER	78.5	87.1
HOM	74	81.1
SPR	49.6	84.7
Total	74.1	84.3

4.3 Work planned for the following reporting period

- Further improvement of the TNO and WMR classifiers (WMR & TNO). This includes adding Mackerel as a recognition species by expanding the training to more numerous data sets.
- Modify classifier to improve the result at the school edges (TNO)
- Development of the fish length estimation method
- Evaluation of fusion methods for the two species classification algorithms.
- Analyse classification robustness (TNO & WMR) to:
 - Calibration errors
 - Catch location
 - Catch season
 - Cell size
- Explore the use of ES80 data (instead of EK80) for classification (TNO & WMR).

5 WP4: Demonstrator software development

5.1 Introduction

The ECHO software (TNO) allows the processing of all raw EK80 FM data real-time on board a vessel equipped with an EK80 sonar (see Figure 2). This allows the skipper to quickly assess the data quality when recording the sonar data to fill the catch database. Additionally, the software applies all processing steps (image processing, automatic school detection etc.) required to produce classification results. Sonar images and classification results are presented to the user in real-time. Since one of the objectives of the skippers against the use of the EK80 FM is the noisy images it creates, the software additionally performs image enhancement to present de-noised sonar images to the skipper.

5.2 Work undertaken and current state of WP4

The ECHO software has been updated, summarizing:

- Compatibility with the database format defined between TNO and WMR (see Section 4.2.1): reading in and out of data, intermediary results location, naming conventions etc.
- Compatibility with the calibration format (also defined in the database) and integration of the new calibration results (see section 2).
- Speed optimization (computation, data reading) in order to account for the new classification algorithms (species and length distribution) developed currently by of WMR
- Addition of functionality based on skipper feedback
- Integration of a first TNO classifier

The software is ready for the integration of new classification algorithms, and integration on a new vessel.

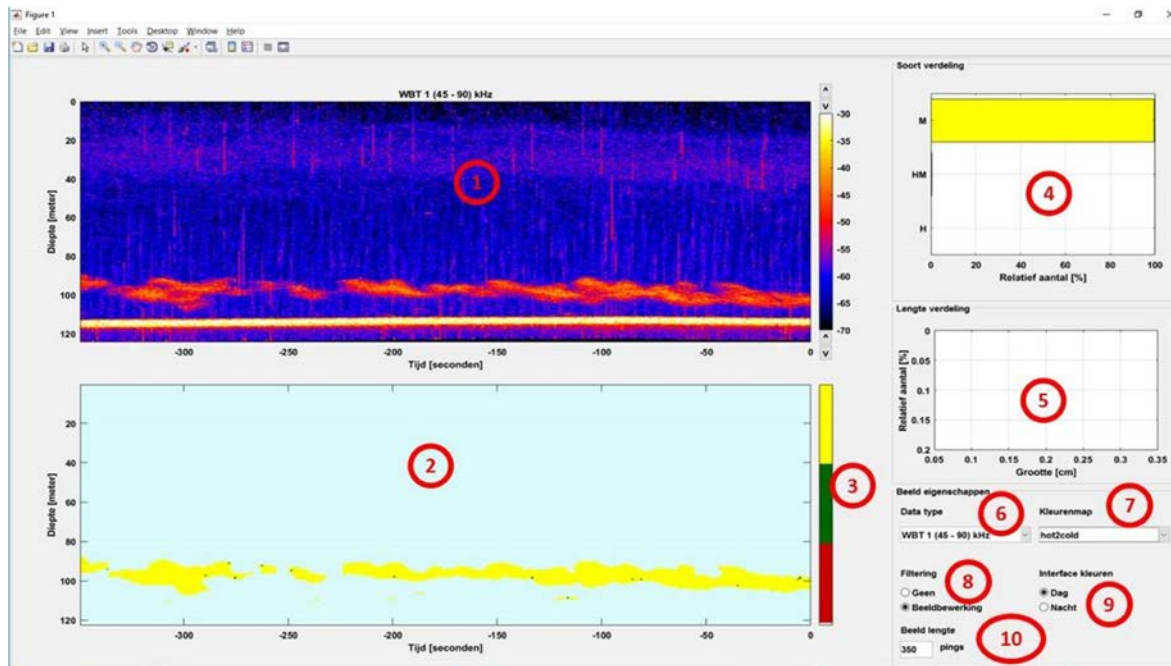
A beta version of the ECHO software has been presented and installed on a demo PC on board the Afrika (SCH24) in 2017 (see [1]). The demonstrator is still running but unfortunately, only a few datasets could be recorded since the vessel had to leave European waters (see Section 3). The classification module has been disabled on this version, since the classification modules had not been trained yet at the time of deployment (more data needed). Figure 14(a) shows a screenshot of the first version of the software with a short description of its functionalities in the caption. Figure 14(b) shows the current version as installed on the Afrika (SCH24), with classification still disabled.

Figure 15 shows the updated software version based on feedback from the skippers, which will be installed in the next iteration (September/October 2018). The updated version implements some skipper requests.

- It allows selection of part of the school to estimate classification and length estimate statistics from.
- It detects and isolates external disturbances in the EK80 data. It displays a combined spectrum of the disturbances with which the skipper may derive the disturbing source.
- To enhance EK80 images in the ECHO software and in the main EK80 display software, the skipper can decide to switch off the disturbing sources.
- Additional filter banks have been implemented and given access to for the user. Horse Mackerel schools tend to be sparse. With high resolution EK80 images, skippers indicate to find it difficult to detect these schools. The filter banks allow filtering and displaying EK80 data in more narrow bands (lower resolution) to allow easier detection by the skipper.
- EK80 settings validation routines have been implemented to warn the skipper in case the EK80 is used in a mode in which no proper data is collected for the project. This concerns

wide-band/narrow-band mode, transceiver types switched on, transceiver frequency bands used etc. This ensures that data collection is done properly.

(a)



(b)

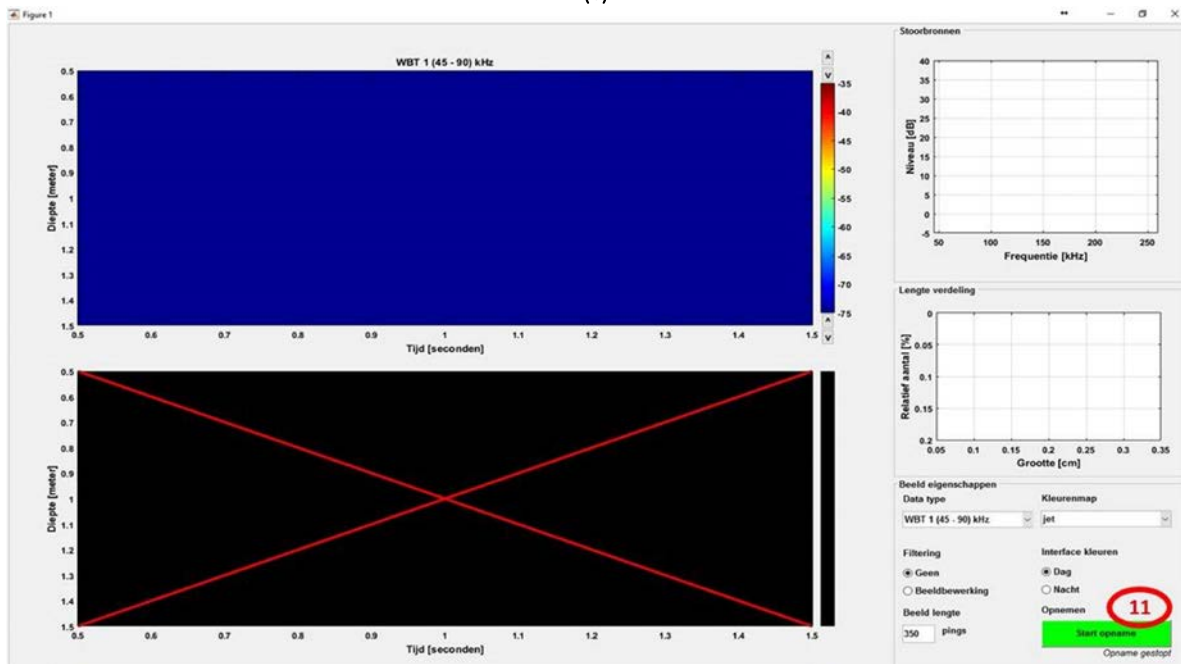


Figure 14: screen shots of the two ECHO software versions implemented to date.

(a): beta version of the ECHO software: (1) Echogram display; (2) Fish school detection and classification display with one color per species (see colorbar (3)); (4) species composition in displayed echogram in % for each of the species (M: Mackerel, HM: Horse Mackerel, H: Herring, 100 % Mackerel here); (5) length distribution estimate (not implemented yet); (6): drop down menu with data type selection, as displayed in echogram; (7): echogram colormap; (8): echogram filtering; (9): day/night mode (i.e. interface background color); (10): number of pings displayed in echogram.

(b): Screenshot of the data record version of the ECHO software installed on the Afrika (Sch24). This software version does not have any classification functionality activated yet. In this phase, it is solely used to filter data recorded by the EK80 through a push button (11) and flag potential errors in the settings of the EK80.

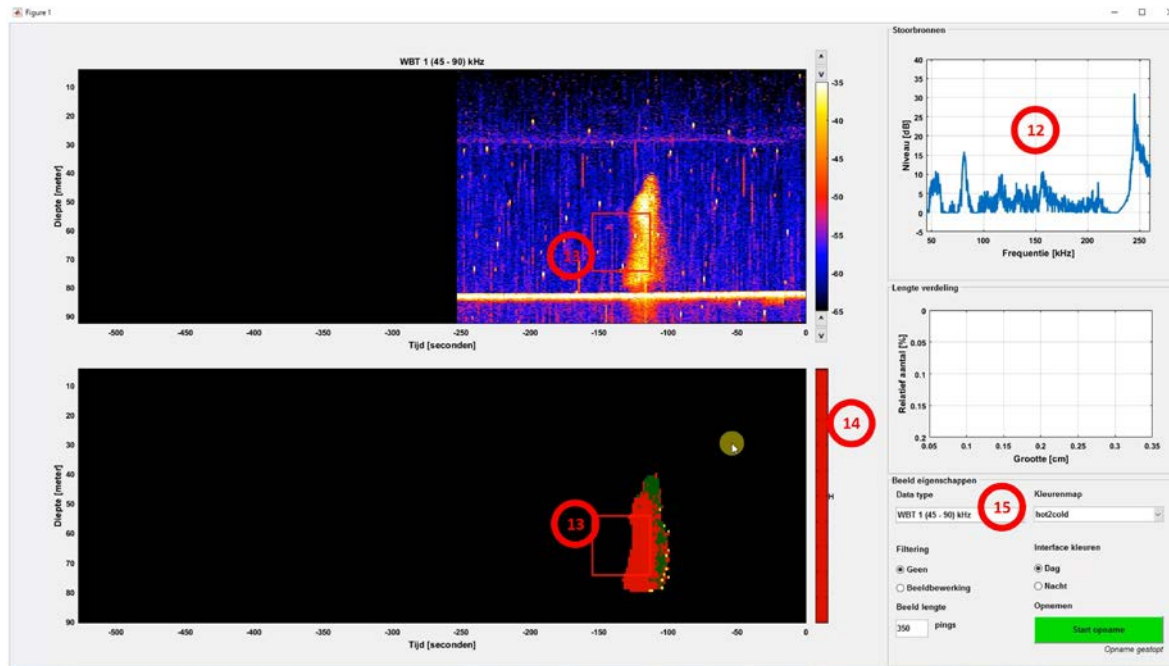


Figure 15: Screen shot of the updated ECHO software based on skipper feedback. Additional or modified items are indicated by the red circles again. (12) A spectral representation of disturbing sources other than the EK80; (13) Ability to select part of the school to compute classification and length estimation statistics on; (14) Colorbar that is split in colors of the different species of which lengths of the different parts of the colorbar represents the classification percentages of the species.; (15) Additional filtered versions (narrow band) of the data to reduce data resolution when desired.

5.3 Work planned for the following reporting period

- Integration of the WMR and TNO classifiers (TNO).
- Update/training of both classifiers with new datasets (analysis and processing in progress).
- Perform software improvements to increase execution speed to support operation in deeper waters (> 150m) (TNO)
- The ECHO software will be installed on a new demonstrator (new PC mounted on board the Alida (SCH06)) (WMR)
- Work on the most effective way to present the classification results to the skipper (classifier fusion) (WMR & TNO)

6 WP5: Management, communication & reporting

6.1 Meetings

- Several project meetings and communications between WMR and TNO took place to discuss project development:
 - Meeting at TNO (01/05/2017)
 - Meeting at TNO (05/10/2017)
 - Meeting at WMR (12/03/2018)
- Meeting at WMR (12/03/2018, WMR, RVZ) and meeting at BCS Harderwijk BV, Amsterdam (29/01/2018, WMR, RVZ, van der Zwan, sustainovate):
 - discuss data recording between the realFishEcho and the SEAT projects. It was decided to have data for the SEAT project being collected by the Willem van der Zwan (SCH302), while the Alida (SCH6) focuses on collecting data for the realFishEcho project.
 - Plan data recording for the next months

6.2 Miscellaneous

- Regular email/phone contacts between WMR & TNO.
- Regular communication with the skippers (mostly from Alida (SCH06)) to provide feedback on the project to skippers.
- Regular visit to the vessel to collect data and gather feedback from the skippers.
- The project was presented at the following events:
 - Bioacoustics day conference (Den Helder, 20/10/2017):
"Fish species identification using wideband echosounders"
 - Species recognition meeting (RVZ, Scheveningen, 08/11/2017):
"RealFishEcho - broadband (EK80) species recognition"
- The development of the fish length estimation method needs to be delayed as more data and especially diversity in the length frequency distributions are needed. To date, fish length measurements with corresponding acoustic data were only collected for two species (Herring, Horse Mackerel) and by only one vessel (Alida (SCH06)) in December 2017, January 2018 and March 2018. Initial work on the fish length estimation method was undertaken based on the December 2017 data set from the Alida (SCH06).

7 Quality Assurance

Wageningen Marine Research utilises an ISO 9001:2008 certified quality management system (certificate number: 187378-2015-AQ-NLD-RvA). This certificate is valid until 15 September 2018. The organisation has been certified since 27 February 2001. The certification was issued by DNV Certification B.V.

On the basis of this accreditation, the quality characteristic Q is awarded to the results of those components which are incorporated in the scope, provided they comply with all quality requirements. The quality characteristic Q is stated in the tables with the results. If, the quality characteristic Q is not mentioned, the reason why is explained.

The quality of the test methods is ensured in various ways. The accuracy of the analysis is regularly assessed by participation in inter-laboratory performance studies including those organized by QUASIMEME. If no inter-laboratory study is available, a second-level control is performed. In addition, a first-level control is performed for each series of measurements.

In addition to the line controls the following general quality controls are carried out:

- Blank research.
- Recovery.
- Internal standard
- Injection standard.
- Sensitivity.

The above controls are described in Wageningen Marine Research working instruction ISW 2.10.2.105. If desired, information regarding the performance characteristics of the analytical methods is available at the chemical laboratory at IJmuiden.

If the quality cannot be guaranteed, appropriate measures are taken.

Annex I: Automatic reporting for fishing trip overview

Drawing on analyzed data in the data base format (Section 4.2.1), R codes were developed to automatically generate tables, maps and plots in the format of a short report. This aims at summarizing the trawl catch and acoustic recordings of each fishing trip.

For this work, the R-studio functionalities and relevant R packages (e.g. "rmarkdown", "knitr" and "ggplot2") were used. This section provides examples of: (1) catch characteristics (Table I.1, Figure I.1-I.4); (2) combination of acoustic and catch data (Figure I.5). Catch data recorded by fishermen are particularly important because they are the base for ground thruthing, i.e. allocating specific species to acoustic records.

Figure I.1 and I.2 show the total catch per species relative to date and duration of haul. In figure I.2, the bar corresponding to 3rd December 2017 indicates high amount of herring in the catch in a short duration in the first haul of the day. This catch was followed by horse mackerel catches as can be seen on figure 3 for the second and third haul. This is an indication that horse mackerel and herring co-occur in the area. Therefore, one should be careful in the interpretation of the schools on the echograms when generating the ground truth table. The histograms in figure I.3 give a summary of the length distribution across the hauls e.g. whether there is a structure in the distribution of length composition across the hauls that can be used for species identification.

Summary of the catch

Table I.1: Catch Summary table, where HER = Herring, HOM = Horse mackerel, MAC = Mackerel, SPR = Sprat.

Date	No.Haul	Tot.Duration	Tot.Catch	HER	HOM	MAC	SPR	OTHER
2017-12-02	1	2.8	10	0	0.0	0	10	0
2017-12-03	3	3.2	290	120	170.0	0	0	0
2017-12-04	1	0.8	120	0	120.0	0	0	0
2017-12-05	3	7.1	155	40	115.0	0	0	0
2017-12-06	3	4.4	110	76	30.0	0	0	4
2017-12-07	4	5.6	160	130	0.0	0	0	30
2017-12-08	3	6.1	90	70	0.0	0	0	20
2017-12-09	3	4.4	145	145	0.0	0	0	0
2017-12-10	1	1.1	40	20	0.0	20	0	0
2017-12-11	2	0.4	95	95	0.0	0	0	0
2017-12-15	4	9.4	55	0	4.8	15	0	35
2017-12-16	1	1.3	5	0	0.0	0	0	5

Catch composition and haul duration

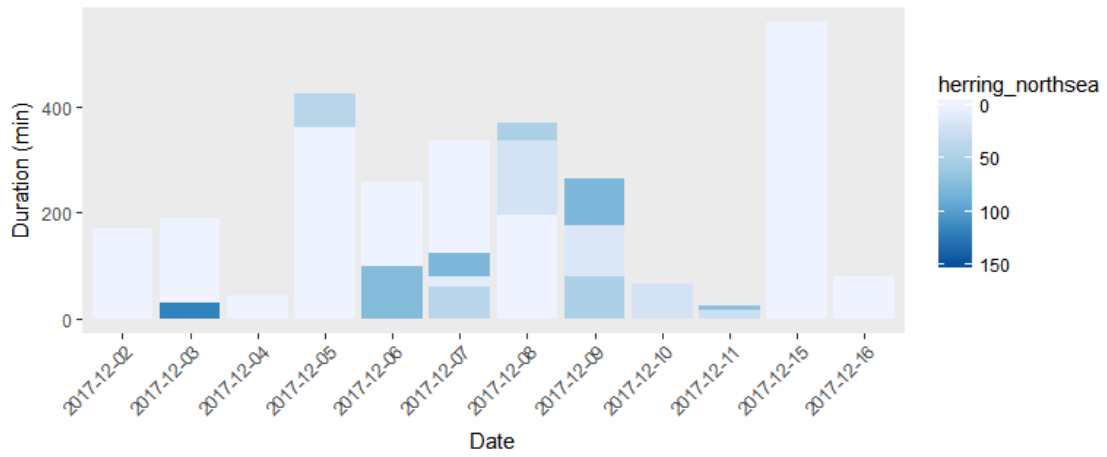


Figure 1.1: Total haul duration (bar length) and total **herring** catch (colour) per date. The blue colour scale represents different catch sizes.

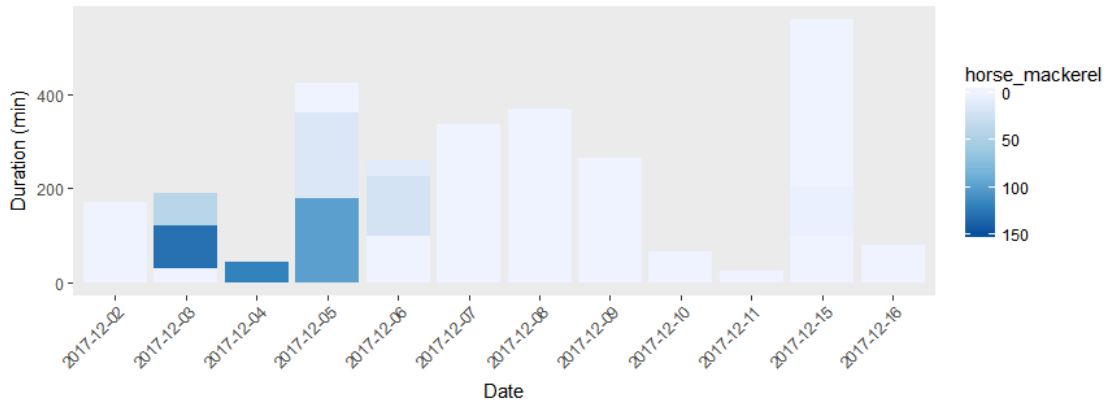


Figure 1.2: Total haul duration (bar length) and total **horse mackerel** catch (colour) per date. The blue colour scale represents different catch sizes.

Length distribution plots

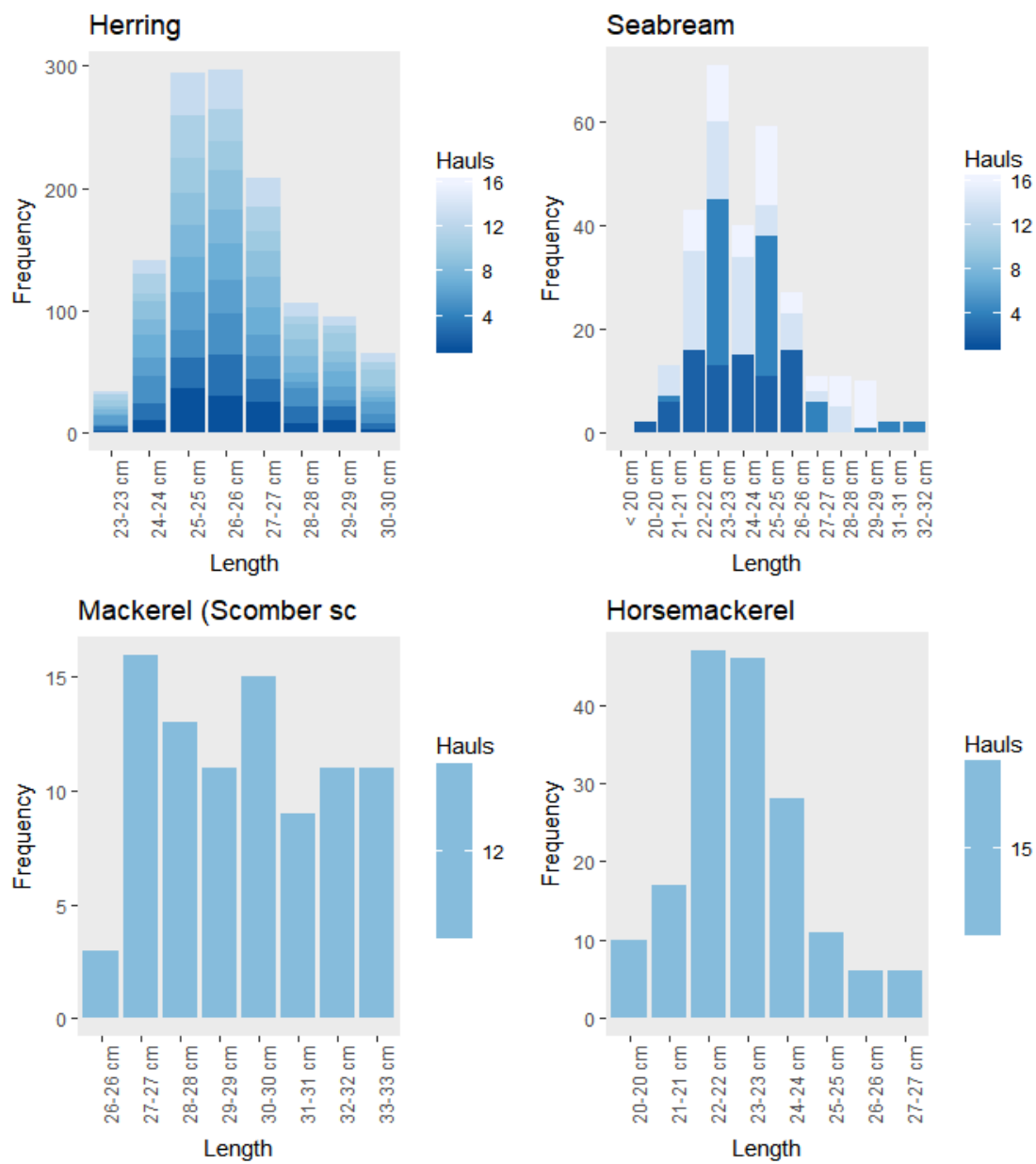


Figure I.3: Length distribution species/hauls

Fishing trip maps

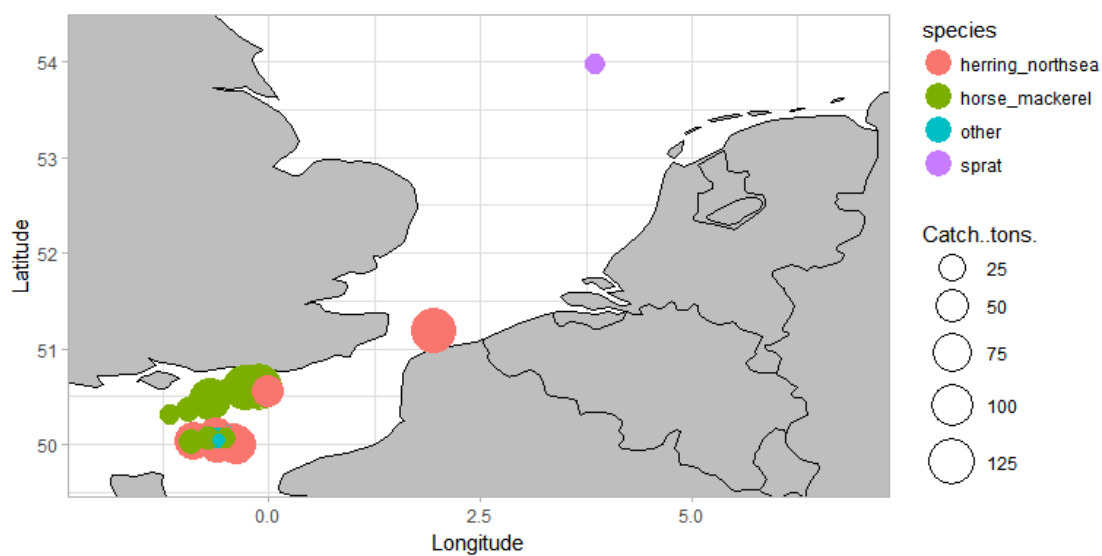


Figure I.4: Location of the trawl hauls and catch composition and amount (bubble plot)

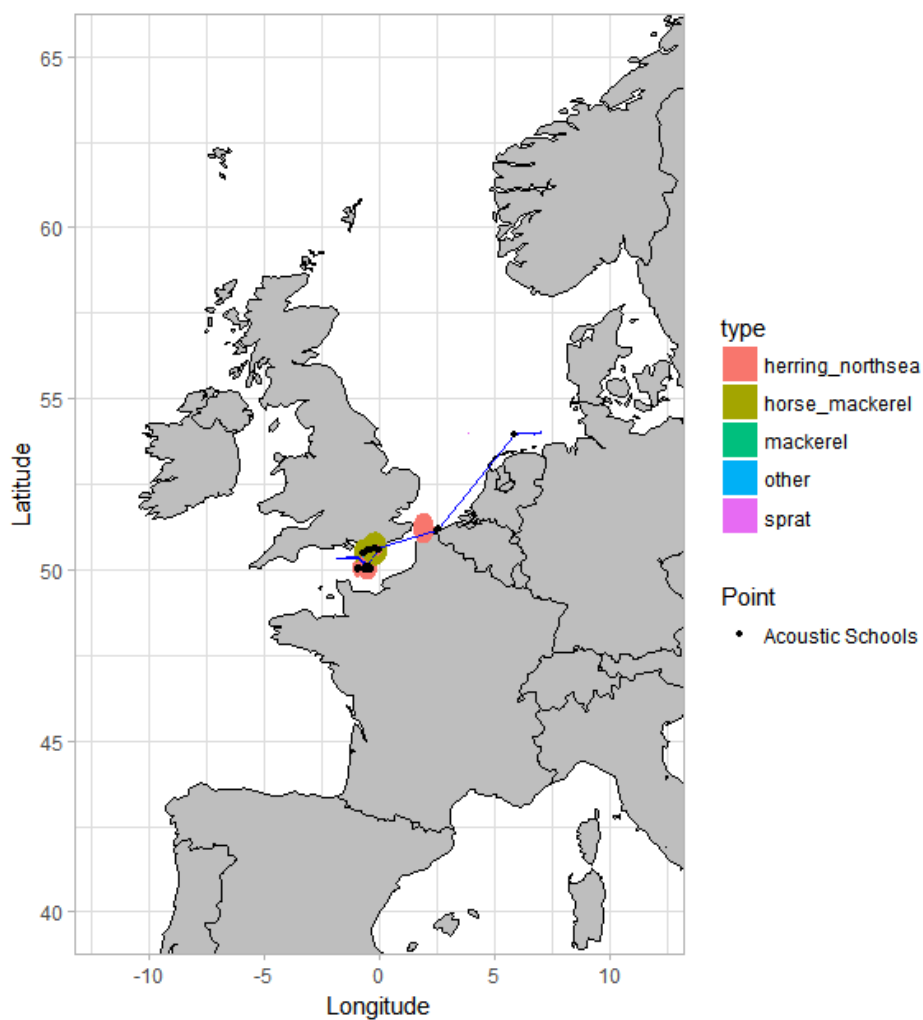


Figure I.5: Overview Map

Annex II

Introduction

The spectral response of fish schools recorded during the December 2017 fishing trip of Alida (SCH6) are assessed in this section. The selected fish schools correspond to those that correlate with catch data. Each fish school is divided into cells (Figure 8) and a spectrum is associated to each cell. ground-truthing of each cell is based on catch data (i.e. catch composition, Table I.1).

The distance between each spectrum within a fish school was calculated using the dynamic time warping method (DTW) which is commonly used in the comparative analysis of time series data. This technique allows one to find the best alignment between two series by accounting for local shifts in peaks and dips. Here, this technique allows the grouping of similar spectra within a school. Further, the main characteristic frequency response of this school could be extracted using clustering methods, eliminating non representative cells (e.g. because of noise).

Based on these extracted components, summary plots were generated by averaging the main clusters within fish schools and between species. Results show that in the overall averages there are differences between species. There is a pronounced difference between herring and horse mackerel in the absolute spectral levels and subtle differences in the normalized spectra. The difference becomes more apparent when the global average is subtracted from species average. Together with apparent distinction between averaged spectral trends of species, there are also considerable overlaps between individual schools of different species. Furthermore, some of the individual schools exhibit opposing trends relative to the average of that species, which indicate either incorrect labelling in from the ground-truthing method file or factors affecting the spectral response other than the species type (e.g. depth).

Method

After the extraction of spectra for each cell, a post processing routine is applied in order to summarize and visualize the main characteristics of the frequency response of each school and species. The workflow used for this process shown in Figure II.1. This is further explained in the following section. A catalogue of summary plots were generated per schools for each species.

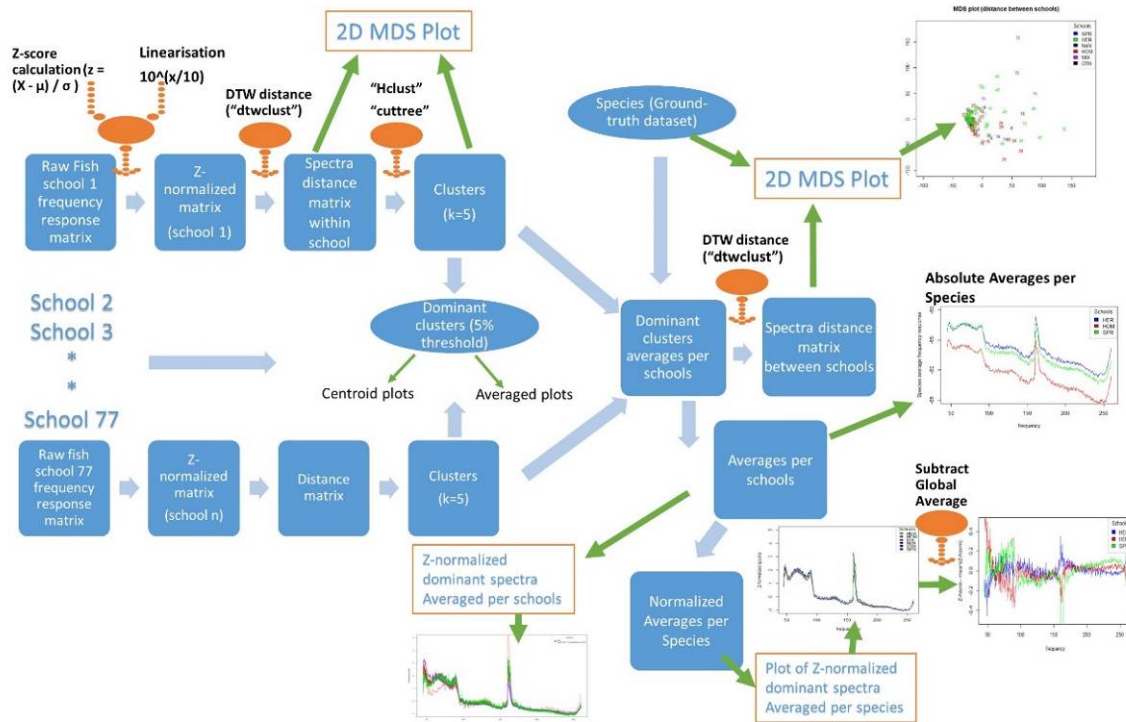


Figure II.1: processing workflow for the spectral analysis.

Description of the workflow

Pre-processing

- Spectral responses are extracted using the demonstrator software (MATLAB code, Section 5). The extracted data consist of a spectrum for each sampling units (depth range vs ping). This includes the full frequency range from 45 kHz to 260 kHz (combination of three channels) with total number of 446 frequency points. The real part of each spectra together with associated depth, ping number and frequency array are exported as MATLAB files and imported to R for further processing

Dynamic Time Warping (DTW)

- Prior to clustering, each spectrum is centred and normalised by its standard deviation (Z-normalization): $z = (X - \mu) / \sigma$. In this equation, X is the spectrum in linear domain, μ the mean of the time series and σ the standard deviation of the time series.
- Normalized spectra are then converted to the R time series matrix format for computation efficiency.
- The distance between each spectrum pair was computed using the Dynamic Time Warping method (DTW) using the R package 'dtwclust'. This results in a distance matrix containing pair-distances between each spectra.
- Using the distance matrix, a hierarchical clustering is computed for each school using the "hclust" R function. The "cuttree" R function is then used to extract different number of clusters ($k=n$), here varying from two to five. All cluster groupings are stored in output files but only $k=5$ was used for further analysis and visualization
- To visually examine the distribution of the different clusters, a two-dimensional multidimensional scaling (MDS) plot is created based on the distance. For visualization, each cluster members are plotted with different colors. Cluster numbering is ranked based on the number of members from 1 to 5 (cluster #1 contained the highest number of members)

Visualization 1- Examination within each schools (Figure II.2)

- Plots are generated using the clusters that contain more than 5 percent of the total number of cells. Subsequent plots include:
 - A Z-Normalized spectrum for each selected cluster
 - An original spectrum in dB scale for each selected cluster
 - Average of the Z-Normalized values of spectra for each selected cluster
 - Average of spectra for each selected cluster (calculated in linear and converted back to dB)

Visualization 2- Examination between schools (Figure II.6)

- Each spectrum was labeled with the cluster id in the first section. These labeled spectra are filtered based on these cluster ids such that those having members less 15% of samples within each school are discarded. Remaining clusters are averaged.
- From these averaged spectra, distances are calculated using the DTW method (as described above).
- Using this new distance matrix, a MDS plot is created using the ground-truthed school labels (e.g. HER = herring, HOM=Horse mackerel, SPR=sprat...). This two dimensional MDS plot gives insight in the differences in spectral response between the schools.

Visualization 3- Spectra averages (Figures II.3-II.5)

- In order to visualize the shape of the spectra for each species, the average of dominant clusters within each school are used. In order to eliminate noisy elements due to insufficient number of samples to average, schools having less than 200 samples are not used. Additionally, if the second cluster have a number of cells of more than 20% compared to the first cluster and have a total number of cells of more than 90, this second cluster is merged with the first one. Otherwise, only the first cluster is used. Decision of these numbers are based on visual examination. As a result, from 77 schools, only 53 schools met the selection criteria.
- Using these 53 schools two additional plots are generated,
 1. average of spectra (belonging to the dominant clusters as described above) for each school
 2. Z-normalization of averaged spectra.

These two results were overlaid on a plot with different colors of each school.

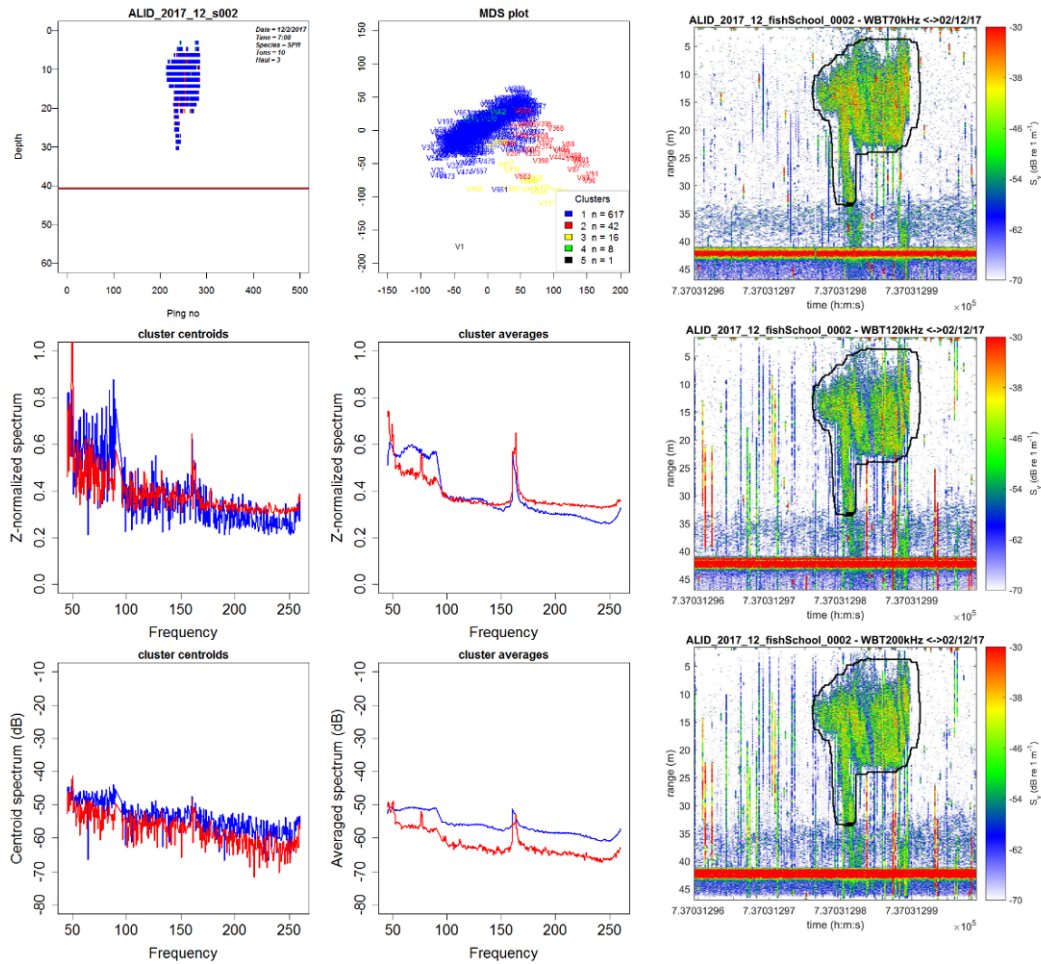


Figure II.2: An example of fish school information catalogue. These plots were created for each schools in order to visually examine the spectrum characteristics. Upper left plot show the cluster members within school the inner legend of this plot also show groundtruth information. Different cluster members are indicated as different colors. Blue is most dominant cluster with highest number of members, red is second yellow is third: green is fourth and black is the last. Mid-left plot show the z-normalized centroid spectra according to the MDS and left bottom plot show the absolute values of these spectra. Mid-centre plot show the average of normalized spectra of each cluster member and bottom-centre plot show the average of absolute spectra. The plots on the right are the echograms for the each centre frequency 8 70, 120 and 200 kHz).

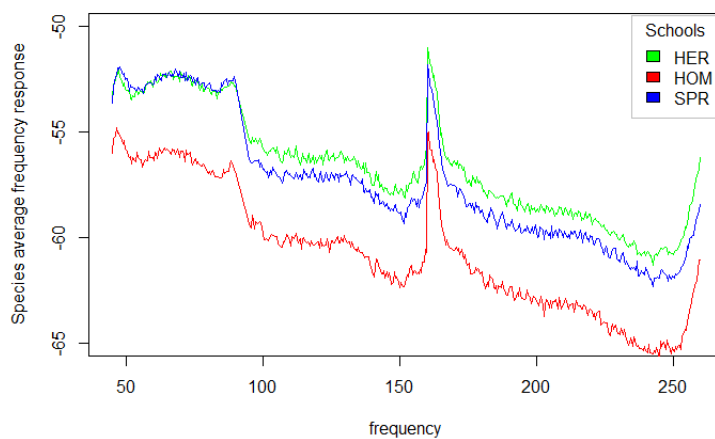


Figure II.3: Summary plot of spectra of species groups (HER = Herring, HOM = Horse mackerel, SPR = Sprat). In this plot, the average of dominant cluster members for each school is then averaged per species for the absolute values.

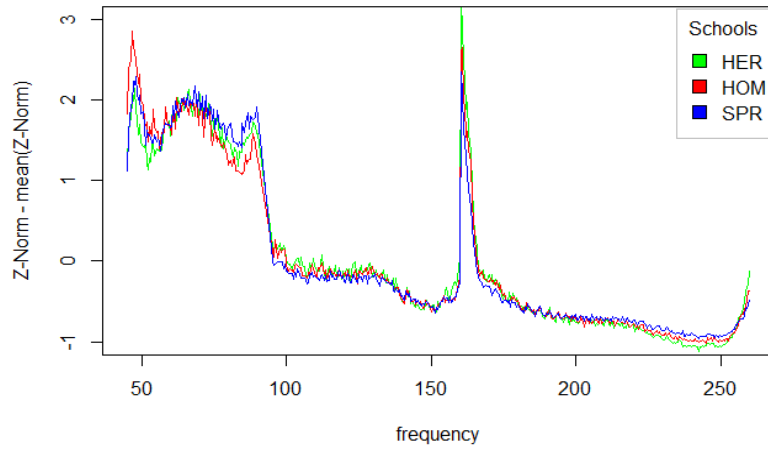


Figure II.4: Summary plot of normalised spectra for each species (HER = Herring, HOM = Horse mackerel, SPR = Sprat). The plot was generating the absolute values to Z-Scores to be able to compare the trends between species.

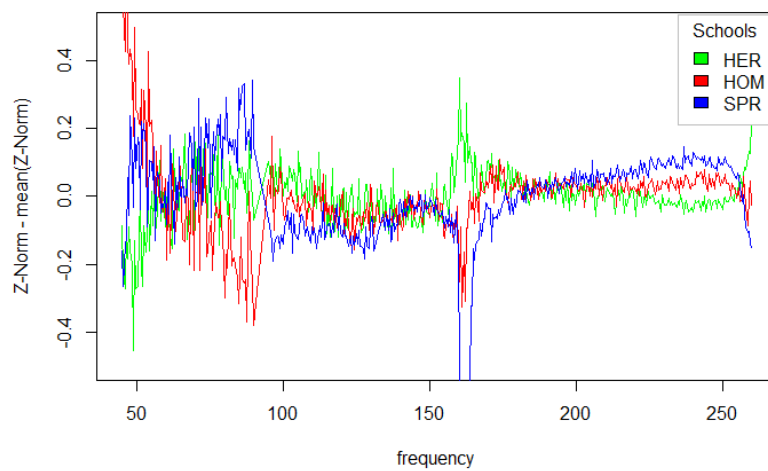


Figure II.5: Global average subtracted from average of species (HER = Herring, HOM = Horse mackerel, SPR = Sprat). Trends in frequency response are more clearly visible compared to previous figure.

MDS plot (distance between schools)

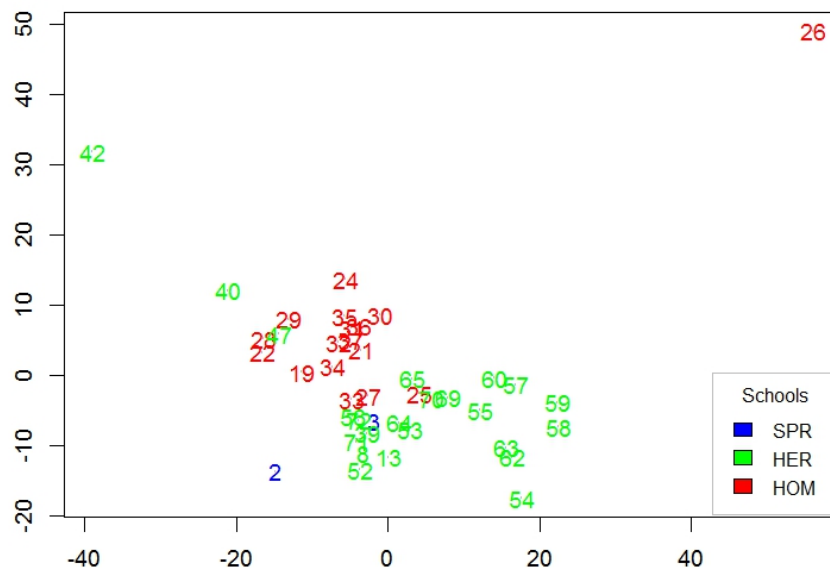


Figure II.6: MDS plot showing the similarity between individual schools according to the DTW distances (HER = Herring; HOM = Horse mackerel; SPR = Sprat). Numbers show the index number of a school in the database and color indicate species.

Results and discussion

The summary plots (Figure II.3-II.5) show a distinction between species. Overall, horse mackerel has consistently lower amplitudes relative to sprat and herring (Figure II.3). There is also a distinction between herring and sprat in the 120-200 kHz frequency band relative to the 50-100 kHz frequency band (Figure II.3). The overall difference of herring and sprat however is relatively less compared to their differences with horse mackerel. Despite the apparent differences in the absolute levels, the normalized spectra (Figure II.4) show only subtle differences and the main shape of the spectra is likely to be a result of the echosounder signature (or calibration). When the global average is subtracted from the average of individual school, the specific trend became more apparent (Figure II.5). In the 50-100 kHz frequency band, horse mackerel shows a decreasing trend while sprat and herring is increasing. Despite this pattern in the averages of the schools, results cannot be generalized as a representative characteristic for all members of a species. This is because there are schools showing the opposing trends of both in herring and horse mackerel species. This discrepancy can be either due to incorrect labeling in the ground-truthing data set (e.g. herring school occurring among horse mackerel schools) or factors influencing frequency responses other than species. Furthermore, similar opposing trends also exist even within a school between the clusters. In the 170-250 kHz frequency band, sprat and herring show opposing trends while horse mackerel show a flat response. The MDS (Figure II.6) shows local clusters of herring that are separated from horse mackerel. The differences in the centroids potentially indicate a pattern specific to each species. On the other hand, large degree of overlap is making a meaningful grouping per species difficult. As a result, whether the local groupings are generated randomly or based on a genuine signal needs to be clarified with more data (e.g. other data sets).

References

- [1] B. J. P. Berges, S. Sakinan, and E. van Helmond, "Practical implementation of real-time fish classification from acoustic broadband echo sounder data- RealFishEcho progress report : Year 1- June 2017," *IJmuiden Wageningen Mar. Res. (Wageningen Mar. Res. Rep. C082/17)*, p. 29, 2017.
- [2] D. G. Reid and E. J. Simmonds, "Image Analysis Techniques for the Study of Fish School Structure from Acoustic Survey Data," *Can. J. Fish. Aquat. Sci.*, vol. 50, no. 5, pp. 886–893, May 1993.
- [3] J. Simmonds, F. Armstrong, and P. J. Copland, "Species identification using wideband backscatter with neural network and discriminant analysis," *ICES J. Mar. Sci.*, vol. 53, no. 2, pp. 189–195, Apr. 1996.
- [4] K. G. Foote, H. P. Knudsen, G. Vestnes, D. N. MacLennan, and E. J. Simmonds, *Calibration of acoustic instruments for fish density estimation: a practical guide*. Copenhagen Denmark: International Council for the Exploration of the Sea, 1987.
- [5] D. A. Demer, L. Berger, M. Bernasconi, E. Bethke, K. Boswell, D. Chu, R. Domokos, A. Dunford, S. M. M. Fassler, S. Gauthier, L. Hufnagle, M. Jech, N. Le Bouffant, A. Lebourges-Dhaussy, X. Lurton, G. J. Macaulay, Y. Perrot, T. Ryan, S. Parker-Stetter, S. Stienessen, T. Weber, and N. Williamson, "Calibration of acoustic instruments," *ICES Coop. Res. Rep.*, vol. 326, p. 133, 2015.
- [6] B. J. P. Bergès and E. van Helmond, "Practical implementation of real-time fish classification from acoustic broadband echo sounder data - RealFishEcho: classification algorithm improvements," *(Wageningen Mar. Res. Rapp. C010/18)*, p. 15, 2017.
- [7] S. M. M. Fassler, B. C. Scoulding, D. Burggraaf, D. de Haan, B. Quesson, J. van de Sande, and P. Beerens, "VIP report 'Use of new broadband echosounder': Techniques for improved ocean imaging and selectivity in pelagic fisheries," *IMARES Rep.*, vol. C171/15, p. 100, 2015.
- [8] A. Antona, "Remote Fish Species and Size Identification Using Broadband Echosounders," Wageningen UR, 2016.

Justification

Report C062/18

Project Number: 4311400004

The scientific quality of this report has been peer reviewed by a colleague scientist and a member of the Management Team of Wageningen Marine Research

Approved: Mw. H.M.J. van Overzee
Researcher

Signature:



Date: 6 September 2018

Approved: Drs. J. Asjes
Manager integration

Signature:



Date: 6 September 2018

Wageningen Marine Research
T +31 (0)317 48 09 00
E: marine-research@wur.nl
www.wur.eu/marine-research

Visitors' address

- Ankerpark 27 1781 AG Den Helder
- Korringaweg 5, 4401 NT Yerseke
- Haringkade 1, 1976 CP IJmuiden

Wageningen Marine Research is the Netherlands research institute established to provide the scientific support that is essential for developing policies and innovation in respect of the marine environment, fishery activities, aquaculture and the maritime sector.

Wageningen University & Research:

is specialised in the domain of healthy food and living environment.

The Wageningen Marine Research vision

'To explore the potential of marine nature to improve the quality of life'

The Wageningen Marine Research mission

- To conduct research with the aim of acquiring knowledge and offering advice on the sustainable management and use of marine and coastal areas.
- Wageningen Marine Research is an independent, leading scientific research institute

Wageningen Marine Research is part of the international knowledge organisation Wageningen UR (University & Research centre). Within Wageningen UR, nine specialised research institutes of the Stichting Wageningen Research Foundation have joined forces with Wageningen University to help answer the most important questions in the domain of healthy food and living environment.

