D5.3. FishData analysis

2022-07-12

SMARTFISH H2020

Innovation for sustainable fisheries

Project number: 773521
Project duration: 1 Jan 2018 – 31 Dec 2022
Project coordinator: Bent Herrmann, SINTEF Ocean
Web site: www.smartfishh2020.eu

Deliverable ID: D5.3 | Due month: M48 | Preparation date: 2022-07-12

Title:

FishData analysis

Open access revision

Lead beneficiary:
SINTEF Ocean

Prepared by:
Lars T. Kyllingstad
Karl-Johan Reite
Joakim Haugen
Jarle Ladstein

Approved by:
Rachel Tiller

Abstract

The goal of SMARTFISH H2020 is to develop, test and promote a suite of high-tech systems for the EU fishing sector, thereby optimizing resource efficiency, improving automatic data collection for fish stock assessment, provide evidence of compliance with fishery regulations and reduce the ecological impact of the industry.

Here, we present a prototype system for visualising and studying information about catch composition and catch efficiency in fisheries. The system mainly consists of two web applications. One is based on data from automatic catch recording technologies and other onboard sensors, and shows information about a limited number of vessels. The other is based on publicly available official catch reports from the entire fishing fleet. SMARTFISH H2020 has also contributed to the development of a third application which provides forecasts of marine environmental conditions. The applications are aimed at three stakeholder groups: fishers, who may use them as decision support when planning their operations; fisheries managers, who may use them for real-time monitoring and control; and marine researchers, who may use them in a stock assessment setting.

This is an open-access revision of a confidential deliverable. Some of the text in the original report has been removed.

Deliverable type

R	Document, report	
DEM	Demonstrator, pilot, prototype	X
DEC	Web sites, patent fillings, videos, etc.	
OTHER	Software, technical diagram, etc.	

Authorship information

Editor	Lars T. Kyllingstad
Contributing partners	SINTEF Ocean, DTU Aqua

Version history

Version number	Date	Description of changes
1	2021-12-29	Initial version delivered to the European Commission.
2-oa	2022-07-12	Revised version for open-access publication. Sensitive information concerning industrial exploitation and commercialisation has been removed. This version also has minor corrections and textual improvements that didn’t make it into the delivered version 1 (and therefore haven’t been reviewed by EU project monitors).

SMARTFISH H2020 consortium

SMARTFISH H2020 (773521) is an Innovation Action within Horizon 2020, the European Union's framework programme for research and innovation, Call H2020-SFS-2017-1, Topic SFS-22-2017.

The following table contains information about the consortium members.

	SINTEF Ocean Postboks 4762 Torgarden NO-7465 Trondheim NORWAY www.sintef.no/ocean	Project Coordinator: Bent Herrmann bent.herrmann@sintef.no Project Manager: Rachel Tiller Rachel.tiller@sintef.no
	AZTI-Tecnalia 48395 Sukarrieta SPAIN www.azti.es	Guzmán Diez gdiez@azti.es +34 667 174 392
	DTU Aqua Danmarks Tekniske Universitet Willemoesvej 2 9850 Hirtshals DENMARK www.dtu.dk	Ludvig Ahm Krag lak@aqua.dtu.dk +45 21 31 64 57
	Marine Scotland Marine Laboratory 375 Victoria Road Aberdeen, AB11 9DB UNITED KINGDOM http://www.gov.scot/Topics/marine/science/	Neil Campbell Neil.Campbell@gov.scot +44 (0)131 244 2902
	Melbu Systems AS Sminesgata 3, 8445 Melbu NORWAY http://www.melbusystems.no/	Einar Roger Pettersen einar@melbusystems.no +47 900 21 883
	Çukurova University Fisheries Faculty 01330, Saricam, Adana TURKEY http://www.cu.edu.tr/Eng/Default.aspx	Gökhan Gökçe gokhan.gokce@ymail.com +90 322 330 6084 ext: 2961
	Cefas Pakefield Road Lowestoft Suffolk NR33 0HT UNITED KINGDOM https://www.cefas.co.uk/	Thomas Catchpole thomas.catchpole@cefas.co.uk +44 (0) 7768041107
	Mersin University Fisheries Faculty Yenisehir Campus 33169 Mersin TURKEY http://www.mersin.edu.tr/	Hüseyin Özbilgin ozbilginh@gmail.com +90 3243412818

	NTNU Institutt for Marin Teknikk, NTNU N-7491 Trondheim NORWAY https://www.ntnu.no/imt	Martin Ludvigsen martin.ludvigsen@ntnu.no +47 91897272
	University of East Anglia Norwich Research Park Norwich Norfolk, NR4 7TJ UNITED KINGDOM https://www.uea.ac.uk/	Michal Mackiewicz M.Mackiewicz@uea.ac.uk +44 (0)1603 59 1206
	Marport France SAS 2 allée Copernic Espace Eurêka Parc Technologique de Soye 56270 Ploemeur, France	Didier Caute dcaute@marport.com +33 622 012 347
	Zunibal Idorsolo, 1 48160 Derio Bizkaia SPAIN http://zunibal.com/en/	Iratxe Arraibi iratxe.arraibi@zunibal.com (+34) 944 977 010 (7:30-15:30) (+34) 608 556 661
	SafetyNet Technologies LTD 29 Shand Street London, SE12ES UNITED KINGDOM http://sntech.co.uk/	Dan Watson dan@sntech.co.uk +44 779 294 9727
	Nergård Havfiske Langnesveien 18 9408 HARSTAD Norway http://nergardhavfiske.no/	Kjell Larssen kjell.larssen@nergard.no +47 48038423
	Interfish Wallsend Ind Est Cattedown Plymouth, Devon PL4 0RW, England	Andrew Pillar andrew@interfish.co.uk +44 (0) 1752 267261
	Larrasmendi Bi S.L. Eqidazu Kaia s/n 48700 Ondarroa (Bizkaia) Spain	Aitor Larrañaga erlaxi@larrasmendibi.e.telefonica.net +34 946830617
	Pesqueras Herbon	Daniel Magán aspin@pesquerasaspin.com
	SINTEF Digital Postboks 4760 Torgarden NO-7465 Trondheim NORWAY www.sintef.no/digital	Karl Henrik Haugholt Karl.H.Haugholt@sintef.no Petter Risholm Petter.Risholm@sintef.no +47 95 89 53 29

Table of contents

List of figures

Executive summary

Using the FishData infrastructure, we have developed a prototype information system for fisheries. The motivation behind the work was to:

- demonstrate the use of the FishData infrastructure developed in SMARTFISH H2020
- develop useful data analysis and decision support tools for the project stakeholders
- present a proof of concept for future fisheries monitoring and control systems based on automatic data collection

The system consists of a web portal with three applications, the first of which is called *Vessel catches*. It provides information about catch composition and catch efficiency for trawlers that have automatic catch recording and data acquisition systems such as those developed in SMARTFISH H2020. The application presents the information in the form of colour-coded trawl tracks in a map. The colour coding represents different interesting measures such as catch amount, catch effort, catch efficiency, and so on. In addition, the user has various options for deeper analysis of both catch composition and catch efficiency.

The second application is called *Fleet catches*. It provides information about catch composition for the entire fishing fleet based on publicly available data from official catch reports. The information is presented as a heat map of catch amounts overlaid on a geographical map. The application has facilities for filtering on geographic area, time period, fish species, fishing gears, and catch quantities.

The third application, *Forecast*, provides near-term forecasts of marine environmental conditions that are thought to affect fishing conditions. This includes physical conditions such as seawater temperatures and currents; chemical conditions such as salinity and nitrate concentration; and abundances of various types of algae and plankton which provide sustenance for fish. The application was co-developed with other projects.

These results are aimed at three stakeholder groups: fishers, fisheries management, and marine scientists who perform stock assessments.

Fishers can use the applications as decision support tools when planning their near-term fishing operations. They need to match the composition of catches to their allocated quotas for specific species, and they need to optimise their catch efficiency for the sake of profit. The applications can help them do both by providing recent information about catch compositions and catch efficiencies for different fisheries, and by providing information about factors that could affect future fishing conditions.

For fisheries management and stock assessment, the web applications are intended as spatiotemporal data visualisation and exploration tools. They can help to gain an overview of recent fishing activities, identify fishing hot spots, identify areas eligible for closing, and guide fisheries inspections. A remote data access API is provided for expert who need direct access to the underlying data to perform their own analyses.

We have concrete plans for exploitation of the results from SMARTFISH H2020 WP5 in collaboration with fishing companies, technology suppliers, and other research institutions. These include:

- extending the methods and applications to new fisheries
- contributing to the development of future fisheries monitoring and control systems
- helping to establish a commercial data sharing and decision support platform for fishers

1. SMARTFISH H2020 motivation and background

With an increasing pressure on marine resource extraction mounting with resultant calls for sustainability in the sector, SMARTFISH H2020 will develop, test and promote a suite of high-tech systems that will optimize resource efficiency, improve automatic data collection, provide evidence of compliance with fishery regulations and reduce the ecological impact of the sector on the marine environment (Figure 1).

SMARTFISH H2020 will exploit and further develop existing technological innovations in machine vision, camera technology, data processing, machine learning, artificial intelligence, big data analysis, smartphones/tablets, LED technology, acoustics and ROV technology. The developments will assist commercial fishers throughout Europe in making informed decisions during pre-catch, catch, and post-catch phases of the harvesting process.

SMARTFISH H2020 will also provide new data for stock assessment from commercial fishing and improve the quality and quantity of data that comes from traditional assessment surveys. This provides the potential for more accurate assessment of fish stocks and allows for the assessment of stocks that are currently data-poor and therefore difficult to manage. In addition, the project will access automatically collected catch data from the fisheries which will also allow for management regulations to gain higher compliance rates.

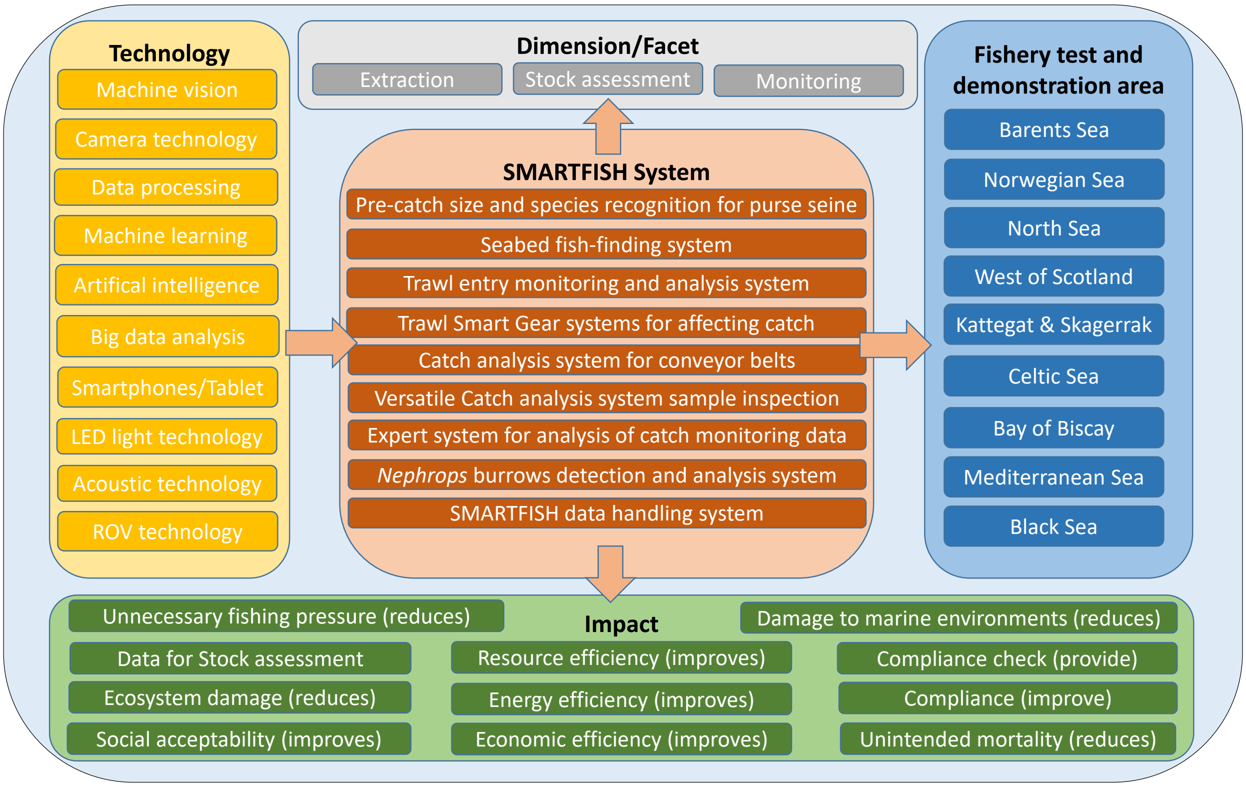


Figure 1: Conceptual structure of SMARTFISH H2020

1.1. Role of the deliverable

This document is the third and final deliverable from SMARTFISH H2020 Work Package 5 (WP5), *Development of data handling systems*. The two preceding deliverables were D5.1 *FishData system specification*, [1] which described the architecture of the FishData infrastructure, and D5.2 *FishData infrastructure*, [2] which provided additional implementation details of the prototype system. In the present document, we describe a prototype information system built on the FishData infrastructure to demonstrate its use and potential benefits to stakeholders.

1.2. Contributors

The following partners have contributed to this deliverable:

- *SINTEF Ocean* have performed the system development.
- *DTU Aqua* have carried out interviews with fisheries management and stock assessment experts to establish information needs.

1.3. Acronyms and abbreviations

Term	Meaning	Explanation
AIS	Automatic identification system	An international system for identifying and tracking vessels, originally intended for safety purposes, but nowadays used for all sorts of vessel tracking. Mandatory for large fishing vessels and increasingly used for smaller ones.
API	Application programming interface	An interface for interaction between computer programs. (Contrast with *GUI*, below.)
CPUE	Catch per unit effort	A ratio between catch size (in mass or number of individuals) and fishing effort (time spent, distance travelled, energy expended, etc.).
ERS	Electronic reporting system	A system for electronic reporting of catches to national fisheries authorities. Mandatory for vessels of length 12 m or more (with some exceptions) in Europe.
GUI	Graphical user interface	A visual interface for interaction between users and computer programs. (Contrast with *API*, above.)
VMS	Vessel monitoring system	A satellite-based system for monitoring of fishing vessels' location, course, and speed. Compulsory for all EU vessels of length 12 m or more.

2. Introduction

FishData is an experimental data acquisition and analysis infrastructure established by and for the SMARTFISH H2020 project. The hardware and software infrastructure itself has been described in two previous deliverables. [1, 2] In this document, we will present a prototype information system built on top of FishData—a system which provides information in both visual and programmatic form about catch efficiency and catch composition in fisheries, and forecasts of marine environmental conditions.

2.1. Motivation

We have multiple aims with this work: Firstly, we want to demonstrate that the *infrastructure* actually works as intended. That is, we want to show that FishData provides the necessary nuts and bolts for building systems that collect, process, and present a variety of fisheries-related data. Secondly, we want to demonstrate concrete examples of such applications because they are useful in their own right.

Ultimately, of course, our work should contribute to progress towards several of the SMARTFISH H2020 project's stated objectives. These are:

- Improve catch efficiency [and] catch composition […] in pelagic and demersal fisheries.
- Improve the accuracy and quantity of fish stock assessment data.
- Advance the monitoring and compliance of fishery regulations.
- Ensure that data collected onboard fishing vessels can be used quickly and efficiently for fisheries monitoring, management and [stock] assessment purposes.

2.2. Even more motivation

In the years since SMARTFISH H2020 was initiated, there has been increasing interest in, and a political push towards, a technological reformation of European fisheries monitoring, surveillance, and control (MCS) systems. A progressive example is the Norwegian Directorate of Fisheries' *FangstID*[1] programme, announced in early 2021. The aim of this programme is to introduce new technologies in the fishing fleet for recording resource extraction as early as possible—preferably in the gear, when the catch is taken aboard, or during on-board processing. [3] Another goal of FangstID is to phase out today's non-verifiable self-reporting of catches in favour of fully automated catch recording and direct quota deduction.

Similar changes are forthcoming across Europe. In 2018, the European Commission issued a proposal for a (now ongoing) revision and modernisation of the EU MCS system. The idea is to introduce electronic tracking of all fishing vessels, implement fully digital reporting of catches, and require video monitoring on a certain minimum percentage of vessels. [4] Furthermore, several of the proposed technical screening criteria for commercial fishing activities under the EU taxonomy point in the same direction. One example is criterion 3.2.2:

> "100% observers' coverage or Remote Electronic Monitoring (REM) is in place on board vessel to monitor compliance with harvesting criteria and better collection of data on by-catch". [5]

The combination of catch scanning technologies and data handling technologies developed in SMARTFISH H2020 work packages 4 and 5, respectively, seem like a perfect starting point for addressing these challenges and developing the fisheries monitoring and control systems of the future. We have taken this on as additional motivation for our work and kept it at the back of our minds while developing the systems described herein.

[1] A direct English translation of "FangstID" would be "CatchID".

2.3. Main outputs

We have focused on the calculation and presentation of two main types of information: *catch composition* and *catch efficiency*. We define what we mean by these terms and explain why they're relevant in the subsections below. The spatiotemporal aspect of these measures is emphasised in our work, both in the way we present them visually and in the way they're generated.

Some of our work has been restricted to demersal trawl fisheries. The main reasons for this are that the FishData-related activities in WP6 (*Testing, demonstration and promotion in Norwegian Sea and Barents Sea fisheries*) are centered around whitefish trawling, and that SINTEF Ocean have access to large and detailed data sets collected on trawlers in other projects which we've been able to use here.

2.3.1. Catch composition

By catch *composition* we simply mean the size and species distributions of the fish that get caught. Knowing the gotten or expected composition of catches in a certain area, at a particular time, using a specific gear, is important for several reasons.

For fishers, it is important because they have limited quotas and a limited bycatch allowance. If they exceed their limit for some species, they can be penalised. It may even prevent them from fulfilling their quota for other species, because different species often coexist and it can be practically impossible to catch one without also catching others. (This is referred to as a "choke" situation.) Knowing what catch compositions to expect at different times and locations can help fishers to target their fishing operations and thus optimise their quota utilisation. It is also useful for financial reasons, because different species and fish sizes fetch different prices in the market at various times. Being able to target one's fishing operations can thus be profitable.

Real-time monitoring of catch compositions is also important in a *dynamic fisheries management* context. Dynamic fisheries management differs from traditional static spatiotemporal management of fisheries (e.g. seasonal area closures and seasonal full-fishery closures) in that it operates on smaller temporal and spatial scales. Typically, it consists of closing small areas or "hot spots" for short periods of time whenever the bycatch in the area exceeds a certain threshold. [6] Other use cases within fisheries MCS include detection of illegal practices like high grading and other forms of illegal discards.

Finally, catch compositions are a key input in stock assessment, since it is to be expected that the composition of catches reflect the composition of the underlying fish stocks. (Note that what gets caught is also affected by gear selectivity, fishing tactics, and more, so the relationship is not linear.)

2.3.2. Catch efficiency

By catch *efficiency*, we mean how much fish (of the desired species and sizes) fishers catch relative to how much effort they put into catching it. A term that is commonly used, especially in a stock assessment context, is *catch per unit effort*, or CPUE. In general, CPUE is just catch size divided by effort. *Catch size* is typically expressed as the total number of individuals or mass of the fish. *Effort* can be the time spent on fishing operations, distance traversed during fishing operations, number of pots or hooks used, area swept by a trawl, and so on. Since fisheries differ significantly in their methods and gear, stock assessors use different CPUE measures for different fisheries and fish stocks. CPUE is traditionally assumed to be proportional to the abundance of fish, and it thus forms one of the main inputs to stock assessment models. [7, 8]

We have chosen to focus on two general measures of effort that are broadly applicable to a variety of fisheries: time spent on fishing operations and distance traversed during fishing operations. The latter is especially relevant for trawlers, where the distance, i.e. the length of a trawl track, is proportional to the seabed area and water volume swept by the trawl (assuming a more or less constant trawl opening). Thus, CPUE becomes a near-direct measure of the density of fish in the area.

For fishers, catch efficiency is an extremely important performance indicator, arguably the most important of all. It is a succinct measure of the relationship between the source of their income—the fish they catch—and their main expenses: payroll (roughly proportional to time spent) and fuel (roughly proportional to distance traversed).

2.4. Overview of results

The main results that will be presented in this document are two web applications contained in a fisheries information web portal. They are:

- *Vessel catches:* an application for visualisation and exploration of catch composition and catch efficiency based on automatically collected data from own and collaborating vessels
- *Fleet catches:* an application for visualisation and exploration of catch data based on official reports available from public sources

There is some overlap in the intended uses of these two applications, but they are complementary in several aspects:

- *Vessel catches* provides more detail for vessels where this is available (private and shared data), while *Fleet catches* presents data from the entire fishing fleet (public data).
- *Vessel catches* demonstrates things that can be done if fully automated catch recording and data sharing technologies are taken up by the fishing fleet, while *Fleet catches* demonstrates things we can do under today's catch monitoring regime.
- Because of the above differences, the two applications demonstrate somewhat different ways of visualising and working with the data.

We will also briefly describe a third application, *Forecast*, which is a kind of "weather forecast" for underwater conditions. This application has for the most part been developed outside SMARTFISH H2020, though the project has made crucial contributions to it. Nevertheless, we consider it to be part of a "suite" of decision support tools for fishers, together with the two above-mentioned applications.

The applications appear as pages within a unified web portal. The web portal provides the application framework, including a common login screen, a menu system, and reusable GUI elements.

2.5. Novelty

2.5.1. Visualisation and exploration of spatiotemporal catch data

We have argued that spatiotemporal catch efficiency and catch composition constitute highly useful information for both fishers, fisheries managers, and marine scientists. However, there is more than one way to obtain such data, and not all of them are equally accessible to all stakeholders. Fishers have access to the data from their own vessel(s), but few have systems in place for analysing the data on a larger spatiotemporal scale, let alone combine data from multiple vessels and locations to identify patterns in fishing potential. Official catch and landing reports may be publicly available information, at least in aggregate and time-delayed form. Usually, however, they are in the form of raw data tables and not in a form which is easily digestible to the average user.

Here, we demonstrate how recent catch and landing data from multiple sources, covering entire fleets of vessels, can be combined, visualised and explored using an intuitive graphical user interface (GUI). The interface allows for easy selection and filtering of data based on species, size, and more, and the results are shown in a map view. Users interested in the underlying numbers have the ability to drill down into detailed tables.

2.5.2. Automatic monitoring of catches and discards

Catch reporting today involves a fair amount of manual work. Fishers are involved in sorting and weighing the fish, and they enter the catch details along with times and locations in their logbooks for reporting purposes. As discussed in section 2.2, there is a need for automation of this process.

Here, we show how *automatically recorded catches* (assuming the use of technologies such as WP4's CatchScanner) can be associated with *automatically identified fishing operations* to produce highly detailed spatiotemporal catch composition and catch efficiency data that can in principle be made available to stakeholders mere moments after the catch has been scanned. Assuming that scanning takes place at an early stage of the process before anything has been discarded, e.g. at the inlet of the on-board process line in a freezer trawler, these data will constitute a detailed record of everything the vessel has extracted from the ocean. This includes landed *and* discarded[2] fractions of the catch, which are both highly relevant for fisheries management and stock assessment.

2.6. Challenges

We have encountered some challenges in obtaining data, both from vessels and official sources.

Concerning the former, our initial hope was to be able to install our data acquisition system[3] alongside a prototype CatchScanner (developed in WP4) on a commercial trawler, and leave them both running for some time period while the vessel was performing its fishing operations. This has turned out to be unfeasible, and the on-board tests of CatchScanner have so far been limited to a short trip where the scanner was tested manually with only a small sample of the vessel's total catches[4]. We felt we could not justify the time and costs associated with installing the data acquisition system on the vessel under these circumstances, given that we would get very little data in return for our efforts. This is what we've done instead:

- When developing algorithms for fishing activity detection based on vessel time series, we have used data from trawlers that are not involved in the SMARTFISH H2020 project. SINTEF Ocean have access to highly detailed data sets spanning several years, collected in the course of other research projects. We have retained the necessary rights to use these data sets in our research (though unfortunately not to share them). However, these data sets do not contain any catch information, which brings us to the next point.
- In order to have catch data for development and demonstration of the web applications, we have made a simple "trawling simulator". This can be used to generate fake data that are sufficiently realistic for our purposes. It has the added advantage that we have full control over parameters such as the number of fish species, the abundance of different species, and the fishing patterns of the vessels; not to mention the fact that we can generate as much data as we need on demand.
- In the hope that we will get some use out of the CatchScanner data set that was collected during the sea trials, despite not having our data acquisition system on board, we have also obtained the vessel's AIS track from the same period, courtesy of the Norwegian Coastal Administration. Nergård havfiske, who own the vessel, have provided us with the "ground truth" about when and where fishing operations were performed during this period. We thus envision using this as a (limited) testing set for the algorithms later on.

When it comes to obtaining official catch reports, we have been expecting the Norwegian Directorate of Fisheries to set up a new data provision service. This service will provide access to complete and detailed catch reports from Norwegian fisheries in near-real time. The Directorate have had plans to do this for some time, but the work has been delayed for various reasons (the Covid-19 pandemic being a significant one). At the time of writing, such a service is still expected to come into existence, but it is uncertain when.

For development and demonstration purposes, we have therefore made do with the data sets the Directorate publishes today. These include data from the Directorate's landing and sales slip register (hereafter referred to as "sales slips" for brevity's sake) and data from electronic reporting systems

[2] Both the EU and Norway now have general discard bans in place, but some species in certain fisheries are exempt from the landing obligation due to the practical difficulties of avoiding them.
[3] I.e., the FishData *on-board infrastructure* described in previous deliverables.
[4] These activities were carried out under WP6 and will be reported separately.

(ERS). Sales slips contain quality information about all catches taken by Norwegian vessels or delivered to Norwegian harbours. They are published for catches from 2011 up to a couple of days old, depending on the time between catch and landing. Unfortunately, the positions of catches are very coarse. On the other hand, ERS data contain accurate positions, but are only available for previous years. The work described here combines ERS data with sales slips. This gives priority to accurate catch position over recentness, as this will better demonstrate the possibilities in the expected future data. We consider this to be an acceptable solution for now. When the new service comes online, we will replace our current data scraper with one that downloads continuously updated reports and make the necessary changes to our processing pipeline to take advantage of the more detailed data.

2.7. Technical components

We will now give a brief account of the main components that make up the web portal: a *data server*, a *web server*, and a *front end*. Figure 2 shows how they relate to each other. The data server and the web server both run as services on the FishData infrastructure and make use of its low-level components and security mechanisms. (For details about these, we refer the reader to previous deliverables. [1, 2]) The front end runs within the user's web browser.

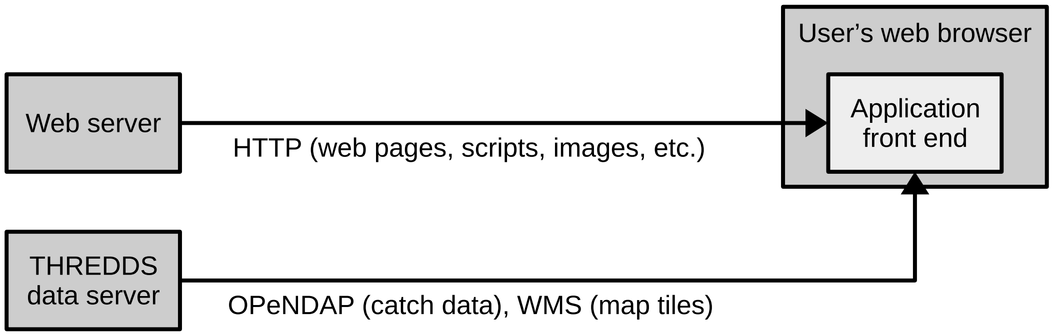


Figure 2: Overview of high-level components and their relationships.

For data distribution, we use *THREDDS Data Server* (TDS). TDS is a server software that provides data and metadata access for scientific data sets. It supports a variety of remote data access protocols, including OPeNDAP and WMS, which are the ones we use in our web applications.

TDS's data model is based on *Network Common Data Form* (NetCDF), a standard for array-oriented scientific data. The server software supports multiple data file formats for its underlying storage, but always presents the data using this common data model. The predominant storage type for TDS is a plain set of NetCDF files, which we also use.

Web Map Service (WMS) is a widely used standard protocol developed by the Open Geospatial Consortium in 1999. This standard specifies how a dataset can be requested and served in the form of multiple georeferenced image tiles to build an image covering a specified area at a specified zoom level. We use WMS to distribute spatiotemporal oceanographic data to the *Forecast* application. The data is provided by a simulation service and distributed through TDS. TDS works "out of the box" for this purpose as long as the data conform to certain standards. The main work in WP5 in this regard has therefore been to provide the necessary standards-compliant data, and to create legends and palettes to present the various parameters visually in the best way.

The front end applications are implemented in the *Vue.js* web application framework using the *TypeScript* programming language and *webpack* module bundler. Continuous deployment mechanisms are set up so that the applications automatically get rebuilt and redeployed when the code changes. The results are served by a service that runs the *NGINX* web server software.

3. “Vessel catches” application

Vessel catches is an application for visualisation and exploration of catch composition and catch efficiency data from vessels that have automatic catch recording and data acquisition systems.

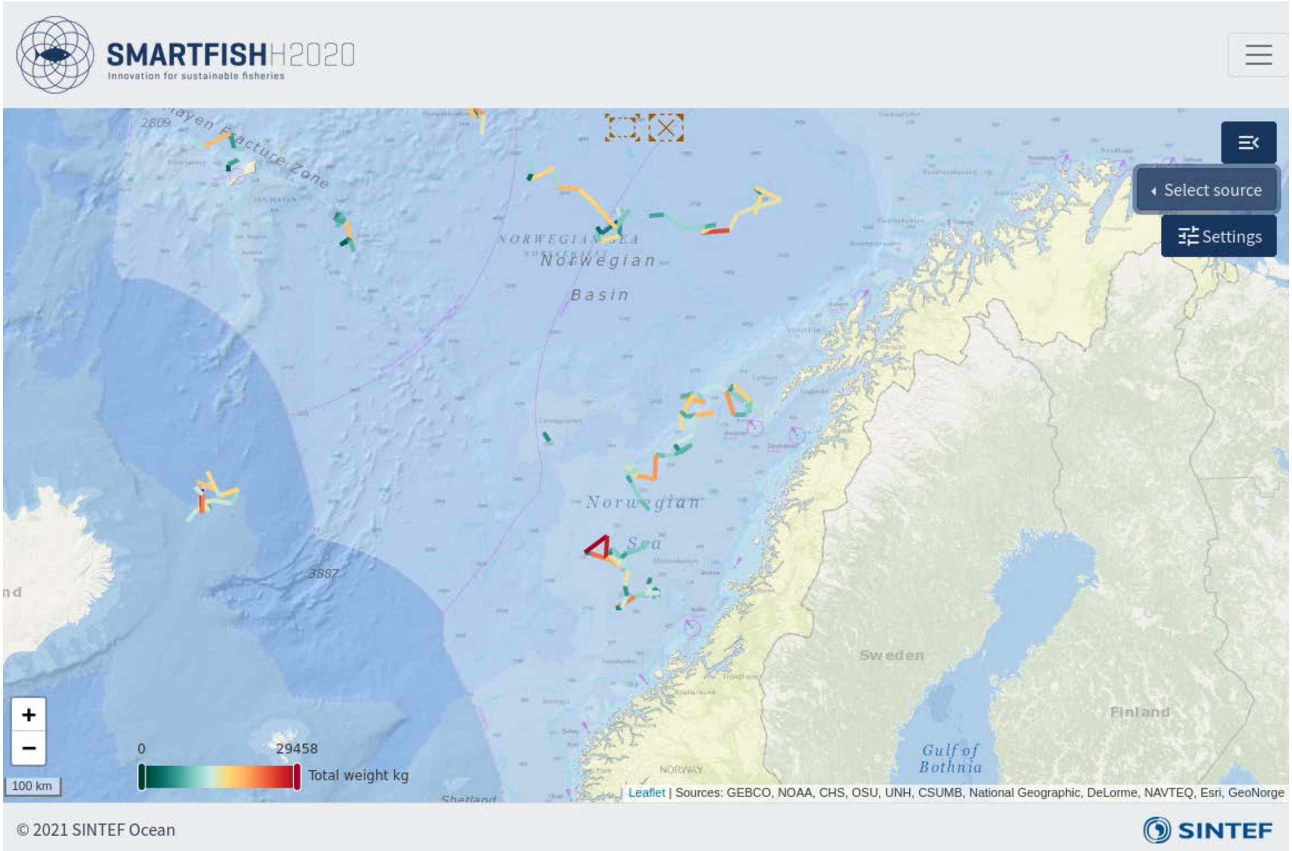
In the near-term future, the potential users are fishers that choose to invest in such technologies when they become available in the market. The application can then be used as a tool for studying and improving the performance of a single vessel, but its real power will be unleashed if and when fishers collaborate to combine the data from multiple vessels. Fishers already share information about fish observations and fishing conditions with each other through informal channels like telephone, e-mail, and social media. *Vessel catches* demonstrates a more systematic and convenient way to go about it. We fully expect that formal data-sharing collaborations among fishers will be established; see section 0 for a concrete example.

In the longer term, we expect that automatic catch recording systems will become mandatory in Europe, as discussed in section 2.2. In such an event, the potential user base of *Vessel catches* widens to include fisheries management and marine researchers. Fisheries management can use it as a tool for real-time monitoring and to guide inspections. For example, if a vessel is consistently reporting different catch compositions from similar vessels operating in the same area, it could be flagged for inspection. For researchers, we mainly see it as a data exploration tool, to be used prior to downloading data sets for detailed analysis. Facilities for direct data access are described in chapter 6.

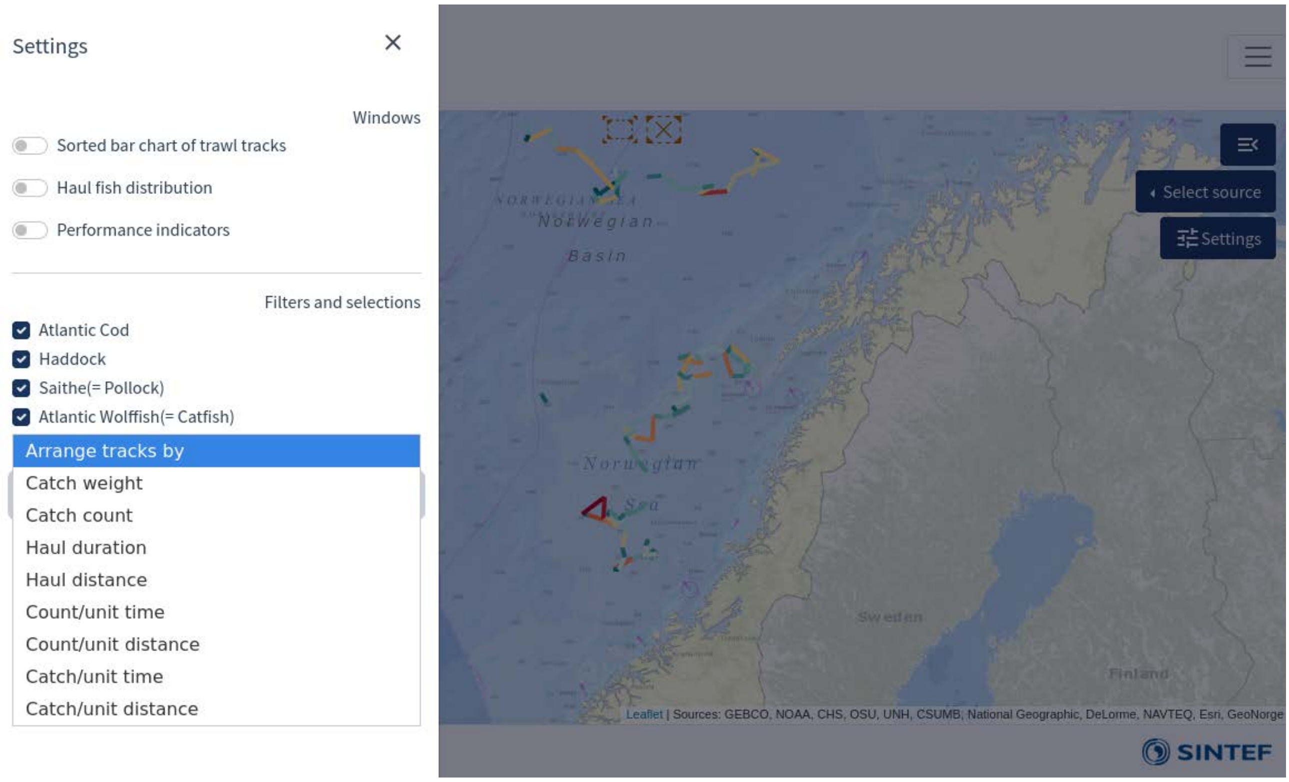
The current prototype version is designed for trawl fisheries, but it should be straightforward to extend it to other fisheries as well. We return to this issue in chapter 7.

3.1. User interface

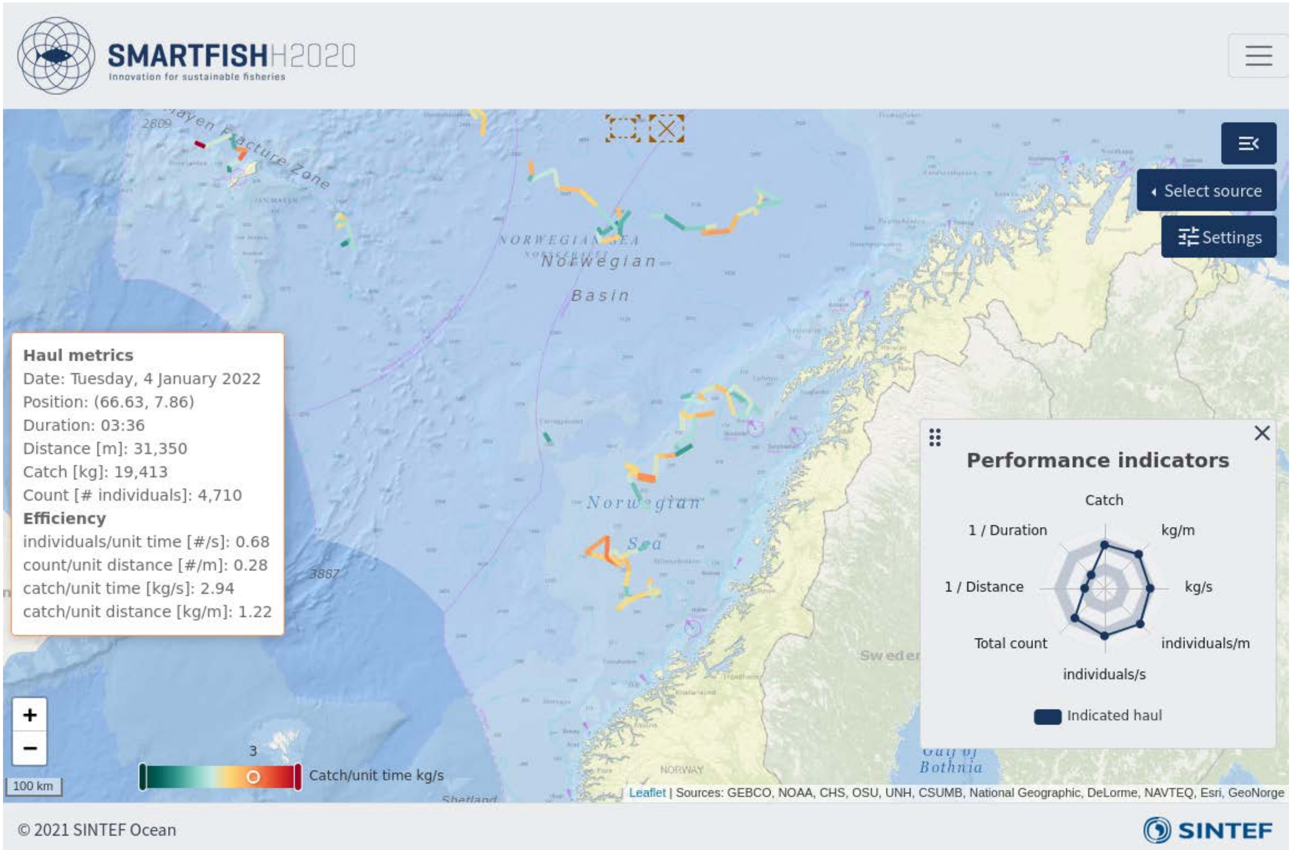
The screenshot below shows the initial view of the *Vessel catches* application. It shows recorded trawl tracks superposed on an ocean map. Each track is colour-coded according to some performance measure, which by default is the total catch weight in the haul. (The data shown in this screenshot and others in this section have all been generated by simulation, for reasons given in section 2.6.)



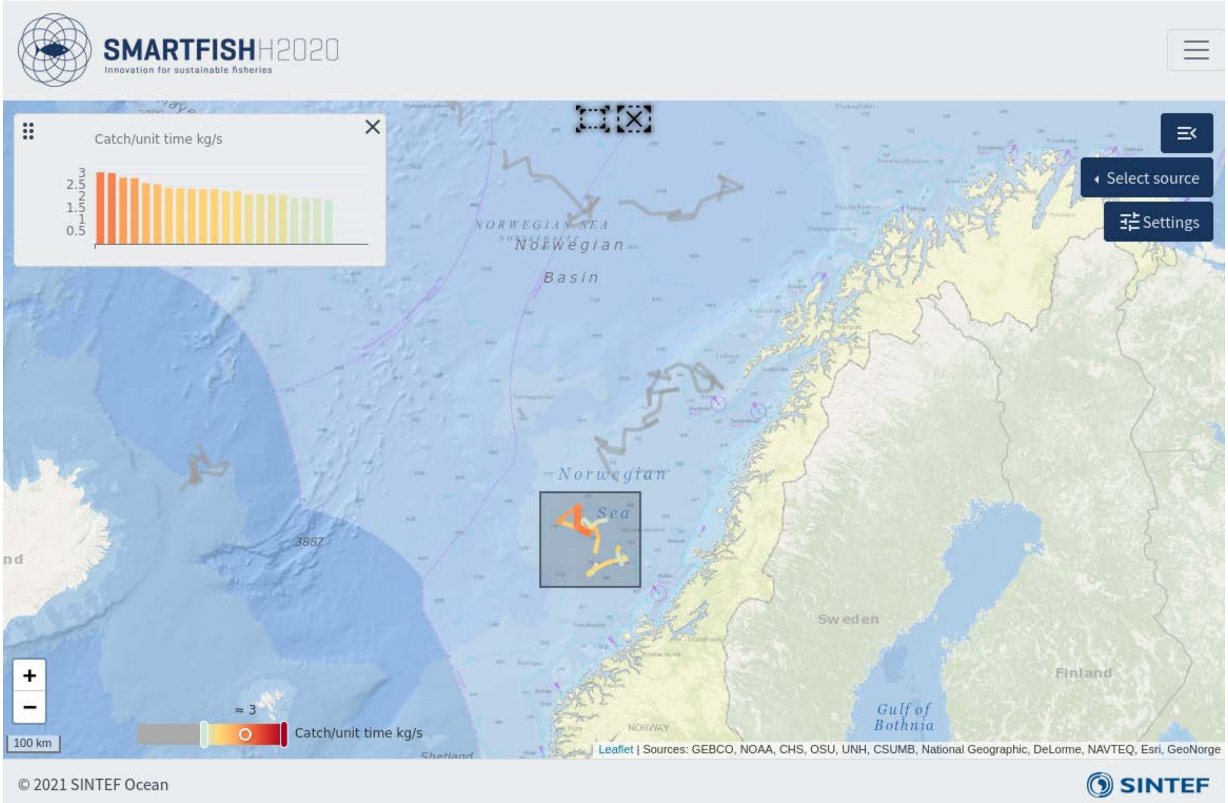
By clicking the *Settings* button, we open a menu where we can turn on various chart insets, select which species to include in the metrics and efficiency calculations, and which performance indicator to use for the colour coding of the trawl tracks.



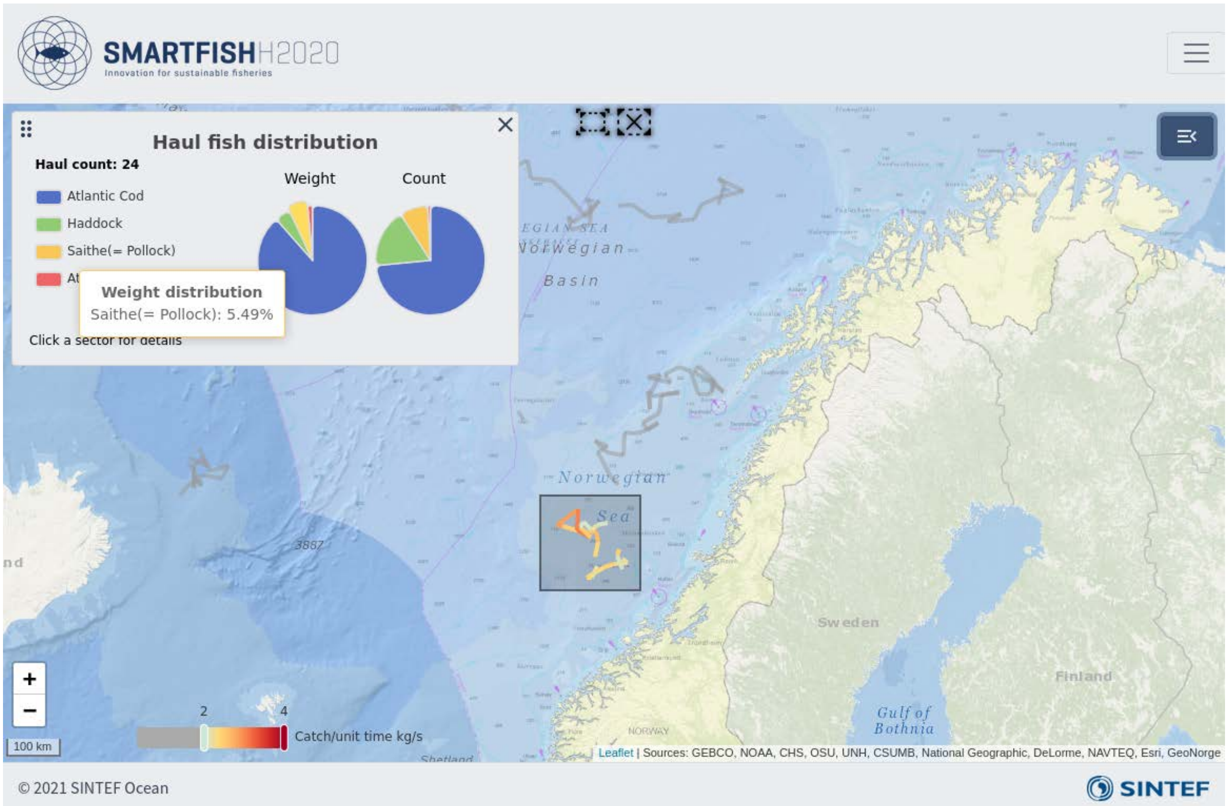
We now toggle the *Performance indicators* slider, which opens a radar chart that displays the various performance indicators. The values shown in the chart are relative to the "best" haul. We also select *Catch/unit time* as the performance indicator to use for track colouring and close the *Settings* window again. Hovering over a track shows detailed metrics and numerical efficiency measures for that haul.



Now, we close the *Performance indicators* chart and open *Sorted bar chart of trawl tracks* instead. This shows a bar chart of all trawl tracks, sorted according to our currently selected performance indicator. We then use the "box selection" control located at the top center of the view to select only a subset of the hauls that seem interesting. The bar chart automatically adjusts to only show the hauls in the selected area.



Then, we toggle the *Haul fish distribution* window to show an overview of the catch composition in the selected hauls. This window can show data for a single haul, for a set of selected hauls, or for all hauls in the data set. Currently it shows the selected ones, 24 in total. The pie charts show the relative proportions of the caught species in terms of weight and count.



As shown, we can hover over pie sections to see the percentages. We can drill further down by clicking a sector. This opens a new view, *Species distribution*. Here we see the detailed distribution of lengths and weights of that species in the selected hauls. The bottom left part is a heat map that shows the amount of fish within certain length and weight ranges. On the top left we show the marginal length distribution, and on the bottom right is the marginal weight distribution. The sliders on the top right control the colour ranges.

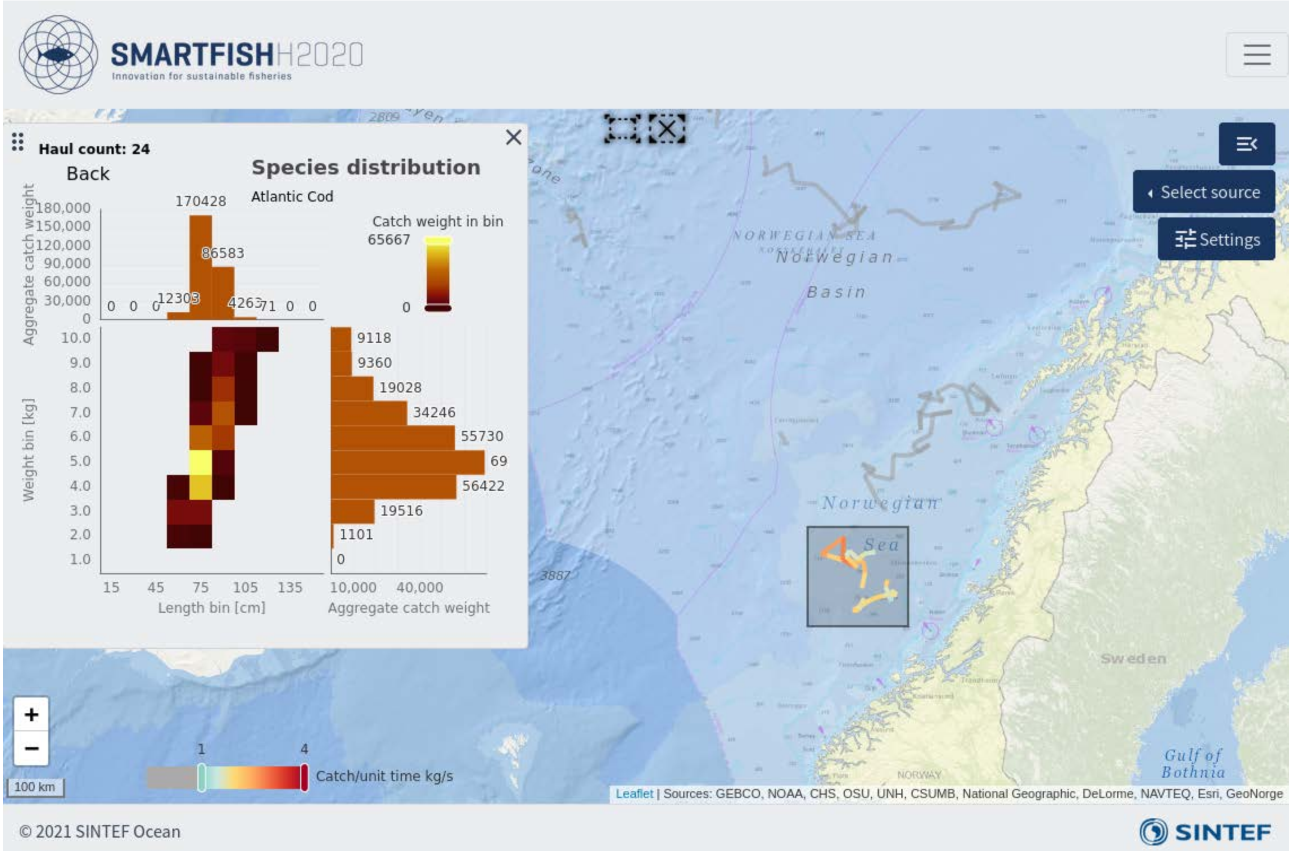


Figure 3 shows an overview of all visual components at once.

3.2. Data processing pipeline

The information presented in *Vessel catches* is based on two types of data which we'll refer to as *vessel data* and *catch data*. The former is acquired from the vessel's navigation instruments and other sensors, and is used to detect when the vessel is actively fishing. The latter comes from a catch recording device such as WP4's CatchScanner, and contains details such as size and species of the caught fish.

An overview of the data processing pipeline is shown in Figure 4. It consists of the following steps:

First, we use vessel data to infer when the vessel is and isn't trawling. For this, we have used *hidden Markov models*, which have been shown to work well for this purpose. [9] The primary input variable is the vessel's speed, which is easily obtainable from the GPS receiver on most vessels. Other variables, like main engine load, can be used when available to increase the accuracy of the detection algorithm. The output of this step is a series of times and locations where the vessel is trawling, i.e., its *trawl tracks*.

Next, the scanned fish are grouped into *hauls*, in most cases merely by associating each fish with the most recent trawl track. The more troublesome cases occur when the vessel takes in a new haul before the previous one is completely processed. In such cases, it may be impossible to tell exactly when scanning of one haul ends and the next one begins. Here, we must rely on simple heuristics like the fact that it takes a certain amount of time to transfer fish from the trawl into the receiving tanks, and from there on into the processing line. At any rate, misclassification of a few fish is unlikely to be a problem for our use cases, given that such hauls will usually be very close to each other in both time and space.

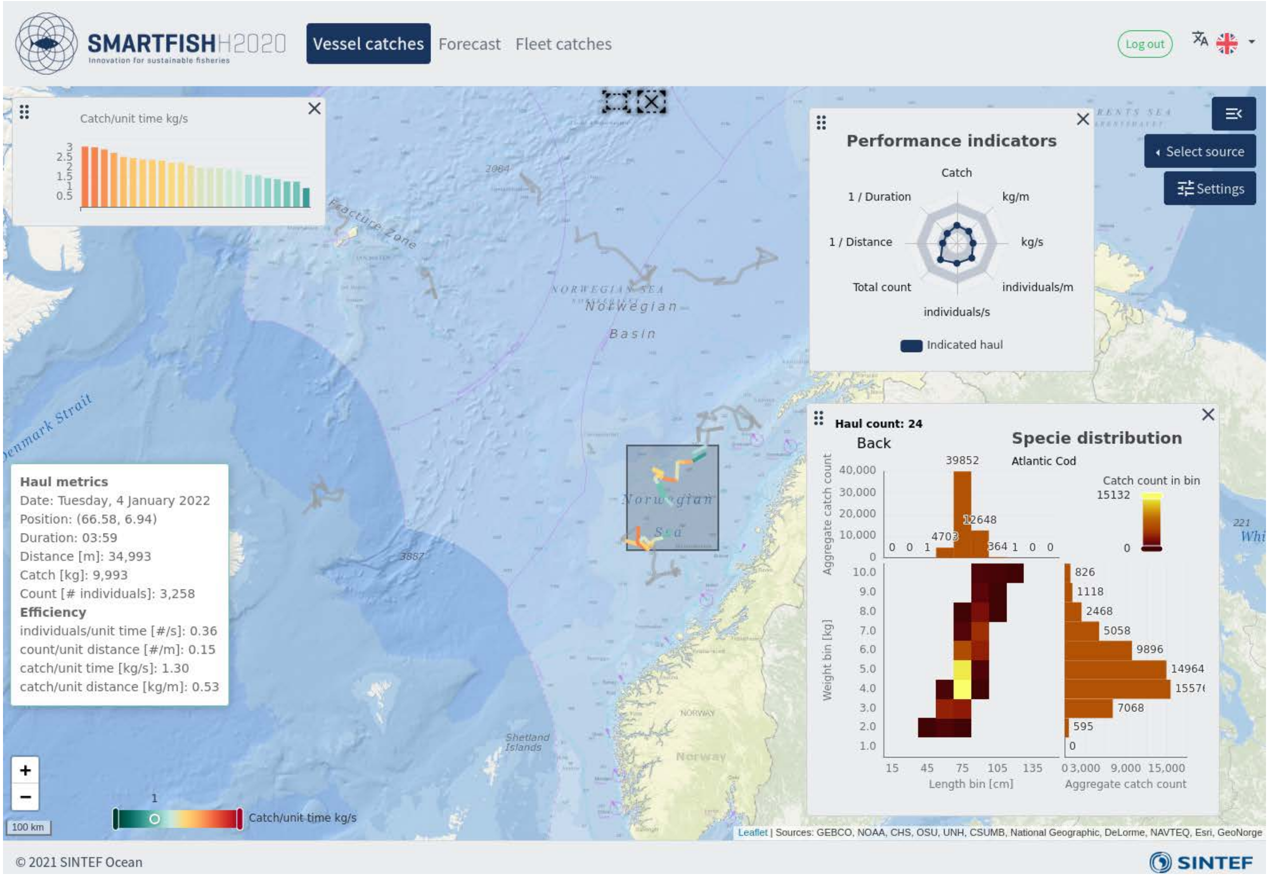


Figure 3: Overview of the *Vessel catches* application.

The output of this step is a list of (haul, trawl track) pairs, which we here collectively refer to as *spatiotemporal catch data*. These are stored in a database for later use in the web application.

The final steps are performed in the front end and consist of:

- filtering and selections according to user preferences
- calculation of derived quantities such as CPUE
- visualisation

CPUE can be calculated directly from the spatiotemporal catch data. The catch amount is given by the haul information, effort in terms of time is merely the difference between the start and end of a trawl track, and effort in terms of distance is calculated by summing over trawl track segments.

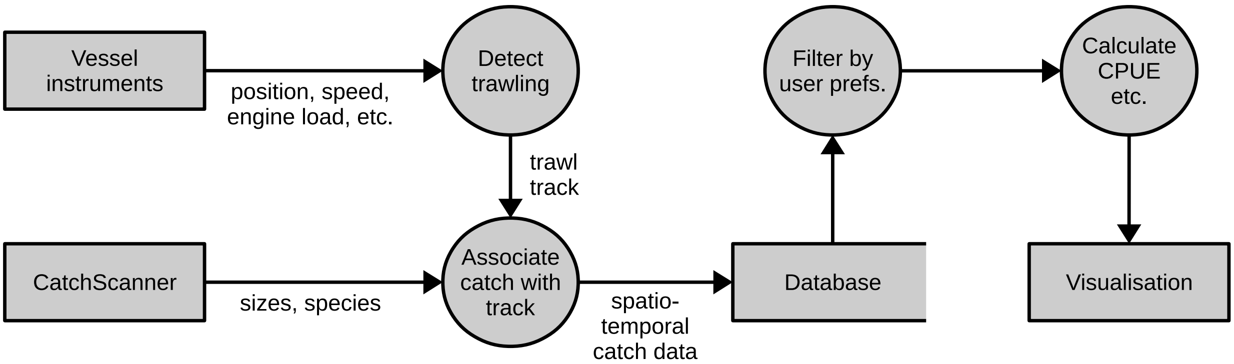


Figure 4: Data-flow diagram showing the data processing pipeline for the *Vessel catches* application.

4. “Fleet catches” application

The purpose of this application is to enable its users to monitor fisheries activity and study recent catches. This can be of use for fishers to find areas with the desired catch composition and for fisheries managers to find areas to consider for changes in regulations, such as closing areas for certain fishing gear due to the presence of unwanted species or high levels of bycatch.

4.1. User interface

Figure 5 shows the application’s user interface. It consists of three main areas: a map, a charts/filter sidebar, and a report area.

The map shows a heat map of the fisheries activity not hidden by the current filters. The heat map gives an impression of the geographical distribution of the catches. It makes no distinction between many small or a few larger catches; such information is displayed in other parts of the page.

The four charts to the right of the map show some characteristics of the activity not hidden by the current filters. The upper chart shows the catch for the individual days; the second one shows the species distribution; the third chart shows the total catch of the different fishing gear in use; and the lower chart shows the distribution of catch sizes. In the lower chart, the horizontal axis represents ranges of catch sizes and the vertical axis shows the total amount of catch for these sizes. Below each chart is a set of controllers that are used to filter what gets shown in the map, described further below.

The four tables in the lower part of the page shows some details of the catches. Each table can be limited to a certain maximum number of rows, selected by a drop-down menu next to the table title. The “Catch report” shows the most recent, individual catches. The “Area report” shows the areas with the most catch, in terms of aggregated amount of catch within each area. The “Species report” shows the most caught species. The “Gear report” shows the aggregated catch for the different fishing gear contributing most to the total catch. It is important to understand that the “Species report” doesn’t take the species filtering into account, and the “Gear report” doesn’t take the gear filtering into account.

4.1.1. Data filtering

An important aspect of this application is its ability to analyse the catch data using various filters.

Catches in a rectangular area can be selected by pressing the *Ctrl* key while selecting the area using the mouse. This will remove all catches outside this area from all data presentations on the page.

Data only within a time period can be selected using the selector below the “Catch per day” chart. The species to include can be selected by clicking on the legend of the “Species” chart. Specific fishing gears can be selected by clicking on the legend of the “Fishing gear” chart. The range of catch sizes to include can be selected below the “Catch sizes” chart.

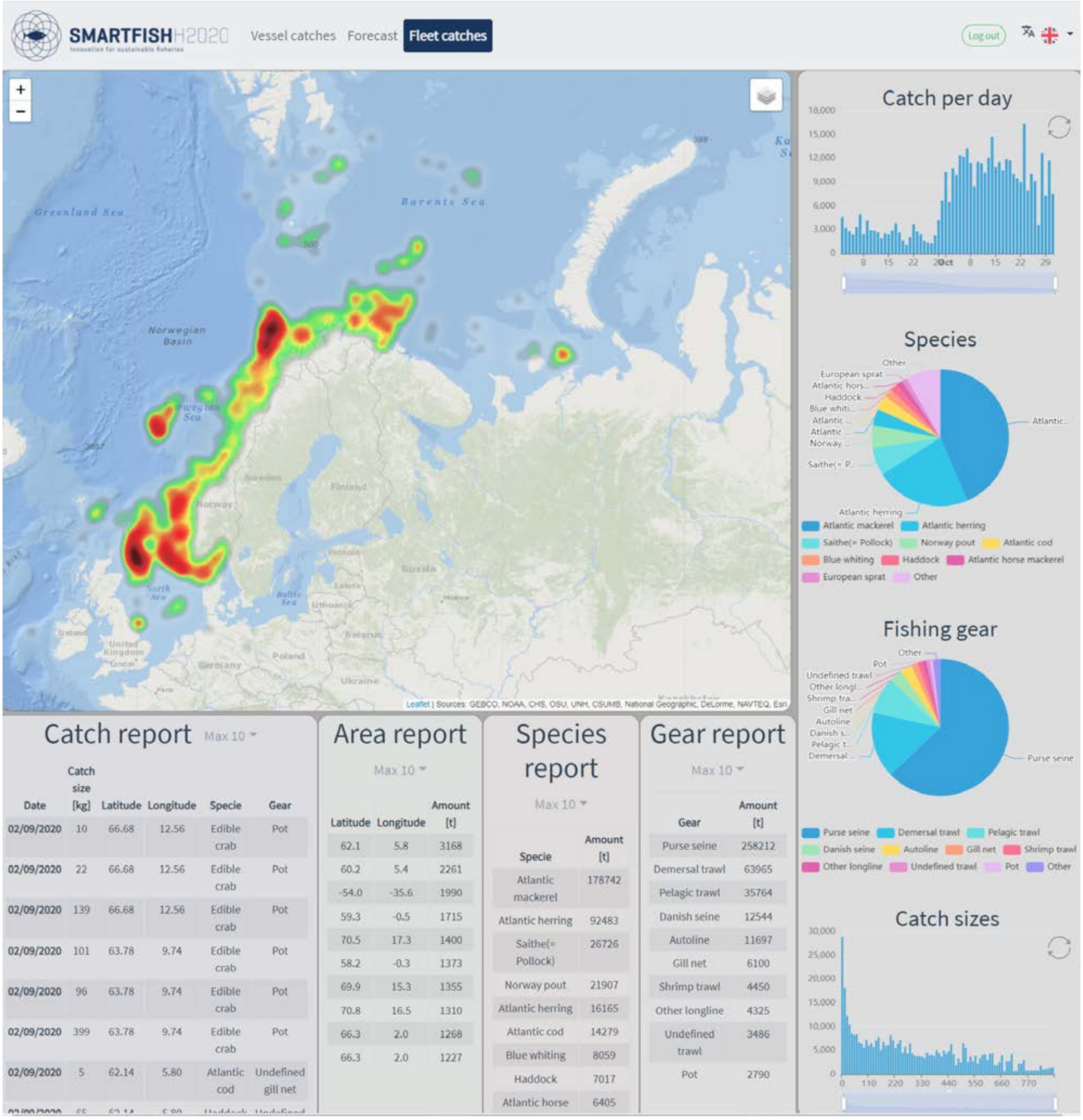
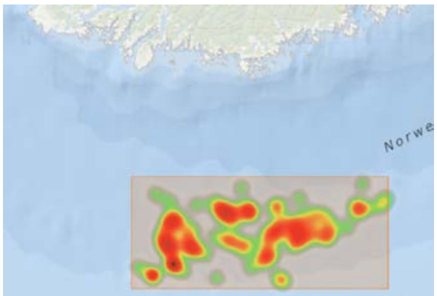


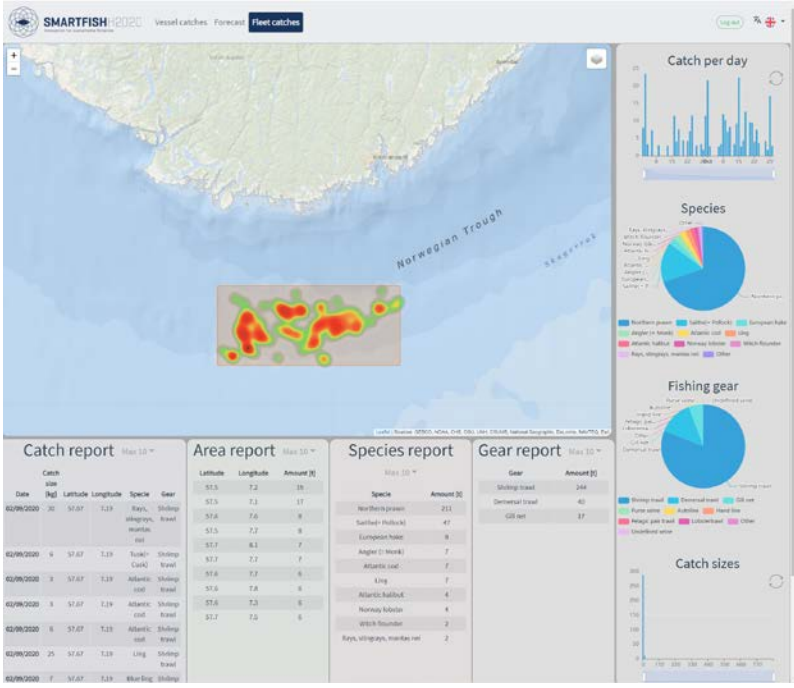
Figure 5: The *Fleet catches* application.

4.2. Example use case 1: Demersal fisheries

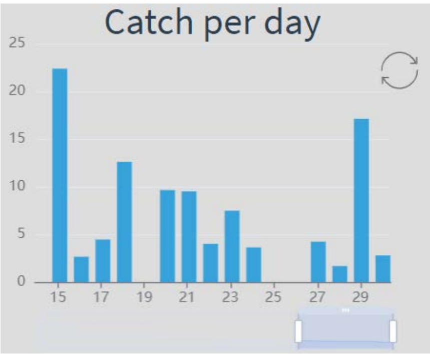
In this example, the user wants to investigate an area in terms of the catch distribution one might expect when using demersal fishing gear.

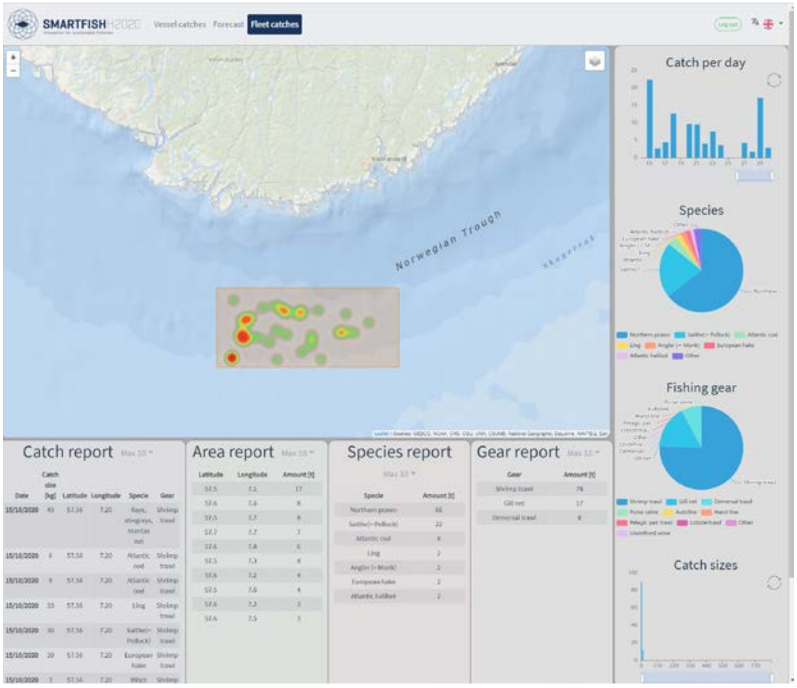
<table>
<tr>
<td>
Step 1: Select the area of interest.



</td>
<td>


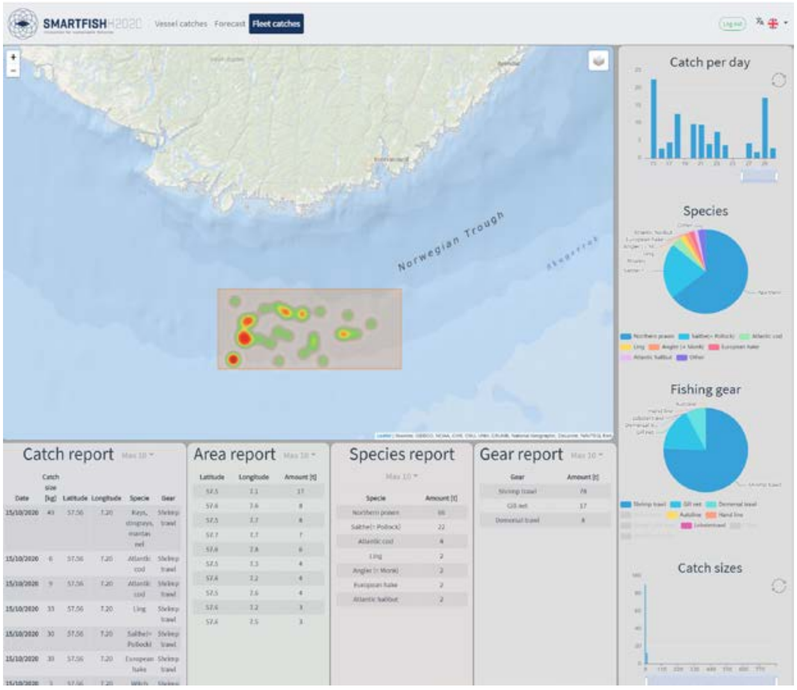
</td>
</tr>
<tr>
<td>
Step 2: Select the time period of interest.



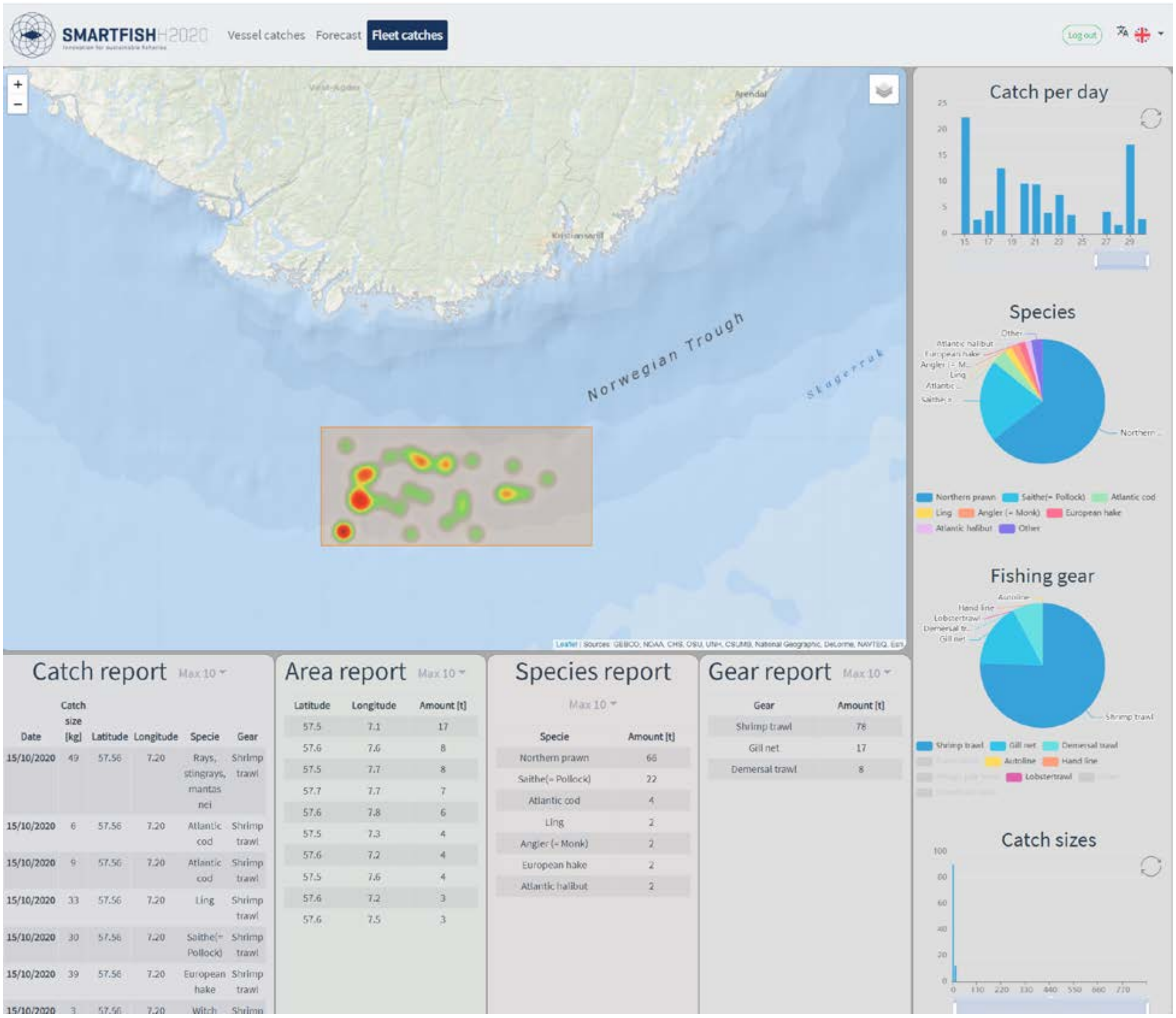
</td>
<td>


</td>
</tr>
</table>

Step 3: Select only demersal fishing gear.



Step 4: Consider the resulting reports and charts.



The analysis shows that in the last two weeks, shrimp trawl has been the dominant fishing gear used in this area. Also, it is seen that 66 tons of shrimp and 34 tons of demersal fish have been caught. Of the demersal fish, saithe is the dominant species.

To investigate further, we can look at only the shrimp trawling to find out how much bycatch these have had. The result is shown in Figure 6, where it can be seen that the total bycatch is about 14%.

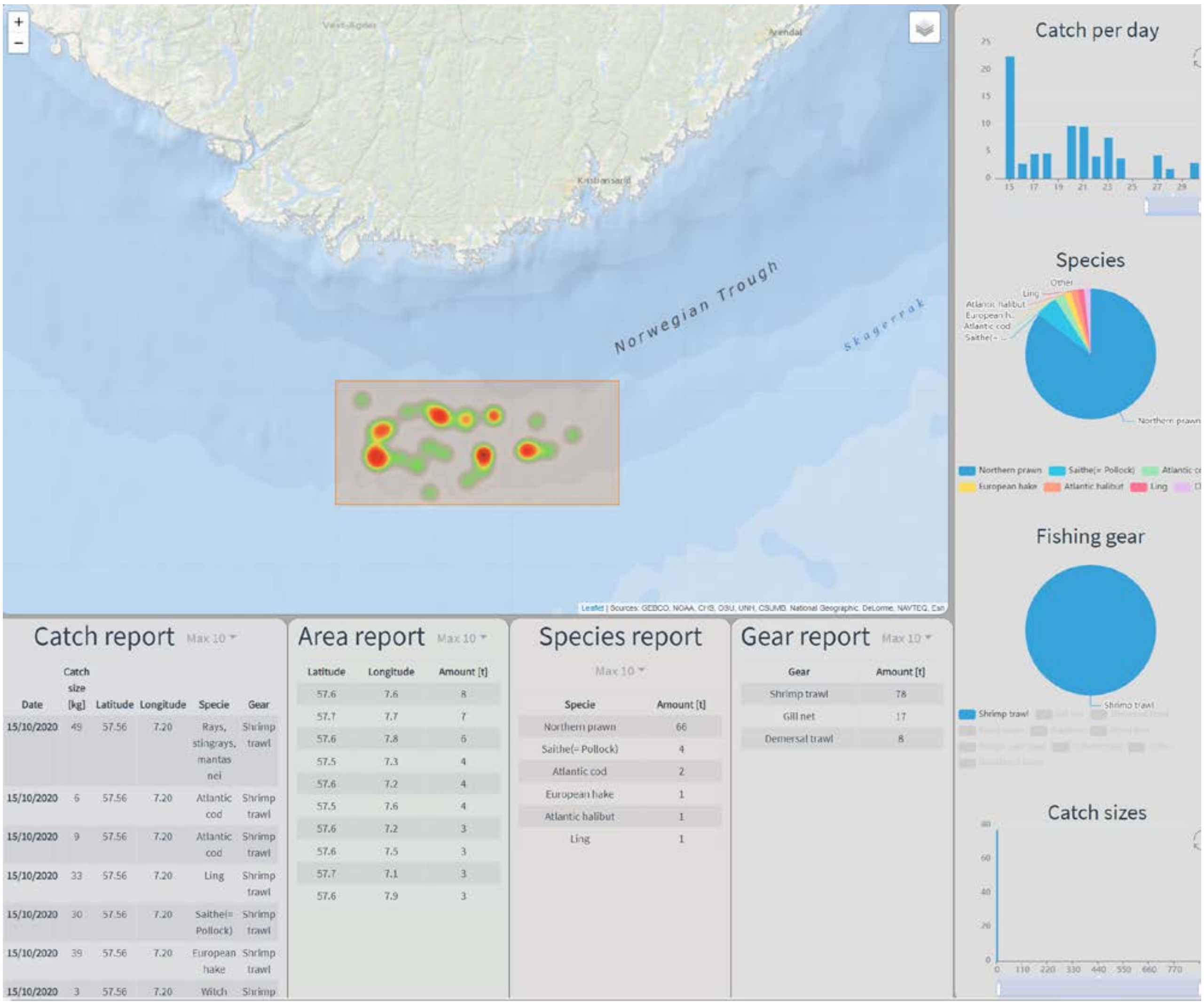
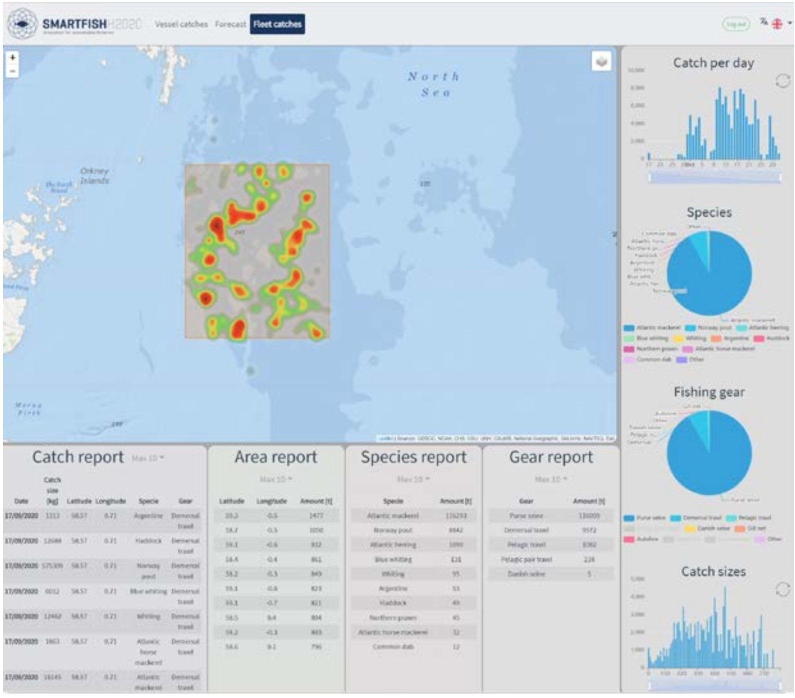


Figure 6: Catch and bycatch for shrimp trawling in a selected area over a two-week period.

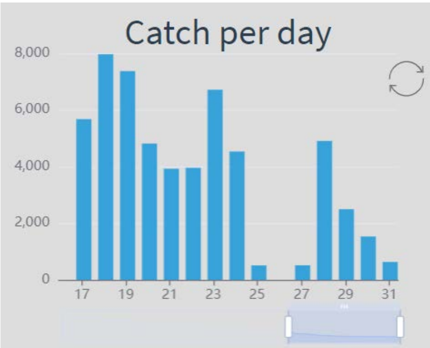
4.3. Example use case 2: Pelagic fisheries

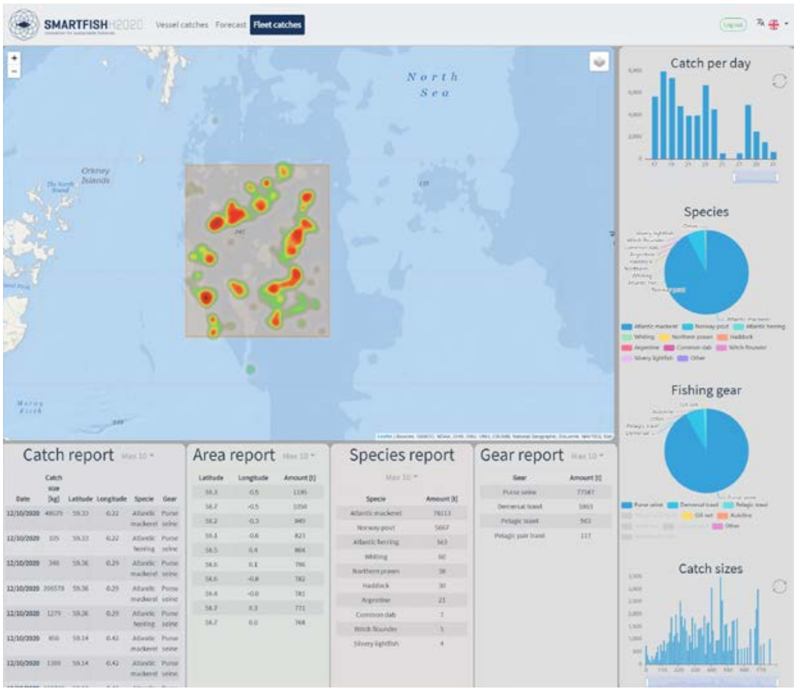
In this example, the user wants to investigate an area in terms of the catch distribution one might expect when using pelagic fishing gear.

Step 1: Select the area of interest.

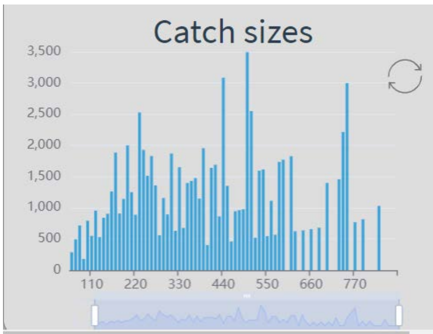


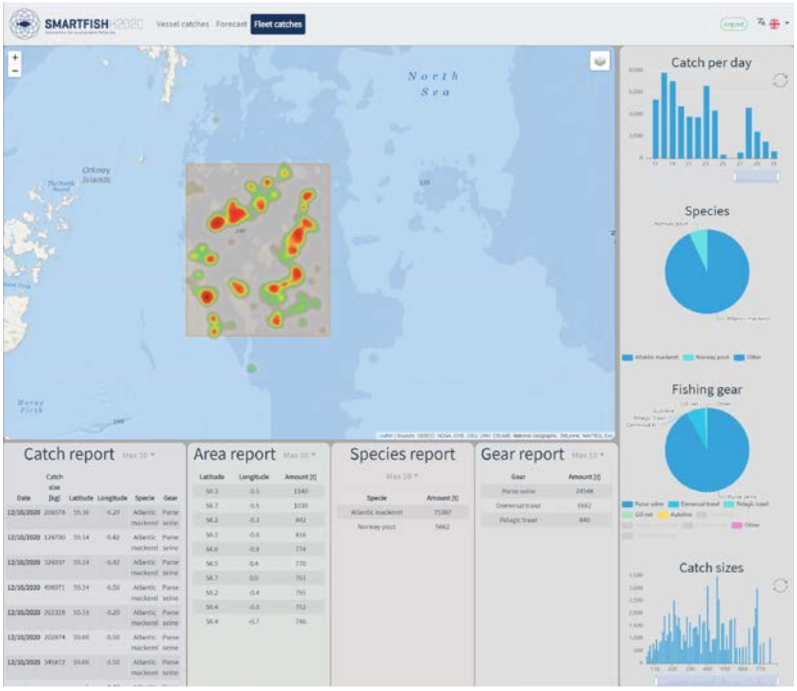
Step 2: Select the time period of interest.



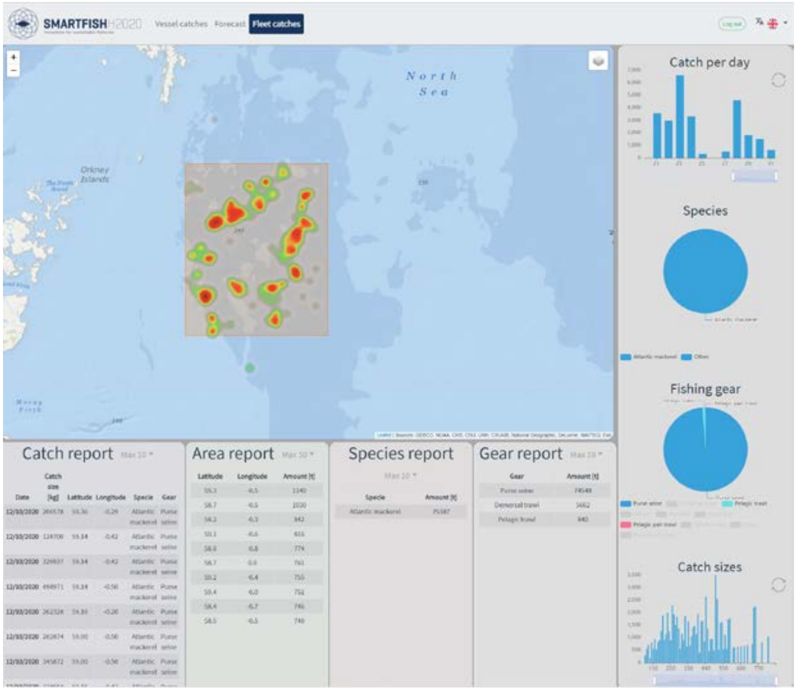


Step 3: Select the catch sizes of interest.

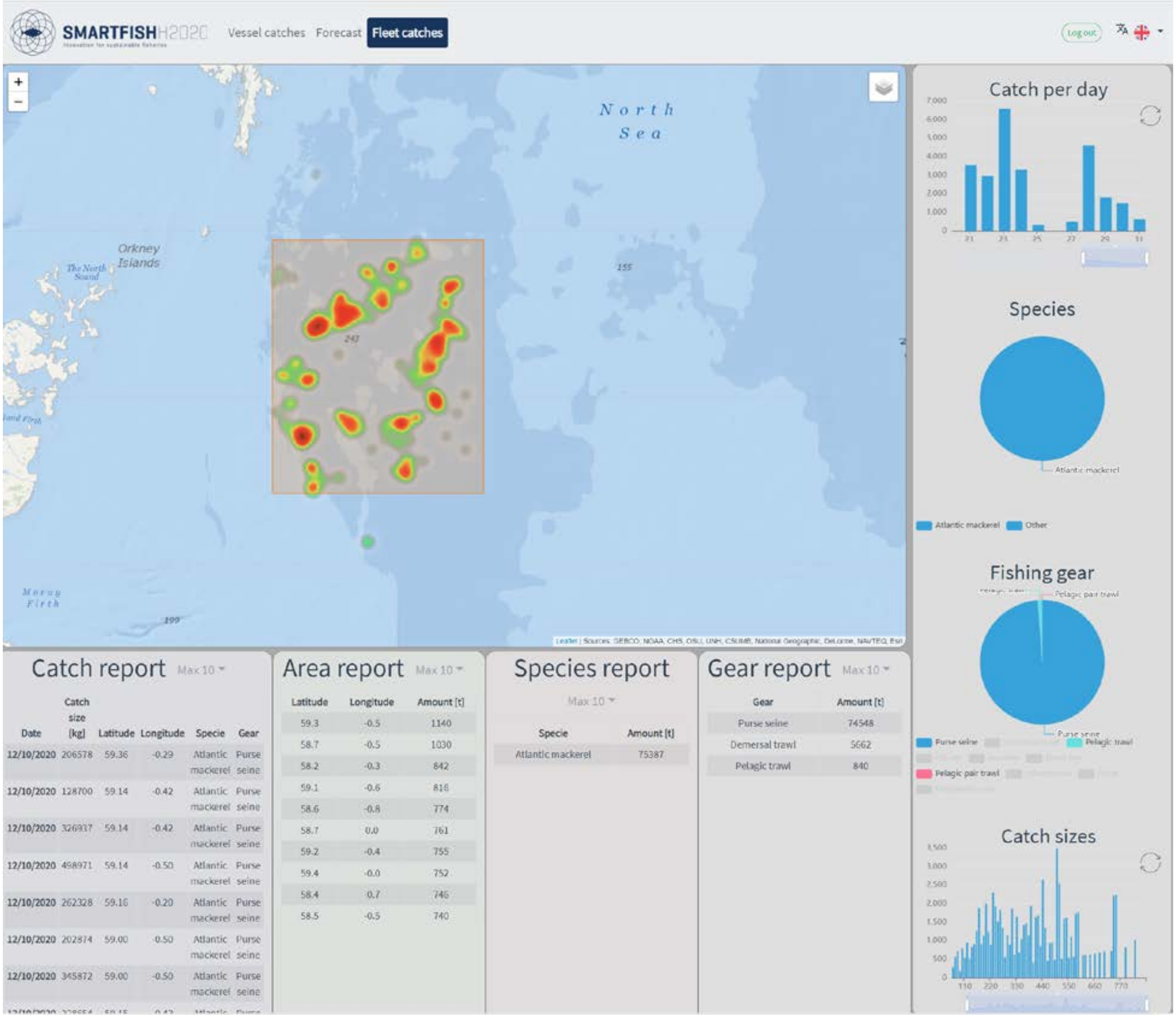




Step4: Select the pelagic fishing gear of interest.



Step 5: Consider the resulting reports and charts.



The results show that the pelagic catches are pure mackerel. Also, the catch amount is very large, which indicates a high catch potential for mackerel.

4.4. Data collection and processing

To give an accurate and detailed presentation of the recent fisheries, we need detailed and recent catch data with accurate positions. Such data are not yet available from public sources, but this is expected to change in the near future, as we discussed in section 2.6. Until then, we use older data for development and demonstration purposes.

Specifically, we have used two data sets published by the Norwegian Directorate of Fisheries. The first is based on sales slips and contains detailed information about catches. The structure of the data is such that the total catch is divided into multiple "lines", where each line describes one catch fraction. These data only contain very coarse catch locations.

The second data set consists of ERS messages and has been used to improve the position accuracy of catches whenever possible. ERS messages contain accurate catch positions, but there is no definite way to match them with sales slip lines. We have therefore assumed that all the fish described in a sales slip line was caught in the position where the vessel reported the largest catch since the previous sales slip. Multiple sales slips can be made for each fishing trip. We have deemed two sales slips to correspond to the same trip if they are less than 12 hours apart.

We considered using available data from vessel monitoring systems (VMS) to further improve the precision of catch locations, but the diversity in vessels and fishing gear made this task too resource demanding at this stage. Since the inaccuracies in position will not be present at a later time when more data are released, and users in fisheries management probably have access to more precise data already, it did not seem sensible to spend resources on this. For this reason, the described procedure was found to be sufficient at this stage and for demonstration purposes.

The sales slips and ERS data are combined into NetCDF files, one for each year and one containing the complete data set. As the purpose of the web portal is to give updated information down to individual catches, we have chosen to focus on presenting a subset of the data in the web portal, and for demonstration purposes we display two months of data. There lies a large potential in tailoring these data to its use, either by increasing the information to include e.g. vessel type or vessel size, or by grouping data to facilitate the handling of larger time spans.

5. “Forecast” application

Forecast is an application for predicting marine environmental factors that are thought to be related to fishing conditions. These include:

- *physical conditions*: seawater temperature, ice thickness, ice fraction, freezing rate, surface elevation, depth, current, and wind speed
- *chemical conditions*: nitrate concentration and salinity
- *abundance of algae and plankton*: diatoms, flagellates, ciliates, and copepods

The primary use case for the application is as a decision support tool for fishers planning their near-term operations. In other words, it can be used more or less in the same way as a weather forecast, except that instead of showing weather conditions, it shows the expected conditions under water. A screenshot is shown in Figure 7. The forecasts currently cover a large section of the North Atlantic and Arctic oceans, and the prediction horizon is 48 hours into the future.

Several projects have contributed to the development of this application:

- *SINMOD*, [10] the ocean model system that feeds the application with data, has been in development for about 40 years [11] through a large number of projects.
- The first proof of concept for the *Forecast* application was developed in *DataBio*[5]. [12]
- SMARTFISH H2020 has provided the application with a robust, stable infrastructure on which to run and an improved mechanism for distributing data from back end to front end.
- The current version of the GUI was developed in the *SFI Harvest* project, which is also where development will continue. (See section 7.2 for more on this.)

Since SMARTFISH H2020 has mainly contributed with infrastructure, which is not the focus of this report, we won’t go into details about the front end and its uses here. Still, we consider it to be part of the “suite” of fisheries decision support tools and have therefore chosen to include it in the web portal.

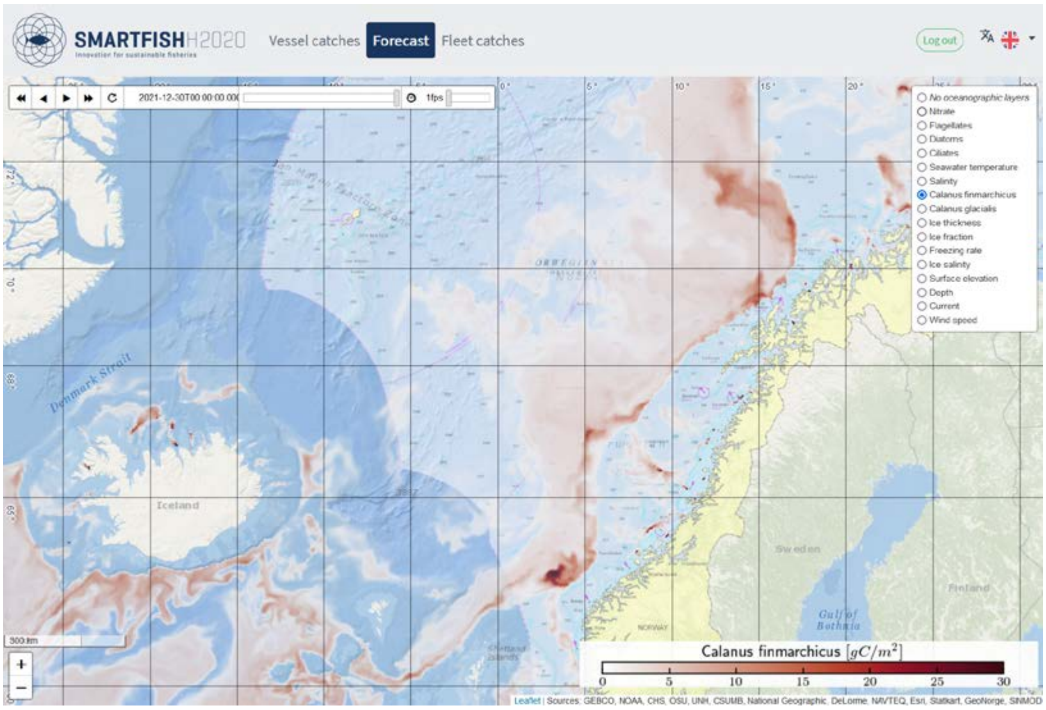


Figure 7: A screenshot of the *Forecast* application showing the expected density of *Calanus finmarchicus*, a zooplankton species which is an important source of sustenance for many commercial fish species. Screenshot taken 2021-12-28, two days before the forecast date.

[5] *Data-driven Bioeconomy* (2017–2019), funded by the European Union’s Horizon 2020 framework programme under grant agreement no. 732064.

6. Remote data access

Certain types of expert users will need access to the underlying data presented in the web applications. They could be researchers who want to perform their own analyses or software developers building new applications. The web portal's structural design (see section 2.7) makes it straightforward to support this use case. Users can simply access the data in the same way that the front end does, since it's based on well-defined, open protocols. The relevant data access protocols are all based on HTTP, so we even use the same access control mechanism as for web content (OAuth).

We have already mentioned that the data distribution service is based on *THREDDS Data Server*, whose data model is an extension of NetCDF. NetCDF is designed to enable self-documenting data sets by providing rich metadata facilities. In addition, various scientific communities have developed conventions for how to structure and describe different types of data. The use of conventions makes data self-descriptive in a machine readability sense too, enabling greater software interoperability. The *CF Metadata Conventions* [13] is the most well-established standard.

Listing 1 shows the structure for the data set used by the *Vessel catches* application, as an example. The structure is designed to be compatible with the CF conventions. The application accesses these data using the OPeNDAP protocol.

Listing 1: Data structure used for *Vessel catches*, described in the NetCDF Common Data Language. Lines starting with a double slash (//) are explanatory comments and not part of the data structure.

```
netcdf catch_stats {
    attributes:
        :Conventions = "CF-1.7 SFH/FishData-0.2";

    dimensions:
        track_point  = [total number of points in all tracks associated with hauls];
        haul         = [number of hauls];
        uuid_byte    = 16;
        species      = [number of species];
        length_class = [number of length bins];
        weight_class = [number of weight bins];

    variables:
        // The first five variables define the trawl tracks using the CF
        // conventions' contiguous ragged array representation.

        ubyte haul_uuid(haul, uuid_byte);
            haul_uuid:long_name = "A UUID for each haul (big-endian byte order)";
            haul_uuid:cf_role = "trajectory_id";
        int track_size(haul);
            track_size:long_name = "Number of points in each track";
            track_size:sample_dimension = "track_point";

        double time(track_point);
            time:long_name = "Time";
            time:standard_name = "time";
            time:units = "seconds since 1970-01-01 00:00:00";
        double lat(track_point);
            lat:long_name = "Latitude";
            lat:standard_name = "latitude";
            lat:units = "degrees_north";
        double lon(track_point);
            lon:long_name = "Longitude";
            lon:standard_name = "longitude",
            lon:units = "degrees_east";

        // Next, some variables that define the species characteristics.
        // Note that the number of length and weight bins is the same for all species.
        // Let's call it N and use B as shorthand for the bin size.  The binning
        // intervals are then [0,B), [B,2B), and so on up to [(N-1)B,+∞).

        string species_code(species);
            species_code:long_name = "Species code";
```

```
        float length_bin_width(species);
            length_bin_width:long_name = "Bin width for individual fish lengths";
            length_bin_width:units = "meter";
        float weight_bin_width(species);
            weight_bin_width:long_name = "Bin width for individual fish weights";
            weight_bin_width:units = "kilogram";

        // Finally, the variables that contain the catch information.

        int fish_count(haul, species, length_class, weight_class);
            fish_count:long_name = "Number of fish";
        float fish_weight(haul, species, length_class, weight_class);
            fish_weight:long_name = "Accumulated fish weight";
            fish_weight:units = "kilogram";

        // Additional, optional information about each haul can be provided
        // in variables whose primary dimension is 'haul'.  The following
        // have predefined names and types (but others can be added if needed):

        float track_length(haul);
            track_length:long_name = "The length of each track";
            track_length:units = "meter";
        float haul_duration(haul);
            haul_duration:long_name = "The duration of each fishing operation";
            haul_duration:units = "second";
}
```

7. Summary and outlook

In this document, we have presented an information system for fisheries. It consists of three applications:

- *Vessel catches*, for visualisation and exploration of detailed catch efficiency and catch composition data based on automatic catch recording systems
- *Fleet catches*, for visualisation and exploration of fleet-level catch data based on public catch reports
- *Forecast*, for predicting marine environmental factors that are thought to affect fishing conditions.

The applications and the surrounding web portal are based on the FishData infrastructure.

One of SMARTFISH H2020's objectives is to improve the catch efficiency and catch composition in pelagic and demersal fisheries. WP5's contribution to this is to give fishers new tools to quantify and visualise their own and other fishers' performance in terms of these measures. The fishers can then use this information as decision support, combining it with their own experience and expertise, in short-to-medium-term planning of their fishing operations.

Other objectives include improving the accuracy and quantity of fish stock assessment data, advancing the monitoring and compliance of fishery regulations, and ensuring that data collected onboard fishing vessels can be used quickly and efficiently for fisheries monitoring, management, and stock assessment purposes. We believe that the implementation of automatic catch recording, automatic fisheries activity detection, and real-time reporting and monitoring are key steps in achieving these objectives. In SMARTFISH H2020 WP5 we have demonstrated how the "information pipeline" can be set up—starting with raw data recorded at the point of resource extraction using technologies like WP4's CatchScanner and other vessel sensors, and ending with useful, actionable information to on-shore stakeholders.

Now, it is time to discuss where we go next.

7.1. Further work in SMARTFISH H2020

We still have some budget remaining for WP5. This will mainly be used for activities that are in the intersection between WP5 and WP6 (*Testing, demonstration and promotion in Norwegian Sea and Barents Sea fisheries*). This includes further testing and refinement of analysis methods, improvements and bug fixes in the software, demonstrations to stakeholders, and so on. We also have plans for scientific publication of some of the results.

7.2. Further research

There are many possibilities for extending the work done in WP5 in various directions. Here are some examples:

- We envision including catch *quality* data in addition to efficiency and composition. This includes things like length/weight ratio, HG-/round-weight ratio[6], fat content, product condition, conservation method, etc. All of this is useful information to fishers; some of it to scientists. Future versions of the CatchScanner may be able to estimate some of these variables, and others may be available from obtainable sales data.
- Any trawl-specific results should be extended to other fisheries as well. Some of our work will generalise quite easily, but given the wide variety of fishing methods and gears, some analyses will require specialised methods (e.g. fishing activity detection [9]).

[6] "HG" stands for "headed and gutted" and refers to the end product of typical onboard processing on a freezer trawler. "Round weight" is the weight of the whole fish, i.e., before processing.

- We want to establish systems for automatic collection and analysis of data from fishing vessels' hydroacoustic instruments. This could provide information about the biomass in the vincinity of the vessels, rather than just what they catch. This would be useful information to fishers (who already share this kind of information informally and inefficiently using telephone, e-mail, and social media). If it was implemented at a large scale across the fleet, it would drastically increase our knowledge about marine biomass, which would be beneficial for both stock assessment and marine research in general.

As for concrete plans, the most direct continuation of SMARTFISH H2020 WP5 will take place in *SFI Harvest*. This is a *centre for research-based innovation* focusing on sustainable harvesting of low-trophic resources such as *Calanus*, krill, and mesopelagic fish. The centre is led by SINTEF Ocean and is funded by the Research Council of Norway for a period of 8 years (2020–2028). Its nearly 30 participants include research institutes, universities, industry, and governmental and non-governmental organisations, national and international. The results from SMARTFISH H2020 WP5 will be exploited in SFI Harvest's Research Area 1, *Survey technology*, to enable marine data collection from ships of opportunity; and in Research Area 3, *Decision support*, to develop decision support systems for fishers. For more information, see the centre's web site at sfiharvest.no.

Several of the SMARTFISH H2020 partners also have plans for continued research and development of digital technologies for fisheries monitoring and control, building partly on results from WP4 and WP5. The research topics we have in mind include:

- automatic catch recording and monitoring in all types of fisheries
- automatic fishing activity detection
- detection of anomalies in catch and fishing activity data, to guide fishery inspections and scientific investigations
- use of blockchain technologies to share data in a secure, transparent and traceable manner between stakeholders.

Both national and EU grant proposals on these topics have been submitted, and more are in preparation.

References

[1] L. T. Kyllingstad, K.-J. Reite and P. G. Auran, "SMARTFISH H2020 D5.1: FishData system specification," SINTEF Ocean, Trondheim, 2019.

[2] L. T. Kyllingstad and K.-J. Reite, "SMARTFISH H2020 D5.2: FishData infrastructure [Open access revision]," SINTEF Ocean, Trondheim, 2021.

[3] Norwegian Directorate of Fisheries, "Hva er FangstID-programmet? (Translation: 'What is the CatchID programme?')," 2021. [Online]. Available: https://www.fiskeridir.no/Yrkesfiske/fangstid/hva-er-fangstid-programmet. [Accessed 23 December 2021].

[4] I. Popescu, "New EU rules on fisheries control," *At a Glance,* March 2021.

[5] Platform on sustainable finance: Technical working group, "Part B – Annex: Full list of Technical Screening Criteria," 2021.

[6] D. C. Dunn, S. M. Maxwell, A. M. Boustany and P. N. Halpin, "Dynamic ocean management increases the efficiency and efficacy of fisheries management," *PNAS,* vol. 113, no. 3, pp. 688-673, 2016.

[7] S. J. Harley, R. A. Myers and A. Dunn, "Is catch-per-unit-effort proportional to abundance?," *Canadian Journal of Fisheries and Aquatic Sciences,* vol. 58, no. 9, 2001.

[8] M. N. Maunder and A. E. Punt, "Standardizing catch and effort data: a review of recent approaches," *Fisheries Research,* vol. 70, no. 2-3, pp. 141-159, 2004.

[9] E. N. de Souza, K. Boerder, S. Matwin and B. Worm, "Improving fishing pattern detection from satellite AIS using data mining and machine learning," *PLoS ONE,* vol. 11, no. 7, p. e0158248, 2016.

[10] D. Slagstad and T. A. McClimans, "Modeling the ecosystem dynamics of the Barents sea including the marginal ice zone: I. Physical and chemical oceanography," *Journal of Marine Systems,* vol. 58, no. 1-2, pp. 1-18, 2005.

[11] D. Slagstad, "Modeling and simulation of physiology and population dynamics of copepods," *Modeling, Identification and Control,* vol. 2, pp. 119-162, 1981.

[12] K.-J. Reite, J. Haugen, F. A. Michelsen and K. G. Aarsæther, "Sustainable and Added Value Small Pelagics Fisheries Pilots," in *Big Data in Bioeconomy*, Springer, Cham, 2021, pp. 389-409.

[13] B. Eaton, J. Gregory, B. Drach, K. Taylor, S. Hankin, J. Blower, J. Caron, R. Signell, P. Bentley, G. Rappa, H. Höck, A. Pamment, M. Juckes and M. Raspaud, "NetCDF Climate and Forecast (CF) Metadata Conventions," 2017. [Online]. Available: https://cfconventions.org/Data/cf-conventions/cf-conventions-1.7/cf-conventions.html. [Accessed 23 December 2021].

[14] European Commission, *Horizon Europe Work Programme 2021–2022, part 9: Food, Bioeconomy, Natural Resources, Agriculture and Environment,* Brussels, 2021.

[15] Norwegian Directorate of Fisheries, “Electronic Reporting Systems,” 19 March 2015. [Online]. Available: https://www.fiskeridir.no/English/Fisheries/Electronic-Reporting-Systems. [Accessed 27 December 2021].