Design goals, technical implementation and pratical use of Another data entry system (Andes)

Daniel Ricard, David Fishman, Lindsay Beazley, Brian Boivin, Jamie Emberley, Ryan Martin, Jenni McDermid, Nicolas Rolland, David Sean-Fortin, Quentin Stoyel and Pablo Vergara

Science Branch
Gulf Region
Fisheries and Oceans Canada
Moncton, New Brunswick, E1C 5K4, Canada

2023

Canadian Technical Report of Fisheries and Aquatic Sciences ####

Fisheries and Oceans Canada | Pêches et Océans Canada

Canada

Canadian Technical Report of Fisheries and Aquatic Sciences

Technical reports contain scientific and technical information that contributes to existing knowledge but which is not normally appropriate for primary literature. Technical reports are directed primarily toward a worldwide audience and have an international distribution. No restriction is placed on subject matter and the series reflects the broad interests and policies of Fisheries and Oceans Canada, namely, fisheries and aquatic sciences.

Technical reports may be cited as full publications. The correct citation appears above the abstract of each report. Each report is abstracted in the data base *Aquatic Sciences and Fisheries Abstracts*.

Technical reports are produced regionally but are numbered nationally. Requests for individual reports will be filled by the issuing establishment listed on the front cover and title page.

Numbers 1-456 in this series were issued as Technical Reports of the Fisheries Research Board of Canada. Numbers 457-714 were issued as Department of the Environment, Fisheries and Marine Service, Research and Development Directorate Technical Reports. Numbers 715-924 were issued as Department of Fisheries and Environment, Fisheries and Marine Service Technical Reports. The current series name was changed with report number 925.

Rapport technique canadien des sciences halieutiques et aquatiques

Les rapports techniques contiennent des renseignements scientifiques et techniques qui constituent une contribution aux connaissances actuelles, mais qui ne sont pas normalement appropriés pour la publication dans un journal scientifique. Les rapports techniques sont destinés essentiellement à un public international et ils sont distribués à cet échelon. II n'y a aucune restriction quant au sujet; de fait, la série reflète la vaste gamme des intérêts et des politiques de Pêches et Océans Canada, c'est-à-dire les sciences halieutiques et aquatiques.

Les rapports techniques peuvent être cités comme des publications à part entière. Le titre exact figure au-dessus du résumé de chaque rapport. Les rapports techniques sont résumés dans la base de données *Résumés des sciences aquatiques et halieutiques.*

Les rapports techniques sont produits à l'échelon régional, mais numérotés à l'échelon national. Les demandes de rapports seront satisfaites par l'établissement auteur dont le nom figure sur la couverture et la page du titre.

Les numéros 1 à 456 de cette série ont été publiés à titre de Rapports techniques de l'Office des recherches sur les pêcheries du Canada. Les numéros 457 à 714 sont parus à titre de Rapports techniques de la Direction générale de la recherche et du développement, Service des pêches et de la mer, ministère de l'Environnement. Les numéros 715 à 924 ont été publiés à titre de Rapports techniques du Service des pêches et de la mer, ministère des Pêches et de l'Environnement. Le nom actuel de la série a été établi lors de la parution du numéro 925.

Canadian Technical Report of
Fisheries and Aquatic Sciences nnn

2023

DESIGN GOALS, TECHNICAL IMPLEMENTATION AND PRACTICAL USE OF ANOTHER DATA ENTRY SYSTEM (ANDES)

by

Daniel Ricard[1], David Fishman[1], Lindsay Beazley[2], Brian Boivin[3], Jamie Emberley[4], Ryan Martin[4], Jenni McDermid[1], Nicolas Rolland[1], David Sean-Fortin[3], Quentin Stoyel[1] and Pablo Vergara[1]

[1]Gulf Fisheries Centre, Fisheries and Oceans Canada, 343 Université Avenue, Moncton, New Brunswick, E1C 9B6
[2]Bedford Institute of Oceanography, Fisheries and Oceans Canada, 1 Challenger Dr., Dartmouth, Nova Scotia, B2Y 4A2
[3]Institut Maurice Lamontagne, Fisheries and Oceans Canada, 850 Rte de la Mer, Mont-Joli, Québec, QC G5H 3Z4
[4]Saint Andrews Biological Station, Fisheries and Oceans Canada, 125 Marine Science Dr., St. Andrews, New Brunswick, E5B 0E4

Cat. No. Fs97-6/nnnE-PDF ISBN XXX ISSN 1488-5379

Correct citation for this publication:

Ricard, D., Fishman, D., Beazley, L., Boivin, B., Emberley, J., Martin, R., McDermid, J., Rolland, N., Sean-Fortin, D., Stoyel, Q. and Vergara, P. 2023. Design goals, technical implementation and practical use of ANother Data Entry System (ANDES). Can. Tech. Rep. Fish. Aquat. Sci. nnn: vii + 42 p.

CONTENTS

ABSTRACT

Ricard, D., Fishman, D., Beazley, L., Boivin, B., Emberley, J., Martin, R., McDermid, J., Rolland, N., Sean-Fortin, D., Stoyel, Q. and Vergara, P. 2023. Design goals, technical implementation and practical use of ANother Data Entry System (ANDES). Can. Tech. Rep. Fish. Aquat. Sci. nnn: vii + 42 p.

An event logger and data entry system to facilitate the capture of information collected during scientific field work conducted by DFO Gulf Region was recently developed. The application, called “ANother Data Entry System” (Andes) was implemented as a browser-based environment that facilitates its deployment in a variety of data capture scenarios. Here, we describe the design goals and implementation details of Andes, and present a number of usage cases. Although the system is tailored primarily for use on fishery-related ecosystem surveys, the flexibility of the application in terms of adapting to new gear types and deployment options promotes its use in a wide variety of applications. To promote transparency, longevity, endorsement and use of Andes by other scientific groups, the application utilizes Open Source software and a shared code development platform.

RÉSUMÉ

Ricard, D., Fishman, D., Beazley, L., Boivin, B., Emberley, J., Martin, R., McDermid, J., Rolland, N., Sean-Fortin, D., Stoyel, Q. and Vergara, P. 2023. Design goals, technical implementation and practical use of ANother Data Entry System (ANDES). Can. Tech. Rep. Fish. Aquat. Sci. nnn: vii + 42 p.

Un système d'entrée de données pour faciliter la saisie des informations recueillies lors des activités scientifiques menées par la Région du Golfe du MPO a été développé. L'application, qui s'appelle "Another data entry system" (Andes), a été mise en œuvre en tant qu'environnement basé sur un fureteur, ce qui facilite son déploiement dans une variété de scénarios de saisie de données. Les objectifs de conception et les détails de mise en œuvre d'Andes sont décrits, et un certain nombre de cas d'utilisation sont présentés. Pour promouvoir l'utilisation d'Andes par d'autres groupes scientifiques et assurer la pérennité du projet, l'application s'appuie sur des logiciels code source ouvert et utilise une plate-forme de développement de code partagé.

1 Introduction

The scientific activities conducted by the Atlantic regions of Fisheries and Oceans Canada (DFO) include ecosystem surveys, oceanographic monitoring surveys, sentinel fisheries programs, coastal surveys, port sampling and many others where field data are collected. These monitoring programs require the collection of a wide variety of information on the marine environment and ecosystems, and are used to inform commercial fishery activities. To support the collection of data and metadata during those surveys, paper-based systems were initially used during field work on research vessels, fishing vessels or at commercial ports. With the advent of computing capabilities, paper data sheets were digitized and the resulting data was stored for subsequent analyses. In the 1980s and early 2000s, a number of computer-based data entry systems were developed, in particular to support the activities of scientific ecosystem cruises where a variety of instruments are used to collect data. For trawl surveys, this meant a tool that could support the activities associated with processing the catches from the net, and that could also be used to manage survey activities while at sea. In the Maritimes and Gulf regions, the Groundfish Survey Entry (GSE) and later the Ecosystem Survey Entry (ESE) were deployed on workstations in the wet laboratory of research ships. These applications supported the capture of data and information on individual specimens, including the collection of length frequencies and more detailed sampling based on the survey goals and associated protocols. In the Québec Region, a similar tool called the "Module des Relevés de Recherche" (MRR) was developed and used to directly enter data in digital format on the research vessel.

The later versions of those data entry systems were built with Visual Basic for Applications (VBA) and relied on Microsoft Access databases. While highly functional, both the ESE and the MRR have important limitations:

- Despite being used over a local area network (LAN), these tools are not designed to work in a networked configuration.
- These tools are unilingual (French or English depending on the region) and their infrastructure does not readily support multiple languages in their user interface.
- Pre-.NET Visual Basic (VB), the main development frameworks used to build the applications, are dated programming languages. To adapt the software using modern libraries and Application Programming Interface (API) in VB will become increasingly difficult over time.
- The use of a Version Control System (VCS) is incompatible with the Microsoft Access applications that VB applications rely on. The absence of a VCS presents serious challenges for managing, disseminating and troubleshooting versions of an application across multiple platforms and is a major impediment for a collaborative approach to development.
- While these applications were built by DFO, multiple attempts to obtain support for these tools from Information Management and Technology Services (IMTS) (now Chief Digital Officer Sector, CDOS) were unsuccessful. This lack of support means that technicians often resort to either Excel or pen and paper to perform activities unsupported by the applications. Performing data entry using separate and disconnected methods means

that the data has to be manually consolidated at a later date which is time-consuming and prone to error.

- The lack of a true server-side application (e.g. web application) meant that agile development was effectively impossible; especially in the context of a mission that was underway.

While the status quo was meeting the immediate data capture needs of scientists, a proactive stance towards addressing the above spurred the development of *An*other *d*ata *e*ntry *s*ystem (Andes). This project effectively began in the summer of 2018 as an application to be deployed on field tablets to support port sampling data collection, and later evolved into an application with a wider range of uses. The first field deployment of Andes took place in the fall of 2019 during the southern Gulf of St. Lawrence September ecosystem survey. In its pilot year, Andes was used in parallel to the ESE on board the CCGS *Teleost*.

This report documents the design principles that guided the development of Andes, provides technical details about its implementation and details the usage cases encountered thus far. It provides guidelines for assessing the suitability of Andes for novel applications (e.g., other field activities and sampling programs) and discusses the ways in which the application could gain further adoption in supporting a diverse range of data collection activities within DFO.

2 Methods

The overall goals for the redevelopment of the data logging system used in the Gulf Region were to utilize a contemporary software development environment, facilitate transparency through use of a shared code development environment, have the ability to support multiple languages, and to support flexible deployments. The new application would be modularized into different consoles to allow for a diverse range of uses, all while leveraging the same core features and infrastructure that the main application was built on. As the starting point of Andes was the ESE, the replacement data entry software had to first replicate the proven functional capabilities of ESE. Early versions of Andes achieved those goals and the system was further developed over its usage by scientists in the Gulf Region.

Initially, Andes was designed with 3 main usages in mind: A) as the main data entry system for ecosystem survey observations B) to track and capture data and metadata from oceanographic observations, and C) for commercial port sampling activities. For use on the ecosystem surveys, the application must allow for the capture of all information and parameters detailed in the sampling protocols of the survey (e.g. Hurlbut and Clay (1990)). This entails the capture of all information related to fishing activities, plus ancillary variables such as weather, sea state, and water temperature. As catch contents are sorted and identified during the ecosystem surveys, the data capture application must support users in obtaining catch weights and abundance by taxon, individual observations of a specimen's length, weight, maturity status, etc. The application should also be designed to flag users when specimens are to be collected, or when additional sampling requirements were present (e.g. collecting fish otoliths).

A large component of the ecosystem surveys conducted by the Gulf, Quebec, Maritimes, and Newfoundland and Labrador Regions includes the collection of physical, chemical, and biological oceanographic data in support of the Atlantic Zone Monitoring Program (AZMP) (Therriault et al. 1998). A CTD/Rosette system is deployed at a subset of fishing stations where vertical profiles of temperature, conductivity (salinity), dissolved oxygen, chlorophyll and other parameters are collected. Water samples (e.g., nutrients, dissolved oxygen, salinity) are collected at predetermined depths using the Rosette system and attached Niskin bottles, and plankton samples are also collected via vertical ring net tows. These data support annual state of the ocean reporting conducted by the AZMP, but are also used for catch interpretation and in various stock assessment processes.

Historically, the event metadata associated with the deployment of oceanographic equipment on the ecosystem surveys was recorded using the Electronic Logbook (ELOG) system, a URL-based logging system developed by the Paul Scherrer Institute, Switzerland. However, ELOG was installed and operated completely separately from logging systems used for the biological data, making it difficult to merge the two datasets upon completion of a survey. Ideally, the new application would streamline the capture of oceanographic data and simplify its association with the corresponding biological data.

In order to support data collection during commercial port sampling activities (Benoît and Daigle 2007), the application would be deployed on field tablets for scientific staff that collect data related to commercial fishing activities. A typical usage case is to obtain length-frequency samples from commercial landings, and to also obtain length-stratified hard parts used for age estimation, and also the collection of whole individuals for later processing in the laboratory. It is not expected that there would be any interaction between this module and the ones listed above.

The following section outlines additional user requirements and design principles that also played a role in decision-making during the development of Andes.

2.1 Flexibility, scalability and reliability

The development environment used for the application must be a contemporary programming language with a proven track record for performance, usability and adaptability. Within the context of DFO Science, there are a high number of use-cases under which this application can be deployed. For example, the number of users might range from a single employee alone in the field to several dozen scientists, technicians and vessel personnel participating in a research cruise. Similarly, the application might need to be deployed on a stand-alone device or accessed from within a LAN or even over a wide area network (WAN). The application must be able to accommodate a wide range of practical scenarios.

2.2 Version control/source control

The utilization of a VCS is an indispensable component of a sustainable development workflow. Version management is especially important in the context of having concurrent instances of the application in production at any given time. Knowing the version of a production instance is necessary to resolve any issues that might arise. Similarly, databases and backup files are intimately linked to an application's version number. In order to successfully re-instantiate a backup file, the precise version under which it was produced must be known. The VCS will also provide an indispensable framework for the coordination, examination and integration of contributions from collaborators.

2.3 Unit Testing

The application performance needs to be reliable, especially considering its potential to be deployed in remote field environments. The implementation of unit tests is a practical way to ensure the maintenance of core functionality over time. At one extreme, the addition of any code can be preceded by the creation of unit tests (i.e., test-driven development). This approach will maximize the stability of an application, but can hinder the momentum of a project, especially in its early stages. At the other extreme, application development in the complete absence of unit testing occurs at a relatively fast pace but will result in a project that is vulnerable to breaking in unexpected ways and one that is difficult to maintain and pass on to other developers.

2.4 Backup strategy

In all deployment scenarios, data of high business value will be captured and there is little to no tolerance for data loss. Accordingly, the application must have a way to facilitate the implementation of a robust data-backup strategy. Capturing numerous snapshots of the application (and database) is ideal since doing so provides redundancy as well as the ability to revert to a specific point in time. If possible, the backup files and snapshots should be stored on a volumes with some form of redundancy.

2.5 Customizable protocols

The flexibility of project leads to design, modify and report on their sampling protocol without depending on developers is very important. The application should allow users to provide a variety of detailed information based on their particular sampling protocols, without the need to change the application source code. Similarly, different protocols utilize different code conventions for the identification of biological catches. The application should offer the flexibility of users to utilize their preferred system of catch codes.

2.6 Quality control

The implementation of quality control checks in a data entry application is of paramount importance. At a minimum, the following quality control checks should be leveraged:

Sets The application should ensure all the required fields on a given set card have been filled in. Users should also be warned if the set's start and/or end coordinates fall outside the expected sampling stratum (if applicable).

Catches The application should verify the validity of catches that do not have any data entry associated with them. This validation is important to help identify catches that might have been entered accidentally.

Specimens The application should flag specimens whose length falls outside an acceptable range. Similarly, a validation of the specimen's length-to-weight ratio should be performed.

Observations Individual observations are characterized by an observation type. Observation types should have predefined data types such as integer, float, string or categorical. The application should ensure that inputted observation values should never be left null and that they respect the data type of the corresponding observation type. In the case where an observation type (e.g., sex) has a set of defined categories (e.g., male, female, unknown), the application should ensure that any entered values fall within the set of available options.

2.7 User interface

The user interface of the application can have a significant impact on user experience and on-boarding. By ensuring the application has a modern and intuitive interface, the barriers related to on-boarding new users are significantly reduced. Furthermore, an intuitive interface will reduce the need for extensive help documentation. Wherever extra annotation is required, documentation should be inserted directly in the application in the form of tool tips and help bubbles. By appealing to the end users' intuitions and by providing in-situ help documentation, we reduce the likelihood fields and features being used incorrectly.

2.8 Reactivity

In the context of being on a research vessel survey, data entry happens at a very fast rate and on numerous devices; often with multiple transactions per second. Accordingly, it is imperative that the application does not create a bottleneck for data entry and is able to keep pace with experienced technicians. The usage of a reactive javascript library in conjunction with an API would allow data entry to occur without webpages having to constantly refresh.

2.9 Multilingualism

The ability for users to choose the language of their choice in the application is of considerable importance to this project. Previous tools that have been used were unilingual, and this by itself would have limited the scope of their use in a national context.

3 Results

The architecture of Andes differs significantly from that of its predecessor (see Figure 1). The application and its associated services are centralized on one or several servers and include: 1) a web-service for handling HTTP requests and responses; 2) a database service for storing data associated with the application; 3) a file-sharing service for handling the storage of backups and related files; 4) a printing service used for printing out specimen labels; and 5) a message-brokering service for handling asynchronous tasks. Client devices, such as data-entry workstations, no longer require the installation of anything more than a modern web browser; i.e., one that is capable of supporting HTML5 and ECMAScript 2016 (Javascript). Accordingly, this increases the range of devices and operating systems that may be used for accessing the application. For instance, the switch to using Andes has allowed for the integration of mobile phones, tablets and linux workstations into the data entry workflow. Finally, the new configuration means multiple stations can simultaneously receive and enter data into the same Andes instance.

The Django Web Framework was selected for the backend of this application due to its modularized nature; virtually all aspects of the programming framework can be decoupled. Furthermore, the Django framework is written in pure Python language; an open-source, generalized object-oriented programming language that is popular for use in data-heavy applications. In addition to standard Django templates used to render HTML webpages to end-users, the web framework also contains an elaborate REST-API component. Django uses an Object Relational Model (ORM) to handle the data layer, and includes APIs for a variety of modern relational database management systems (i.e., PostgreSQL, MariaDB, MySQL, Oracle, SQLite, etc.). Several of the application's frontend templates contain reactive components implemented in Vue.js.

The frontend of the application is built in HTML5, javascript and Cascading Style Sheet. Most users will be familiar with the flow and functionality of a web browser and will be comfortable navigating and entering data into a website. Facilitated by the Django model and form classes, all controls (i.e., fields) on the website contain verbose descriptions and help text. The Bootstrap v5.0 CSS and javascript libraries were utilized in order to give the application a sleek, modern look and to ensure compatibility with different types of devices (e.g., personal computers, tablets and mobile devices). The Bootstrap library also provides palatable styles for displaying help text in the form of popups and tool tips.

While web applications are most often used over a network, the Django library comes with a development web-server that permits users to serve and use the application locally. In this scenario, a single computer acts simultaneously as server and client. While there are important limitations to the use of the Django development web-server in a full-scale production environment, this configuration is adequate for stand-alone use-cases (e.g. field tablet used for port sampling).

The architecture used by Andes creates networking requirements that were not previously present in the ESE and MRR. The server and the client devices must be connected to the same network. The network does not need to have access to the WAN (i.e., the internet) connection. As long as they are connected to the same LAN, they can be configured for work together. This suits the networking environment on board remote vessels that can have sporadic connection failures with the wide area network.

The Andes application provides access to its various components based on an internal system of authentication and authorization. The credentials of a given user will affect what action they are able to do. For example, while the chief scientist is able to modify sampling requirements for the mission, other users are not.

3.1 Andes Modules

The main index page of the Andes user interface (Figure 2) allows users to access several different modules that loosely correspond to the main use cases defined in the above Methods section. Access to the different modules is determined by a system of authentication and authorization. Consequently, the index page will appear different to users depending on the permissions they have been granted.

While Andes has been implemented using a modular design, there is a core set of components that are used across modules. This is true at the level of the database, where in addition to a number of shared tables, sub-modules will have tables that are specific to a given use case. For example, while the Mission table is used across several modules, the Specimen sub-module is specific to the Ecosystem Survey module.

Similarly, Andes has modules that are used across multiple user scenarios; in particular, the Bridge module (Figure 3) and the Cruise Dashboard (Figure 4). These modules display high-level information to end users such as queued stations, vessel speed, position and heading and various summaries of Science activities that are underway or that have already taken place. All core pages of the application can be toggled to night mode, as desired.

3.1.1 Ecosystem survey

The Ecosystem Survey module is the main entry point that technicians will use to input survey data. A depiction of the main tables involved in the Ecosystem Survey module, and their relationships are displayed in Figure 5. This component of Andes replicates the capabilities of the ESE for capturing detailed information on length, weight, ageing material, maturity, etc. about fish and invertebrate specimens. This module, typically accessed from the wet laboratory of a survey vessel, is used for all entry of data related to measurements and observations of marine organisms. New catches are first entered into the Active Set page, as portrayed in Figure 6. The Active Set page accepts regional catch codes as a way to input new catches into a set. If a code is not known, a search feature is available. Next, baskets and their corresponding weights and statuses (e.g., sampled vs. not sampled) are then entered into the Catch Card page (Figure 7). Finally, specimens are entered into the Data Entry page (Figure 8). As the data entry progresses, users are dynamically prompted with observation fields that follow the catch-specific sampling protocol. An overview of the sampling protocol is displayed on the right-hand side of the Data Entry page.

3.1.2 Oceanography

The Oceanography module is a stand-alone component of the Andes application that is capable of being deployed independently of the Ecosystem Survey module. This module is used to track and record the deployment of CTD/Rosette systems and plankton nets (Figure 9). Basic metadata, such as the location and date/time of deployment, 'on bottom' and recovery of gear can be collected. Additionally, the Oceanography module allows for users to enter which water samples/parameters were collected from pre-determined depths after CTD/Rosette deployment (Figure 10). The simplified Entity Relationship Diagram of the Oceanography module of Andes is presented in Figure 11. Upon conclusion of a survey, Andes is designed to provide a series of oceanographic summary reports that summarize aspects of the data collected:

Mission instrument report Provides a summary of the oceanographic equipment used on the survey, including component type, model and serial number, and date of last calibration (for CTD sensors).

CTD metadata report for each fishing set and station where the CTD/Rosette system was deployed, the date, time, position, sounding, and bottle sample IDs are provided.

Hydrolog report For each fishing set and station number, a summary of the CTD/Rosette and ring net deployments are provided, including the event numbers corresponding to each gear deployment, comments entered into Andes regarding each gear deployment, and summary information related to each set and station, including surface temperature, sounding, day of year, and date/time.

Plankton report This report provides a summary of the plankton net deployments conducted in relation to each fishing set and stratum. The wire out and wire angle, as well as flow meter start and end are provided.

Bottle report This report provides a detailed summary of each CTD/Rosette operation (e.g., altimeter height, bottle height and depth, max. CTD depth), the water samples/parameters collected, and also includes the results of the Winkler titrations for dissolved oxygen samples, if entered into the Oceanographic Activity detail page.

TS report This report provides a summary of the surface and bottom temperature and salinity data from each CTD cast in relation to each set, station, and stratum.

ELOG report This report emulates the .log summary report produced by ELOG. This report is required in order to upload the survey data into the Microsoft Access template used by AZMP to load data to BioChem, DFOs national repository for discrete and plankton data.

These reports facilitate the post-processing of the data and its distribution and upload to various open source platforms.

3.1.3 Port sampling

Andes was adapted from a previous port sampling module to support activities where technicians obtain length frequency samples from commercial fishing activities. The protocols used in the Port Sampling module are much simpler than those belonging to the Ecosystem Survey module. Through the protocols, users are able to control collection quotas (e.g., “keep two specimens per bin”), the flow of data entry (e.g., which field should be displayed in the sample form) and the layout of the data entry page (e.g., length bins organized in a vertical or horizontal configuration) (see Figure 12). Typically, production instances of this module are deployed on ruggedized field tablets that are suitable for use in wet environments.

The design of the user interface is simple and intuitive, where each length bin is a large button on the display. As specimens are tallied, the corresponding buttons on the touchscreen are pushed. A screenshot of the data entry page is presented in Figure 12. The various tables of the Port Sampling module and how they relate to one another is shown in Figure 13.

3.1.4 Shrimps

The Shrimp Ecosystem Survey (“Shrimps”) module resulted from an early adoption of Andes in the Québec Region. The niche workflow requirements of shrimp survey operations justified the creation of its own Andes module.

3.1.5 Bridge

The Bridge module (Figure 3) is used by navigation officers in the wheelhouse (usually set up on a tablet) to input fishing set metadata. The data entry occurs in two ways: 1) by directly editing the set form (a.k.a. the set card); and/or 2) via the Fishing Console, which is displayed in Figure 14. The Fishing Console can capture a number of different events that take place during fishing as well as information coming from sonar and trawl mensuration system sensors (e.g., Scanmar system).

Part of the task-specific suite of utilities is a forecasting tool that provides real-time prognostics of survey completion targets based on assumed transit speeds and time spent fishing and processing the trawl catches (Figure 15). This feature provides a useful tool for the chief scientist to evaluate different sampling objectives and survey route. As the conditions change regularly during a survey, this tool links with the set manager to provide an estimate of the amount of time required to complete planned stations. This tool can be used for short-term planning (i.e. a day’s worth of sampling) or longer-term planning (i.e. a mission’s worth of sampling).

Another task-specific utility is the completion map which shows what strata have been completed based on target and minimum number of sets per stratum. This map provides the chief scientist with a clear visual depiction of what has been accomplished, and what remains to be accomplished, during the survey. The number of sets conducted in each stratum is compared to the minimum and target number of sets per stratum to determine the colour that each stratum will appear in the progress map (Figure 16).

3.1.6 Maps

What about the maps module?

3.1.7 Set manager

What about the Set Manager module?

3.2 Technical Implementation of Other Design Goals

3.2.1 Version control/source control

It was decided to use Git as the VCS for this project due to the fact that it is both widespread in use and open-sourced. The remote repository for this project is currently hosted as a private project on the Gulf Science organizational GitHub account. The project is additionally making use of GitHub infrastructure including, pull requests, issue management (e.g., bugs, feature requests and general enhancement requests), security alerts and version releases. The Andes documentation is also being served using GitHub Pages and can be accessed here.

3.2.2 Unit testing

For Andes development, we use a mixture of test-driven development for critical components of the application, and are also continually adding unit tests for more user-specific components. Using the built-in Python/Django testing framework, this strikes a middle ground between the two approaches described in the above section. While the goal is not to implement test-driven development for the application, the use of unit tests is highly encouraged, especially to back up the core functionality of the application.

3.2.3 Backup strategy

The Andes application has the capability to perform manual and automatic backups. These backups consist of two parts: a JSON export of the entire database and the recording of the current version of the application (i.e., the git hash). The structure of the application models will change over time and thus, so will the structure of the data. Accordingly, in order to re-instantiate a particular data snapshot, it is critical to know the precise version from which it was exported. This combination of data export and git version number, gives users the perpetual ability to recreate the exact application environment from the time of the snapshot, no matter how much the application has changed in the interim. In the application, backups are automatically created upon closing sets. Moreover, users also have the ability to manually trigger a backup at anytime.

3.2.4 Customizable protocols

Andes provides project leads the ability to create and modify sampling protocol through the user interface. By doing so, project leads are able to shape the flow and control the behaviour of the application during data entry. This including deciding which fields to display in a form (e.g., set cards), importing stations and other geographical features (e.g., sampling strata, NAFO areas, Marine Protected Area (MPA)s, etc.) and the quotas and observation fields associated with different catch items. Examples of catch-specific sampling requirements that can be programmed by project leads can be found in Table 1.

Andes also allows users to specify a preferred system of catch codes by attaching the code collection to a sampling protocol. When a given protocol is active, all catch items are referenced by their corresponding entry in the attached catch collection list.

3.2.5 Quality control

Andes successfully implements a suite of quality control checks. All quality control flags that are raised during data entry are appended to a report and require sign off before a set can be closed.

Sets

The Andes provides very clear feedback regarding the completeness status of a given set. Flags are raised if a set's start and/or end coordinates are outside the expected sampling stratum. Additionally, an alarm is sounded in the bridge console module when deviating from the desired stratum in real-time. The bridge console also provides personnel with fishing timers and tow distance displays (see Figure 14).

Andes will flag when the tow distance as calculated by the cruise track, differs from the tow distance as calculated by a straight line between the start and end coordinates by more than 5%. Andes will flag when the start or stop coordinates of a set are not within the expected NAFO area (if applicable).

Catches

Flags are generated for catches that do not have any data entry associated with them. In addition to this, Andes will flag when the weights entered for baskets that are considered suspect. This is determined by either the default maximum basket weight (mission level), or the maximum basket weight for a given species. Andes will also flag when the difference of total weight of *sampled* basket differs by more than 25% from the total calculated specimen weight. The total calculated specimen weight is a combination of actual weights (when collected) and those which were estimated from length measurements. The latter is achieved by using regression coefficients estimated from historical length (L in centimeters) and weight (W in grams) observations using the following equation:

$$W = aL^b \tag{1}$$

Andes offers the option to specify separate regression coefficients (a and b) for males, females or unspecified individuals (as shown in Figure 17).

A recent version of Andes implemented a new optional layer of quality control. This feature allows project leads to assign allow-lists and restrict-lists, commonly known as “whitelists” and “blacklists”, to a mission or to a specific geographic feature (e.g., a stratum). In this way, the validity of each catch entered into a set can be assessed. For example, if a catch being recorded is *not* on the that set's associated allow-lists, the end user will be notified that this is an unusual observation and will be prompted to collect documentation. Similarly, if a project lead adds a catch to the mission's restrict-list, users who enter this catch will receive a warning message, asking them to double-check the assignment. This is useful when project leads want to limit the usage of certain taxa during data entry, e.g., *Alosa* sp. is preferred over the use of *Alosa pseudoharengus*.

Specimens

Specimen lengths and length-to-weight ratios are validated against the parameters entered in the sampling requirements for that catch. Acceptable length-to-weight ratios are assessed by comparing the actual weight to the estimated weight, as described in the section above. Additionally, Andes will flag when there is a mismatch between fish maturity and somatic length. Project leads can specify mature length thresholds for males, females or unspecified individuals in the sampling requirement of a given species (as shown in Figure 18).

Observations

When entering observations, Andes enforces the data type of the corresponding observation type; invalid entries are not accepted. When entering an observation for an observation type that has categories, Andes will display the list of options to the end user and inhibit users from entering invalid selections. Andes will also ensure observation types are not left blank. However NaN entries are permitted when a particular observation is meant to be skipped. Certain observation types are meant to be unique, e.g., unique tag number. If this is flagged at the level of the observation type, end users will be notified if there is a violation of this structure.

3.2.6 Reactivity

The Django REST framework was used to construct the WebAPI component of the application. The project takes a hybrid approach, combining the use of standard Django views and Vue.js frontend applications embedded in the templates. The latter were used to avoid the need for constantly reloading webpages and to optimize the flow of traffic across the network. Reactive javascript frontend applications also provide a better experience from the point of view of an end-user.

3.2.7 Multilingualism

The Django framework has excellent support for internationalization and localization, including the translation of text and the formatting of dates, times and numbers. It achieves this using a system of ‘hooks’ used by developers to indicate which parts of the code should be localized.

See Django - Internationalization and localization for more details on this process. In our application, an end-user can toggle between English and French by simply clicking on a button. In this way, each client can view the application in the language of their choice.

3.3 Usage cases

The Andes Django project contains multiple apps that loosely correspond to the different use cases defined in the above methods section. The different modules can be clearly observed in the index page of the site (Figure 2. As noted above, the access to different apps is determined by a system of authentication and authorization. Accordingly, the index page will appear different to users depending on the permissions they have been granted.

While Andes has been implemented using a modular design, there is a core set of components that are used across modules. This is true at the level of the database, where in addition to a number of shared tables, sub-modules will have tables that are specific to a given use case. For example, the Mission table is used across several modules however the Specimen module is specific to the Ecosystem Survey application.

Similarly, Andes has apps that are used across multiple user scenarios; in particular, the Bridge Console (Figure 3) and the Cruise Dashboard (Figure 4). These apps display high-level information to end users such as queued stations, vessel speed, position and heading and various summaries of Science activities that are underway or that have already taken place. All core pages of the application can be toggled to night mode, as desired.

3.3.1 Ecosystem survey data entry

The Ecosystem Survey app is the main entrypoint that technicians will use to input survey data. A depiction of the main tables involved in the Ecosystem data entry application, and their relationships are displayed in Figure 5. This component of Andes replicates the capabilities of the ESE for capturing detailed information on length, weight, ageing material, maturity, etc. about fish and invertebrate specimens. This app, typically accessed from the wet laboratory of a survey vessel, is used for all entry of data related to measurements and observations of marine organisms. New catches are first entered into the Active Set page, as portrayed in Figure 6. The Active Set page accepts regional catch codes as a way to input new catches into a set. If a code is not known, a search feature is available. Next, baskets and their corresponding weights and statuses (e.g., sampled vs. not sampled) are then entered into the Catch Card page (Figure 7). Finally, specimens are entered into the Data Entry page (Figure 8). As the data entry progresses, users are dynamically prompted with observation fields that follow the catch-specific sampling protocol. An overview of the sampling protocol is displayed on the right-hand side of the Data Entry page.

3.3.2 Shrimp ecosystem survey app

Shrimp data is collected by the Quebec region in the course of its yearly ecosystem survey in the estuary and northern gulf of the Saint-Lawrence as well as during other occasional surveys. These data are collected for the stock assessment of the northern shrimp (*Pandalus borealis*)as well as to assess the biodiversity and abundance of shrimp species. In the case of the Quebec region the biological measures related to shrimps are done onboard where in other regions shrimps subsamples are frozen and biological data is collected in the laboratory at the end of the mission. The shrimps application was developed to allow the entry of shrimps data as it follows a different workflow compared to the ecosystem survey application. At every set, a subsample of between 2 and 3 kilograms is first collected from the total catch. This sample is then separated by species, then by maturity stage and only then are the biological data (cephalothorax length for all species and occasionally weight for *Pandalus borealis*) taken.

As noted above, the Bridge console is used by navigation officers in the wheelhouse (usually set up on a tablet) is used to input fishing set metadata. The data entry occurs in two ways: 1) by directly editing the set form (a.k.a. the set card); and/or 2) via the Fishing Console, which is displayed in Figure 3. The Fishing Console can capture a number of different events that take place during fishing as well as information coming from sonar and trawl mensuration system sensors (e.g., Scanmar system).

Part of the task-specific suite of utilities is a forecasting tool that provides real-time prognostics of survey completion targets based on assumed transit speeds and time spent fishing and processing the trawl catches (Figure 15). This feature provides a useful tool for the chief scientist to evaluate different sampling objectives and survey route. As the conditions change regularly during a survey, this tool links with the set manager to provide an estimate of the amount of time required to complete planned stations. This tool can be used for short-term planning (i.e. a day's worth of sampling) or longer-term planning (i.e. a mission's worth of sampling).

Another task-specific utility is the completion map which shows what strata have been completed based on target and minimum number of sets per stratum. This map provides the chief scientist with a clear visual depiction of what has been accomplished, and what remains to be accomplished, during the survey. The number of sets conducted in each stratum is compared to the minimum and target number of sets per stratum to determine the colour that each stratum will appear in the progress map (Figure 16).

3.3.3 Oceanographic metadata collection

4 Discussion

Since its initial deployment during the 2019 September ecosystem survey in the southern Gulf of St. Lawrence, the capabilities and performance of Andes have significantly improved. The software has also been fully adopted as the main logging software used during the winter and summer ecosystem surveys in the Maritimes Region, and is also utilized by the Quebec Region.

The current capabilities of Andes have evolved over the last three years where the system has been used to support field activities in the Gulf, Maritimes and Quebec regions.

The forecasting tool and progress map presented in this report were not in the original deployment of the application and were added during the cruise based on the needs of scientific and Canadian Coast Guard (CCG) personnel. The reporting facilities provided by Andes were developed to support the many users of the data collected during scientific activities. Two types of reports are available for Andes: 1) reports meant to be used during field activities and 2) reports meant to be used after field activities are completed.

If a vessel is located within range of cellular phone signals, it is possible to make Andes accessible via the internet. While this has some important advantages, exposing an IP to the internet comes with security risks that have to be mitigated.

Andes is designed so that data backups are immune to the potential pitfalls of relying on earlier versions of required packages. Having backups that are associated with a specific git version, which includes a comprehensive list of all dependencies and an exact match of the application data model, allows users to readily bring back Andes to where it was at the time of the backup. Users can then access the reporting features that will allow them to create a version of the required data that suits their needs and facilitates inclusion in existing relational database management systems or other types of data solutions.

When present, the client-server separation lends itself nicely to be adaptable to new requirements. One can potentially expand upon the *system* with minimal changes to Andes by interacting directly with its exposed REST-API. For example, an imaging system can be aware of current GPS coordinates or current Set/Station. Photos or videos acquired by such an imaging system could automatically append these as metadata, or even add media identifiers directly to the database as they are captured.

Most modern vessels will already provide the minimum necessary networking infrastructure necessary for system deployment where connectivity between a dry laboratory, a wet laboratory and the ship's wheelhouse are necessary. In the case where static networking options are not available, this requirement can be met relatively easily using inexpensive consumer-grade, portable networking equipment (wireless router, switch, etc.). However, as installing an ad-hoc network for scientific needs in a vessel may lead to code violations, we urge Andes users to collaborate with vessel management body like the CCG or chartering authority. Most modern vessels will already provide the minimum necessary networking infrastructure necessary for system deployment where connectivity between a dry laboratory, a wet laboratory and the ship's wheelhouse are necessary. In the case where it doesn't, this requirement can be met relatively easily using inexpensive consumer-grade networking equipment (wireless router, switch, etc.). However, installing as ad-hoc network for scientific needs in a vessel may lead to code violations there we urge Andes users to collaborate with vessel management body like the CCG or chartering authority.

The addition of allow-lists and restrict-lists was an important addition to Andes that came out of a Regionalization workshop that took place in December 2022 at the Gulf Fisheries Center. While these features will help improve the quality of data collected, it is important that if implemented, they should be detailed explicitly in the survey sampling programme. For example, if project leads add certain catch items to the restrict-list, this should be mirrored in the survey sampling programme.

Andes also now supports the use of various electronic tools such as calipers, electronic measuring boards and scale. Drivers for the measuring boards and scale were developed by the Gulf and Quebec regions and are available on Github.

While there is always an effort to make the user interface intuitive and self-explanatory, there is still the need to maintain documentation for the project. The Andes docs are maintained through GitHub Pages and are publicly available here: https://dfo-gulf-science.github.io/andes/

4.1 Integration of Andes with Existing Data Repositories

Andes is a standalone application with an underlying relational database management system, and the data collected during research activities must be extracted from the application and integrated into existing databases. This step has proven to be challenging but has also led to close inter-regional collaboration since the challenges associated with integration were shared by the different groups involved.

4.2 Future Directions

Onboarding efforts for coastal surveys (scallop, sea-cucumber, whelk) is underway.

4.2.1 Atlantic Zone (Offshelf) Monitoring Program (AZMP/AZOMP) oceanographic surveys

Andes was first trialed in parallel to ELOG on a dedicated oceanographic survey led by the Maritimes Region Atlantic Zone Monitoring Program (AZMP) in the spring 2022, and was recently used as the sole event logger during the 2023 Atlantic Zone Offshelf Monitoring Program (AZOMP) survey. During its initial trial on the 2022 spring AZMP survey, a number of limitations were noted. ELOG assigns each gear deployment a sequential 3-digit 'event' number. In contrast, Andes groups all gear deployments in relation to a Set number, which is a primarily fisheries-related term that is not applicable to oceanographic survey data collection. However, Andes could easily be adapted to the AZMP's normal metadata naming convention by renaming Set to 'station', given that multiple sampling events normally occur at the same AZMP station.

The initial version of Andes included only two oceanographic gear options: CTD/Rosette and ring net deployments. However, Argo floats, oceanographic buoys, moorings, multinet, and other gear are routinely deployed during AZMP and AZOMP surveys, and the order of operations must be flexible. In previous versions of Andes, the order of operations under each Set could not be modified between stations. With developer support, Andes was recently modified to allow for the inclusion of additional gear options and the ability to toggle and select between gear types should the order of operations change between stations.

ELOG allows users to create various 'logbooks' that allow for custom recording of metadata. On AZMP surveys, a custom ELOG logbook was created to capture metadata related to samples collected from underway or flow-through thermosalinograph (TSG) systems used on these surveys. A sub-module could be developed in the future to allow for a similar level of data capture using Andes.

Andes shows great promise for enhanced data recording and oceanographic sample tracking relative to ELOG. Future modifications of Andes could incorporate information on water budget requirement for each CTD cast, and could allow for the tracking of laboratory post-processing. However, making changes to Andes on the fly would not be possible on local installations (i.e., not on the custom servers permanently installed on the ecosystem survey vessels) of Andes without on-board developer support. With this, Andes could be adopted as the full-time event metadata logger on AZMP and AZOMP's oceanographic surveys in the future.

- use for ecosystem surveys by multiple DFO regions
- use of barcodes and/or QR codes for identifying and tracking down samples (e.g. otoliths, tissue samples, . . .)
- imaging system

4.3 Governance

Andes does not currently have a well-defined governing body. This makes it, as well as any project without governance, vulnerable. To alleviate this situation, we ask that this technical document as well as the developer's guide be considered when outsourcing development efforts towards Andes. This will ensure that the core vision is maintained and that software sustainability principles are upheld.

Andes reached a critical point in its lifecycle. It has been used as the main data entry strategy for numerous missions over the span of a few years. Although new features can always be added, its core functional requirements can now be locked-in. The boundaries are now well-defined and described in this report.

5 Acknowledgments

We thank all DFO personnel who were involved in the early testing and deployment of Andes prior to and during the 2020 southern Gulf of St. Lawrence ecosystem survey. We thank the Gulf Region publications coordinator for his/her assistance with getting this document published.

6 References

Benoît, H.P., and Daigle, D. 2007. Procedures for commercial catch sampling of finfish and shrimp in the southern Gulf of St. Lawrence. Can. Manuscr. Rep. Fish. Aquat. Sci. 2833: iv + 63 p.

Hurlbut, T., and Clay, D. 1990. Protocols for research vessel cruises within the Gulf Region (demersal fish) (1970-1987). Can. Manuscr. Rep. of Fish. Aquat. Sci. 2082: 143 p.

Therriault, J.-C., Petrie, B., Pepin, P., Gagnon, J., Gregory, D., Helbig, J., Herman, A., Lefaivre, D., Mitchell, M., Pelchat, B., Runge, J., and Sameoto, D. 1998. Proposal for a Northwest Atlantic Zonal Monitoring Program.. Can. Tech. Rep. Fish. Hydrogr. Ocean Sci. 194: vii + 57 p.

7 Tables

Table 1. Example of sampling requirements for five species that are regularly captured in the annual southern Gulf of St. Lawrence September survey.

Species	Example of sampling requirement
Atlantic Cod (*Gadus morhua*)	Collect length, sex, weight, maturity. Collect otoliths from: - one specimen per cm per set for specimens 25 cm and under; - two specimens per cm per set for specimens between 26-45 cm; - and three specimens per cm per set for specimens 46 cm and over. Collect 10 specimens from the Banc des Américains MPA that are between 40-70 cm.
Atlantic Herring (*Clupea harengus*)	Collect length from every specimen; preserve 2 specimens per every 5 mm length bin per set.
Atlantic Halibut (*Hippoglossus hippoglossus*)	Collect length, sex, weight, maturity, stomach weight and otoliths from every specimen. Collect 300 fin clips per mission. Collect 25 female gonads per mission.
American Lobster (*Homarus americanus*)	Length and sex from one specimen per sex per 3 mm per set. If female, check for lobster eggs. If female size is greater than 70 mm, check molt stage and shell disease index. If female size is greater than 70 mm AND has eggs, check abdominal width, egg stage and clutch fullness rate.
Winter Skate (*Leucoraja ocellata*)	Collect length, sex, weight, maturity, disk width and tail description from every specimen.

8 Figures

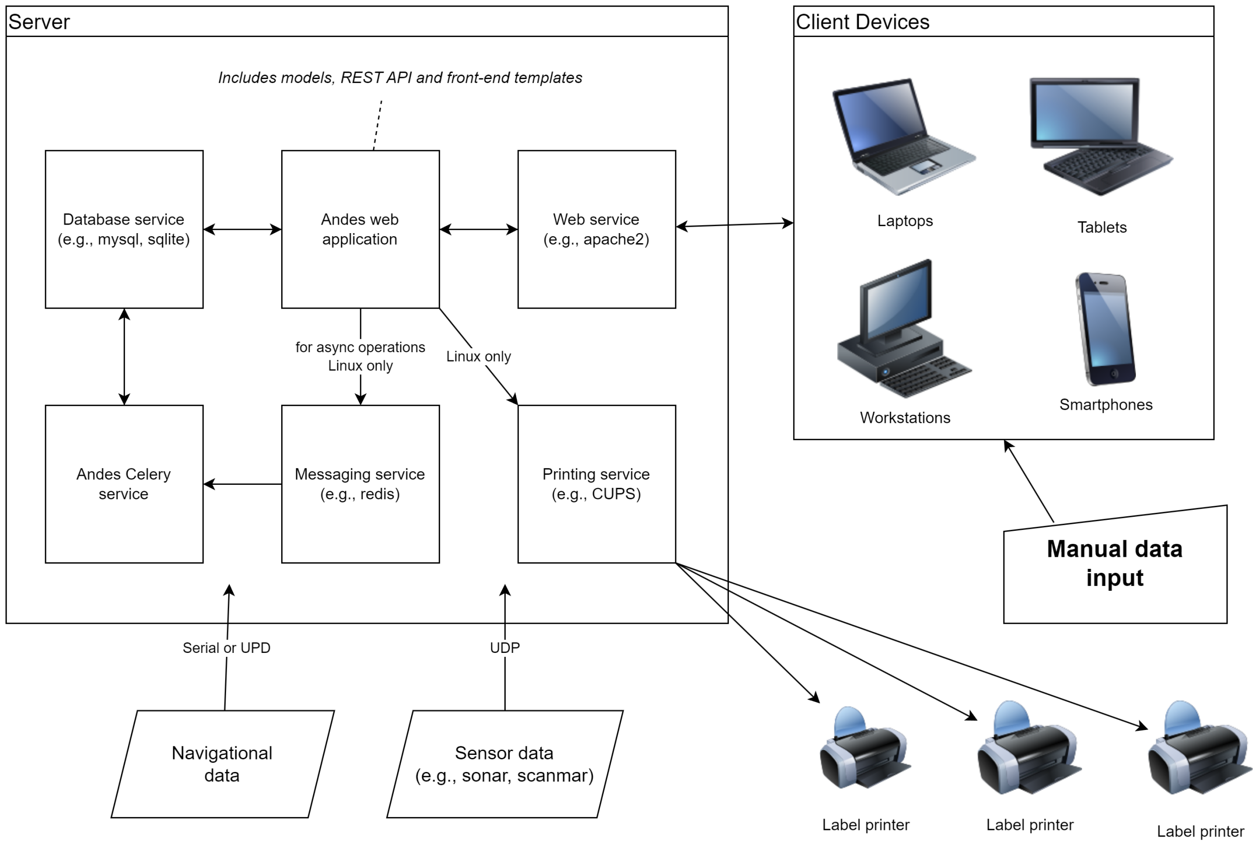

Figure 1. The generalized system architecture diagram of the Andes application and its associated services.

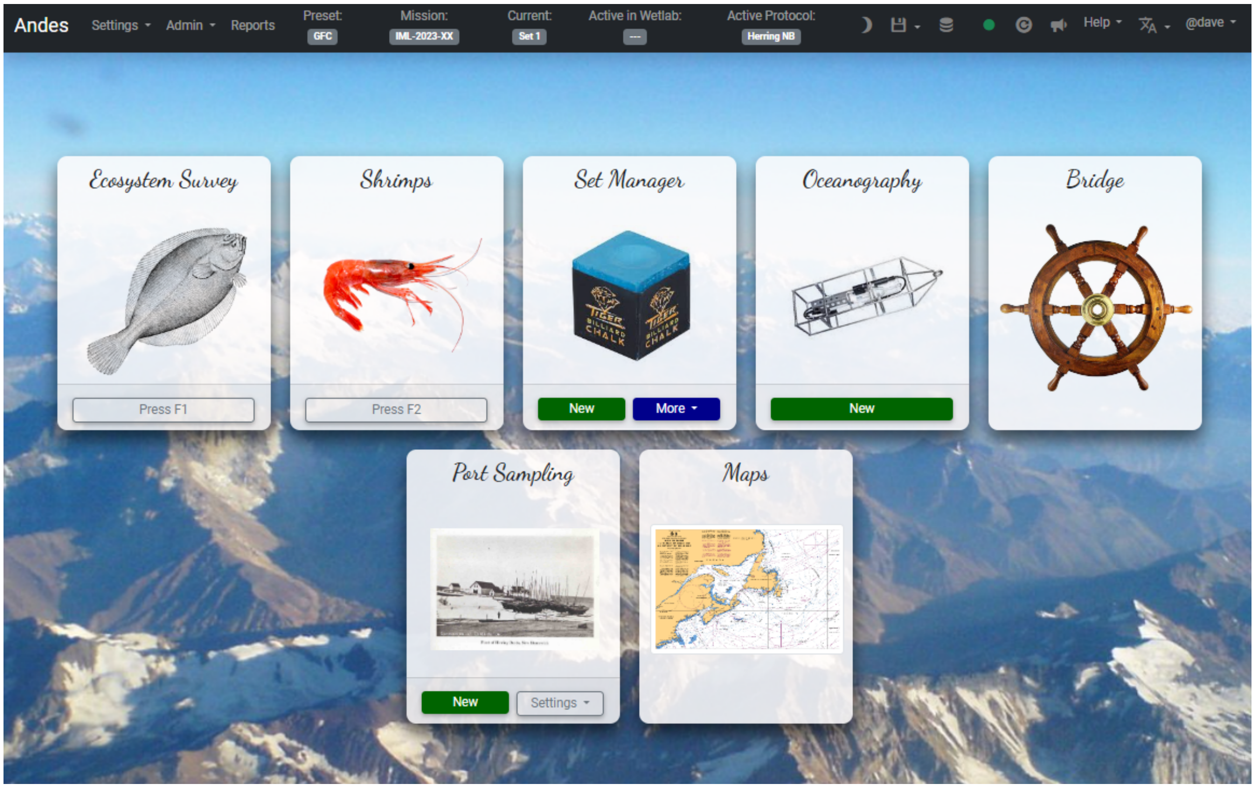

Figure 2. A screenshot of the main index page of Another data entry system (Andes) showing the principal modules of the application. Which are displayed on the index page will depend both the users permissions and profile.

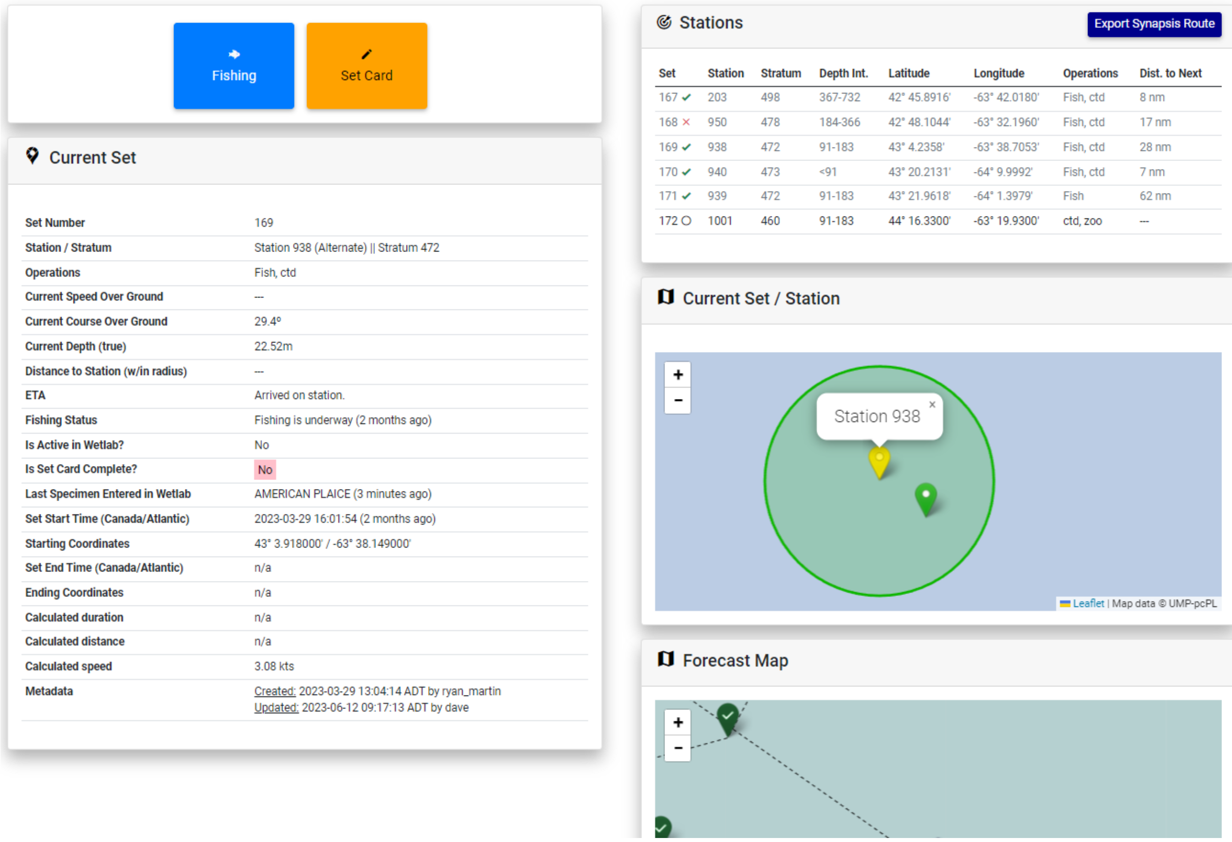

Figure 3. A screenshot of the Andes Bridge console. This app is used by bridge personnel across multiple user scenarios. This dashboard is a critical link in the communication between the scientist and crew personnel. An example of this is the station list; as a chief scientist plans out his or her route, queue stations and the associated activities will appear in the Bridge console. The Bridge console is also used by bridge officers to trigger the start and ends of operations and to input set metadata.

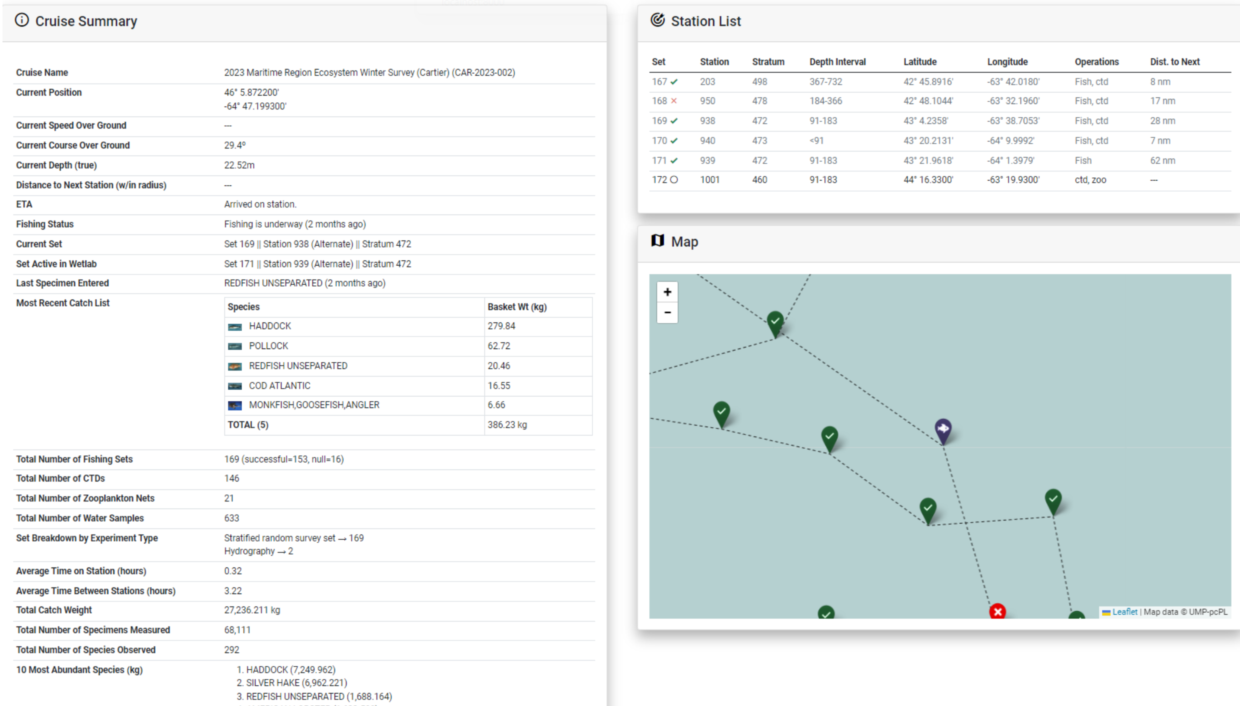

Figure 4. A screenshot of the Andes Cruise Dashboard. This dashboard helps to communicate queued stations, fishing activities underway and estimated times of arrival to science staff and crew personnel. The dashboard also displays high level summaries of Science activities that were conducted.

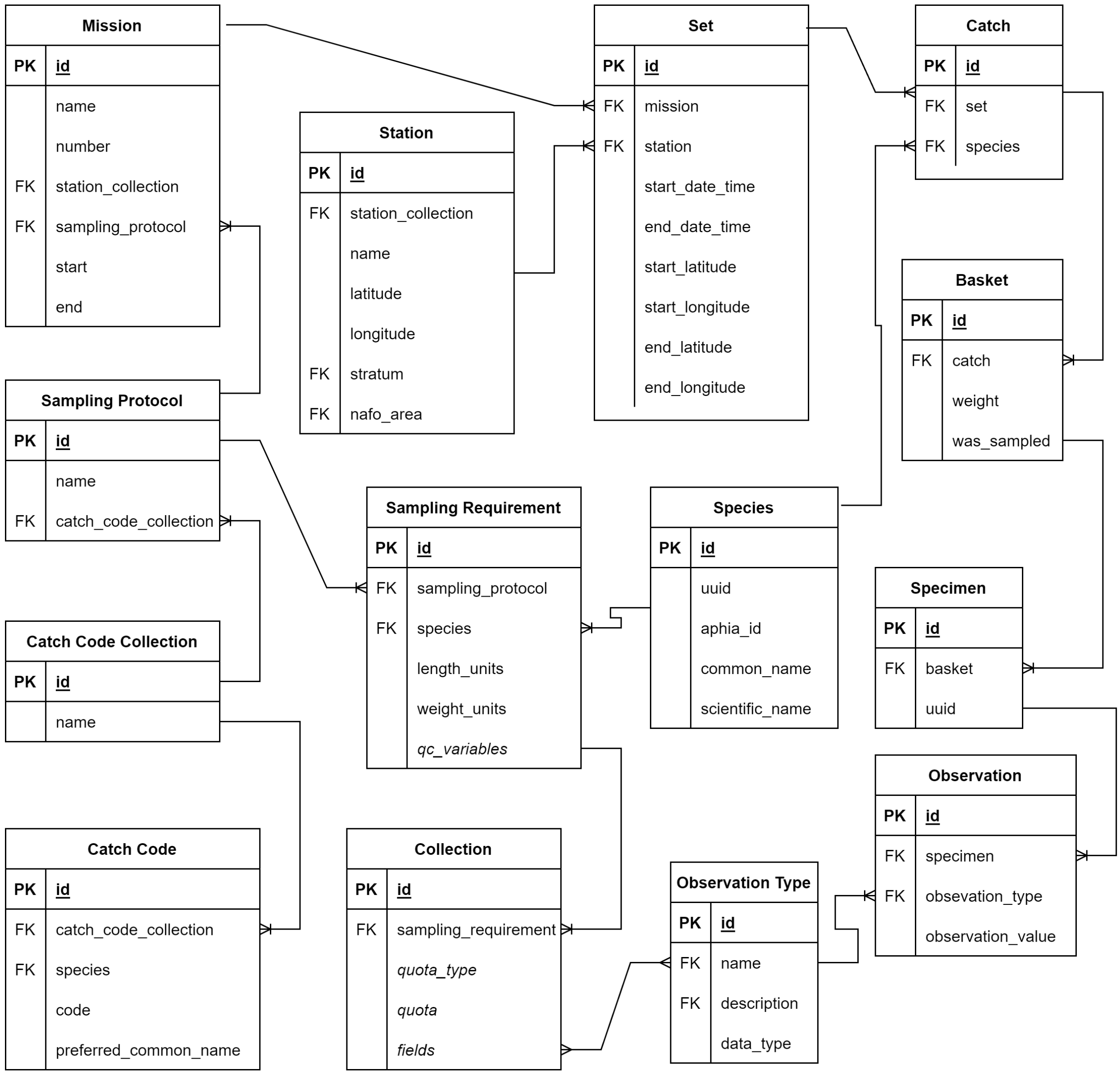

Figure 5. A simplified entity relationship diagram of the data model used by the Ecosystem Survey module of Andes.

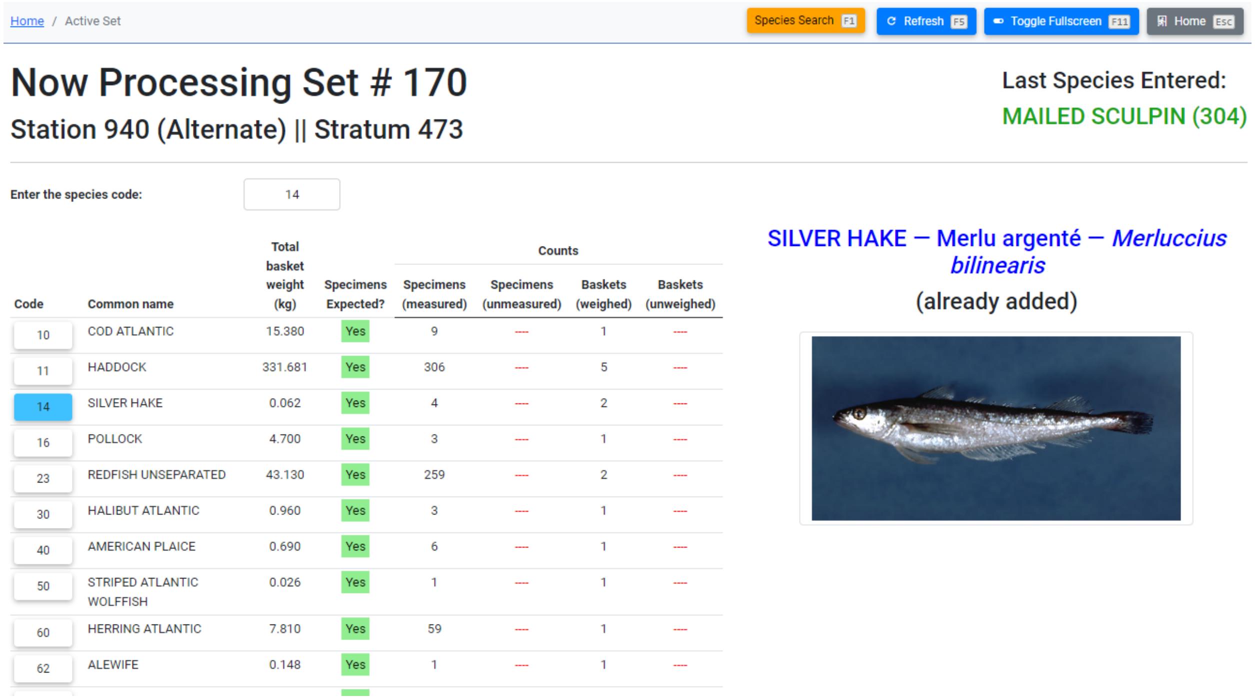

Figure 6. A screenshot of the Active Set page. This page is a component of the Ecosystem Survey module. Users in the wet laboratory enter new catches into this page as they are identified.

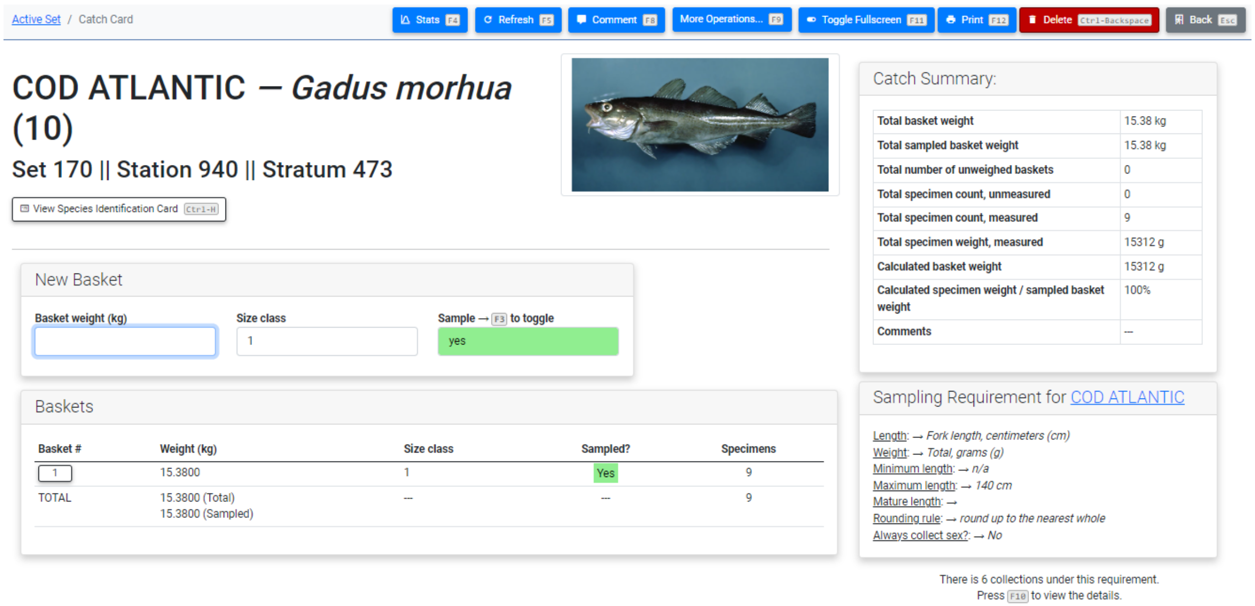

Figure 7. A screenshot of the Catch Card page of the Ecosystem Survey module of Andes. This is where new baskets and their associated weights are entered. This example shows the catch card of Atlantic Cod from a Maritimes Region survey of the George Bank.

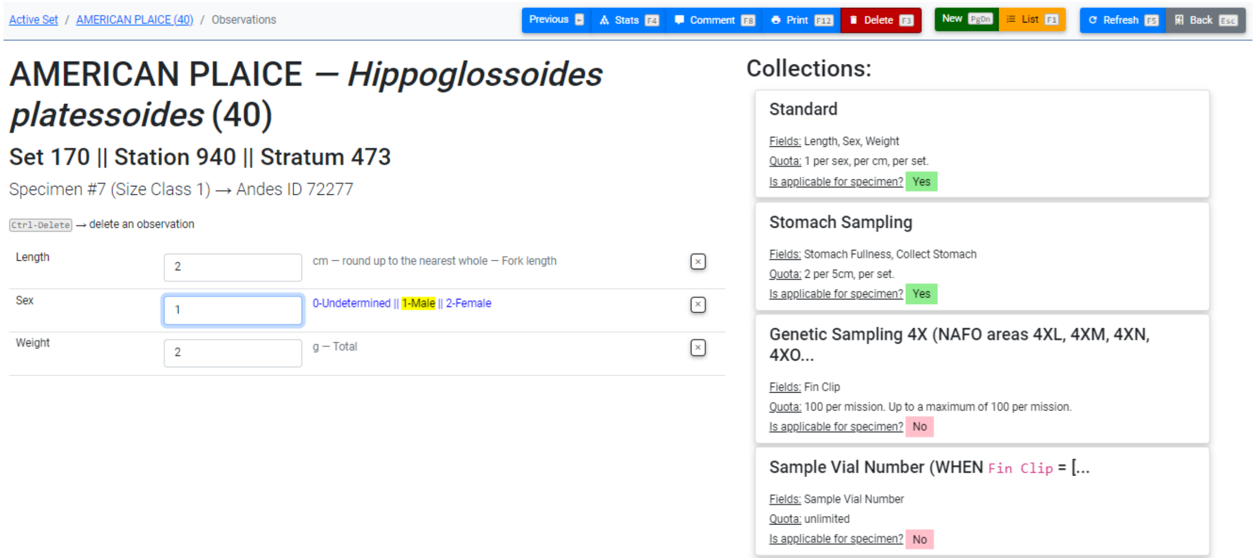

Figure 8. A screenshot of the specimen data entry interface of the Ecosystem Survey module of Andes. The specimen observations are edited on the left-hand side while the sampling collections are displayed on the right-hand side.

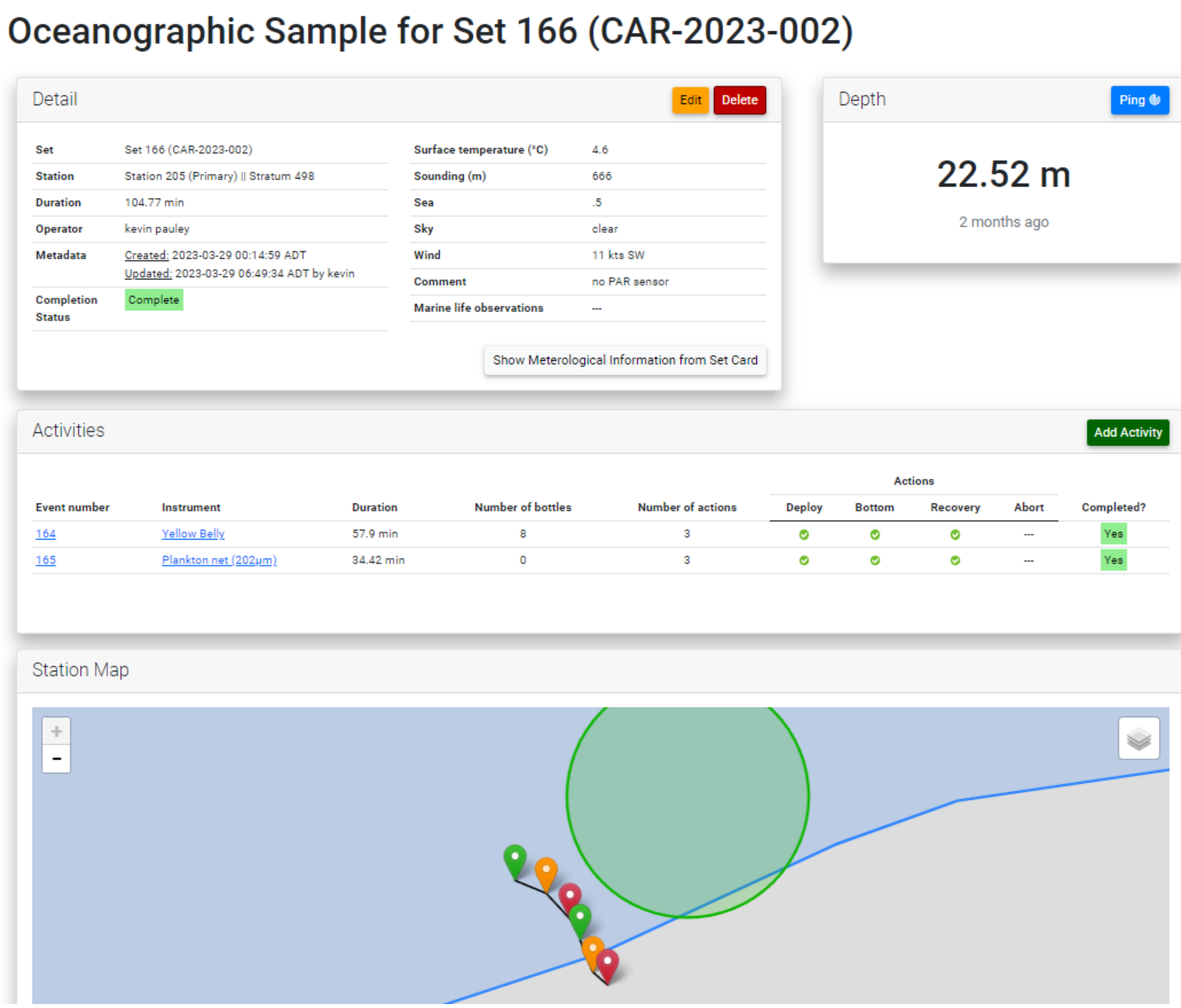

Figure 9. A screenshot of the Oceanographic Sample detail page from the Andes Oceanography module.

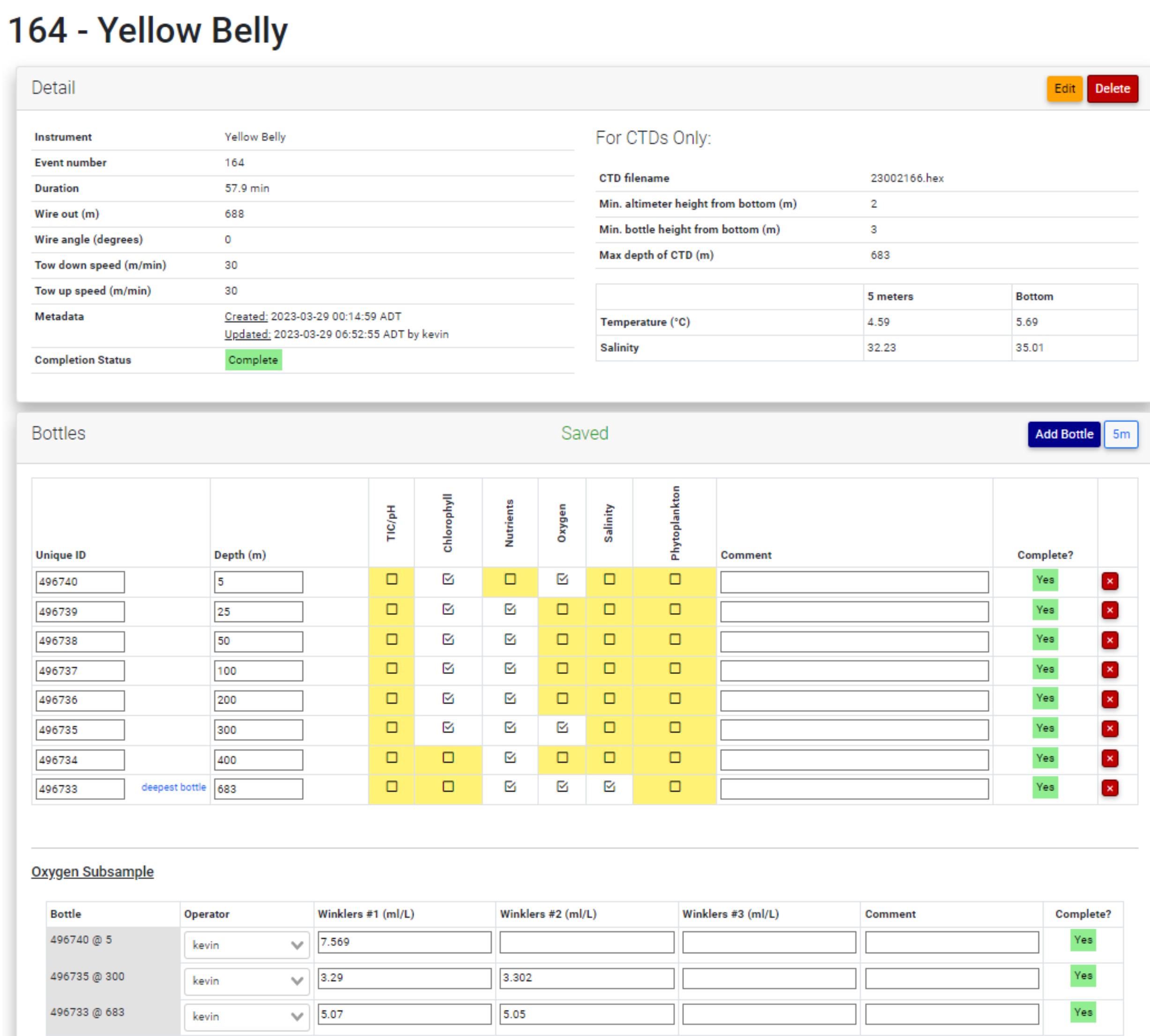

Figure 10. A screenshot of the Oceanographic Activity detail page from the Andes Oceanography module. Note that 'Yellow Belly' refers to the colloquial name of the CTD/Rosette system.

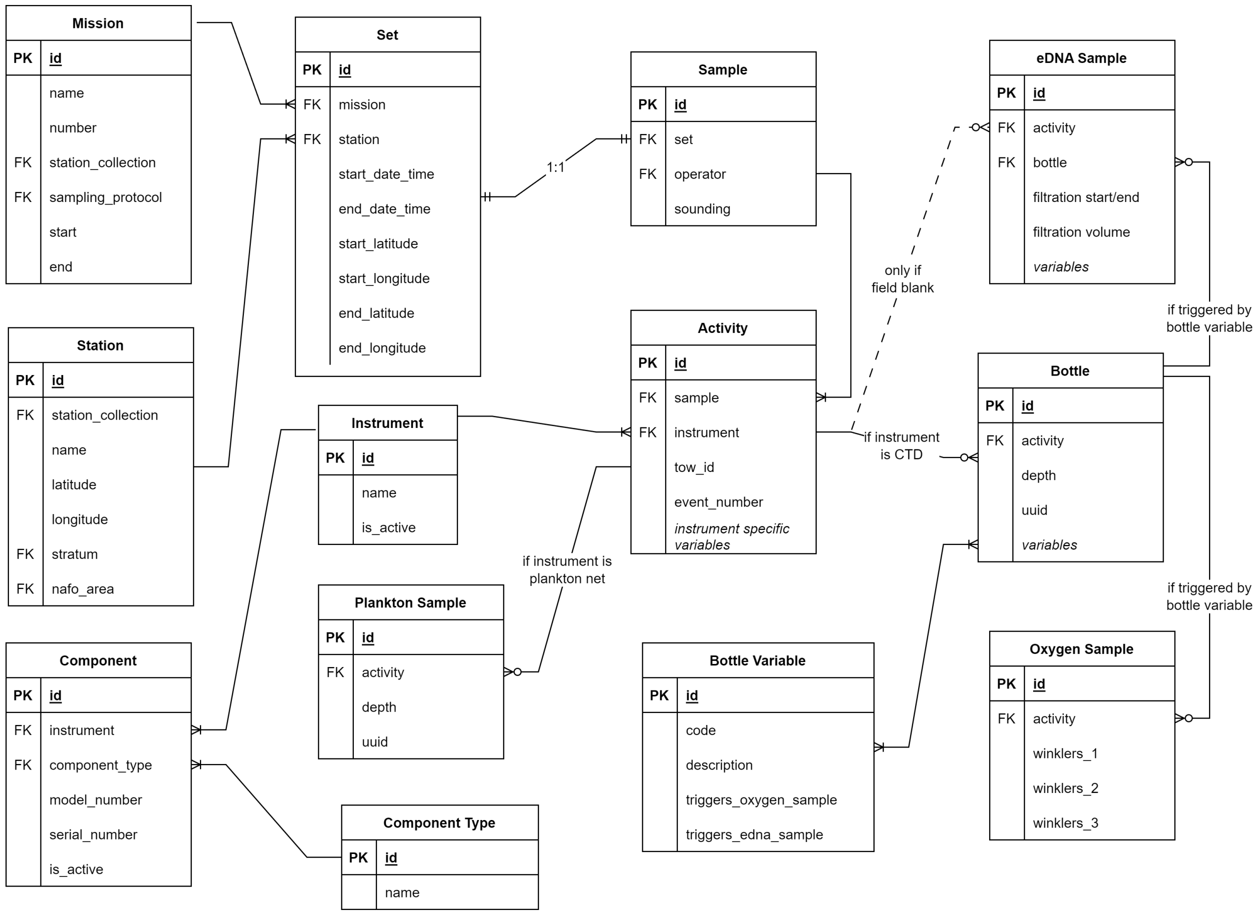

Figure 11. A simplified entity relationship diagram of the data model used by the Oceanography module of Andes.

Figure 12. A screenshot displaying the Port Sampling module protocol form. The selections made in this form will affect the behaviour of the module during data entry.

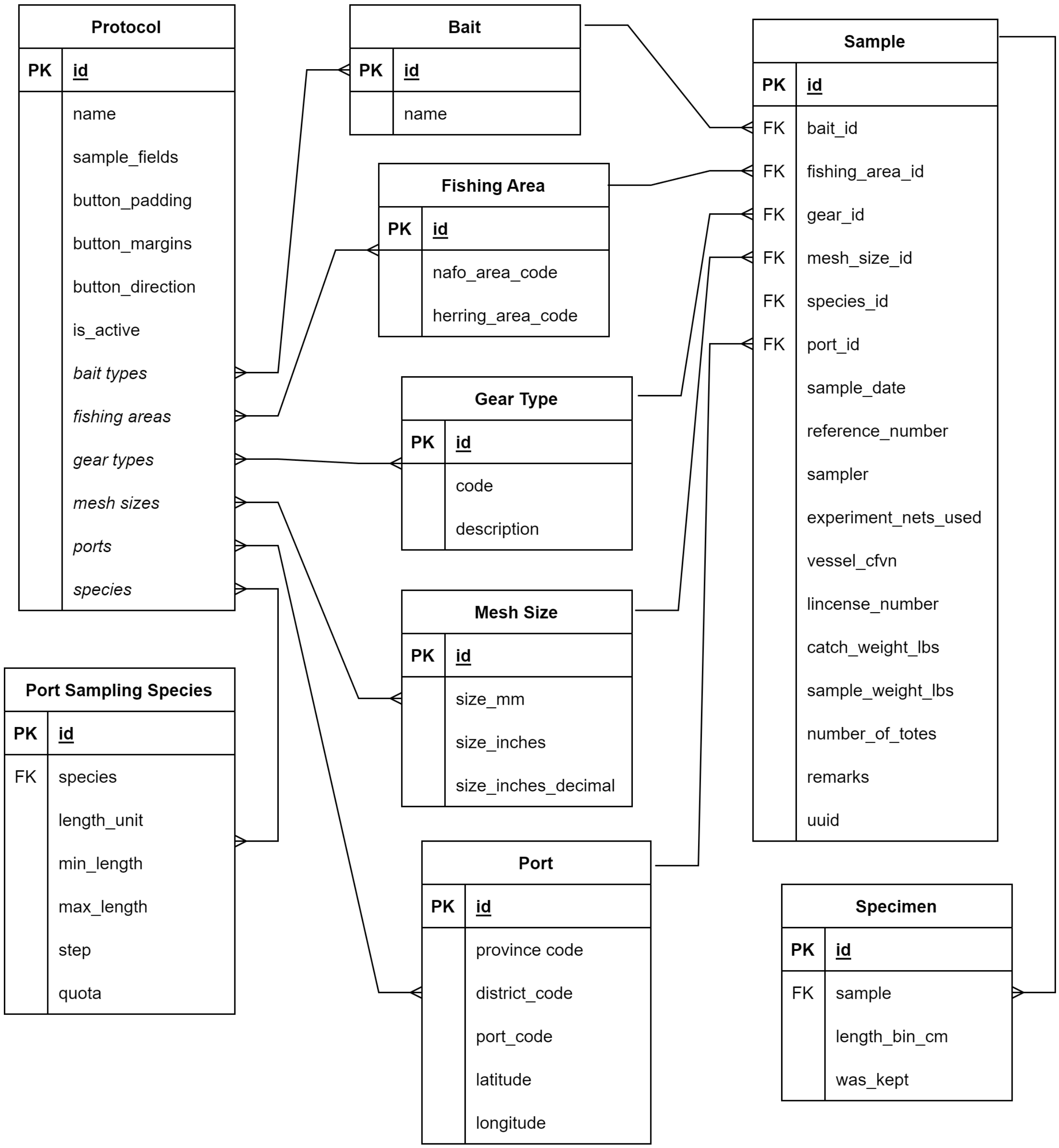

Figure 13. A simplified entity relationship diagram of the data model used by the Port Sampling module of Andes.

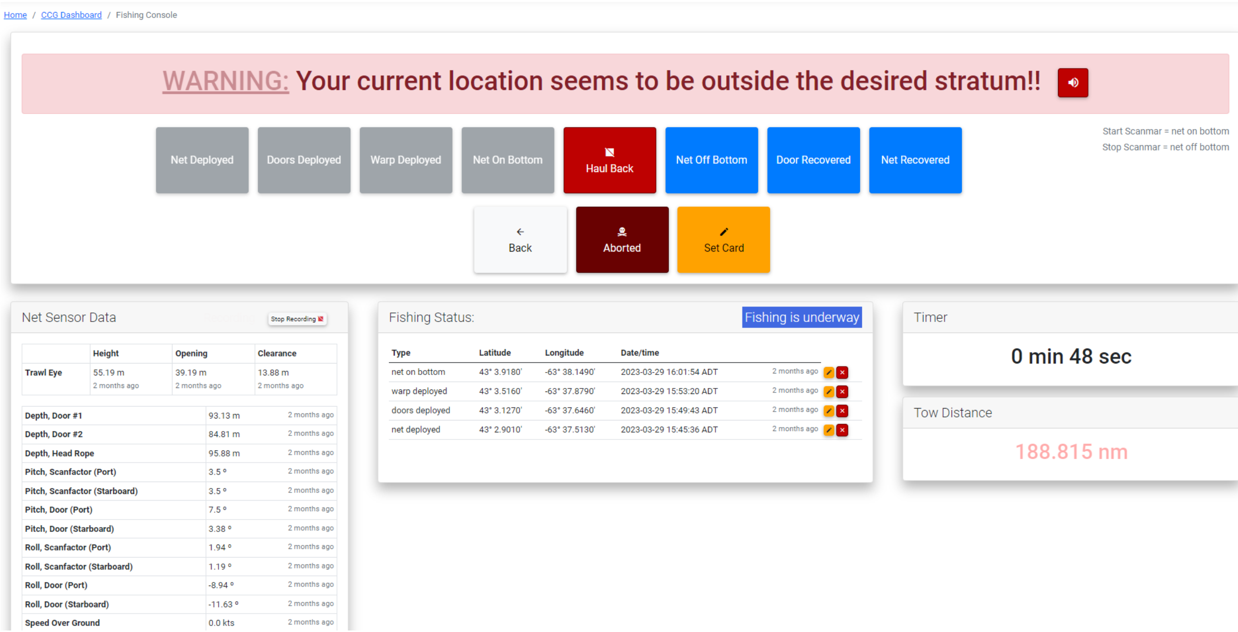

Figure 14. A screenshot of Andes Fishing Console withing the Bridge module. This screenshot shows some quality control validation features including: the fishing timer, the distance display and an alarm message for when the vessel is travelling outside the target stratum. On the bottom left of the screenshot, output from the net monitoring system is displayed. The large buttons in the center of the display are used to trigger fishing events.

Home / Track Forecast

Track Forecast

Sets and Queued Stations

Show Completed Sets

		Time / Date (Canada/Atlantic)					
Set	Station	Arrival	Departure	Distance to next (NM)	Odometer (NM)	Time to next (hours)	Status
---	---		2023-06-28 13:00	98.47	0.00	9.85	Start
---	---	2023-06-12 16:07	2023-06-12 16:07	45.91	98.47	4.59	Current Position
1	601	2023-06-12 20:43 (*)	2023-06-12 22:43 (*)	6.43	144.38	0.64	Queued
2	208	2023-06-12 23:21 (*)	2023-06-13 01:21 (*)	13.32	150.80	1.33	Queued
3	1011	2023-06-13 02:41 (*)	2023-06-13 04:41 (*)	88.05	164.12	8.81	Queued
---	---	2023-06-13 13:30 (*)		0.00	252.18	0.00	End

(*) estimated value

Figure 15. A screenshot of the Track Forecast tool.

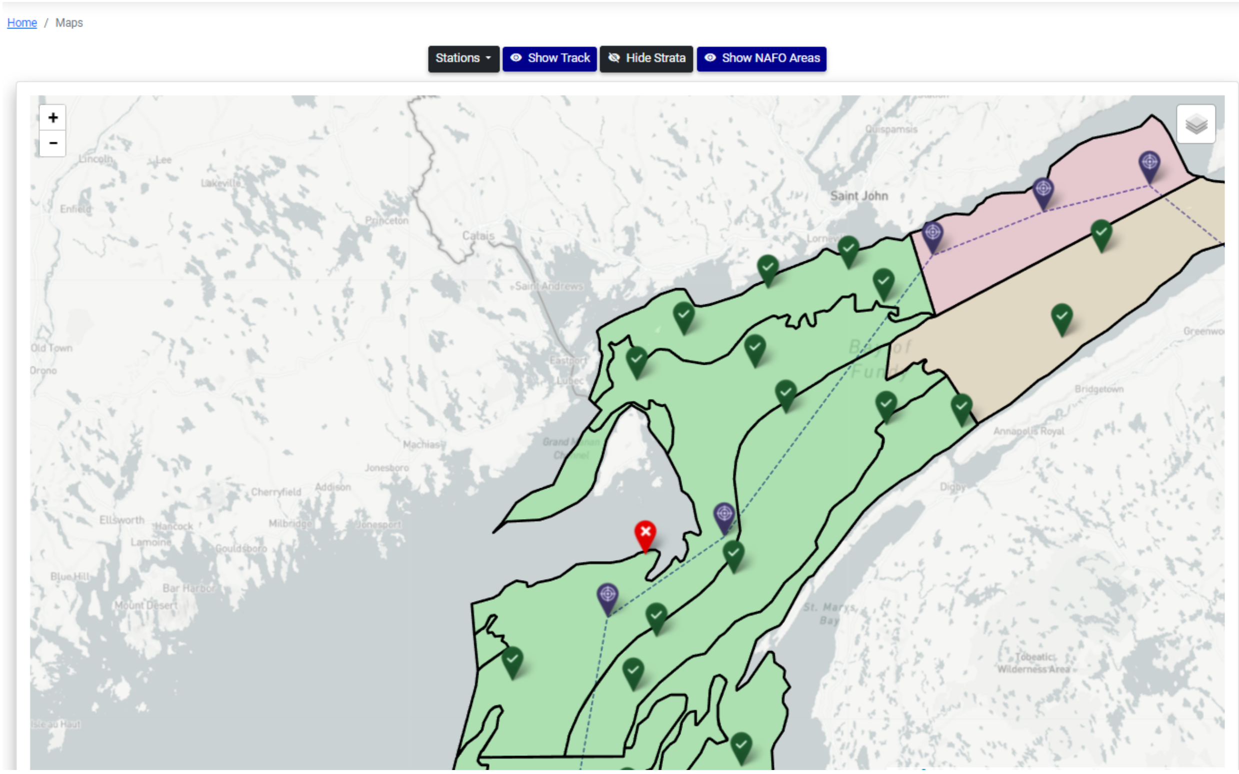

Figure 16. A screenshot of the Andes Progress map. The map colors strata by the status of their set quotas. The green strata are those which have met their specified targets of set conducted; those in yellow have not met the target but have surpassed the minimum; and those in red have not conducted the minimum specified number of sets.

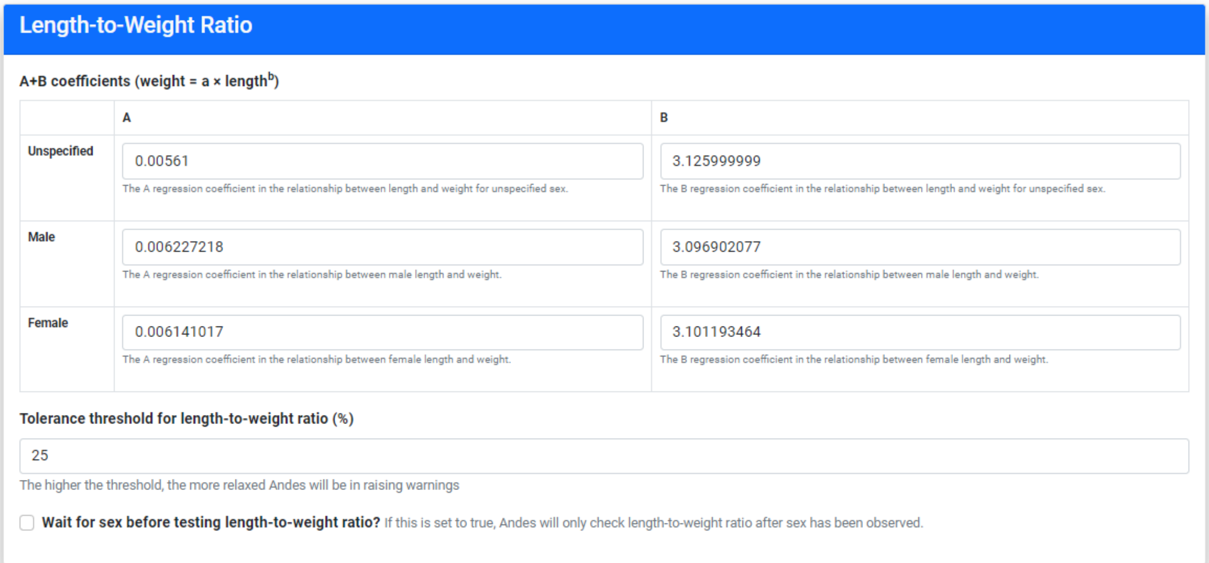

Figure 17. A screenshot of the “Length-to-Weight Ratio” section of the sampling requirement form. Andes offers the option to specify regression coefficients for males, females or unspecified individuals to be used in quality control.

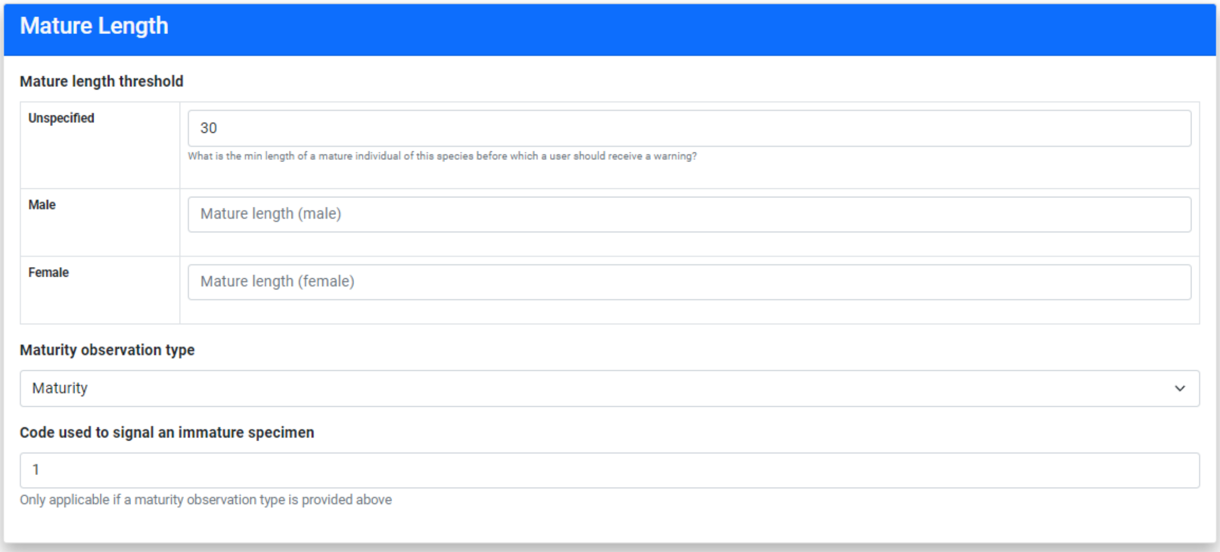

Figure 18. A screenshot of the “Mature Length” section of the sampling requirement form. Andes offers the option to specify mature length thresholds for males, females or unspecified individuals to be used in quality control.

9 Acronyms

Andes Another data entry system.

API Application Programming Interface.

AZMP Atlantic Zone Monitoring Program.

CSS Cascading Style Sheet.

DFO Fisheries and Oceans Canada, formerly the Department of Fisheries and Oceans.

ERD Entity Relationship Diagram.

ESE Ecosystem Survey Entry.

GSE Groundfish Survey Entry.

html5 Hypertext Markup Language.

IMTS Information Management and Technology Services.

LAN Local Area Network.

MPA Marine Protected Area.

MRR Module des Relevés de Recherche.

REST Representational state transfer (REST).

REST-API Representational state transfer (REST) Application programming interface (API).

VCS Version Control System.

WAN Wide Area Network.

10 Glossary

Andes Another data entry system is an application developed by Fisheries and Oceans Canada to support data collection for a variety of scientific programs.

Andes lead A person with a functional understanding of the inner workings of Andes. This person has the knowledge and user-rights that permit them to configure a Mission, define a Sampling Protocol, Sampling Requirements, etc.

API An Application Programming Interface (API) is a particular set of rules and specifications that a software program can follow to access and make use of the services and resources provided by another particular software program that implements that API.

AZMP The Atlantic Zone Monitoring Program collects and analyses the biological, chemical and physical oceanographic field data in the four Altantic regions of Fisheries and Oceans Canada.

Bridge The Bridge app is meant to be used by the navigation officer while the fishing officer deploys and retrieves the trawl. The operator inputs fishing events (eg., net deployed, doors deployed, winches locked, net on bottom, haul back, net off bottom, doors recovered, net recovered) or actions (eg., net on/off bottom) directly to Andes via the Bridge app. .

Catch A table of the Ecosystem Surveys app. A Catch table contains a link to Set and Species entries, as well as an optional: speciment count, unweighted baskets, relative abundance category and an invertabrate catch ratio. Catches can be subdivided into sub-catches having a parent-children relationship.

Catch Card Catch details recorded prior to detailed sampling.

CCG The Canadian Coast Guard.

Closed Set An Open Set Can be closed if the data it contains (eg, from Catch Card) does not trigger a Set Flag.

Configuration Preset Configuration for a particular deployment scenario. Configuration presets include details on the backup locations, and peripherals like GPS, sonars, label printers, etc. .

Cruise Same as Mission.

CSS A Cascading Style Sheet is a style sheet language used for describing the presentation of a document written in a markup language such as HTML or XML.

CTD Conductivity, salinity and density.

Current Set The Set that is currently occuring in time. It has begun and is not finished. There can only be one Current Set at a time.

DFO Fisheries and Oceans Canada is a department of the Government of Canada that is responsible for developing and implementing policies and programs in support of Canada's economic, ecological and scientific interests in oceans and inland waters.

Ecosystem Survey The Ecosystem Survey app is used to as the main data input interface. .

ESE The Ecosystem Survey Entry replaced the GSE in the early 2000s.

Fishing Event An event related to fishing activty of Set. Metadata such as GPS coordinates and time to are associated with the folllowing event types: *net deployed*, *doors deployed*, *warp deployed*, *net on bottom*, *haul back*, *net off bottom*, *door recovered*, *net recovered*, and *aborted*. A Set may contain a list of such event types. The Fishing Events are added to the Set by a crewmember using the Bridge application. .

GSE The GSE is a data entry tool developed in the 1980s.

html5 The Hypertext Markup Language is the standard markup language for documents designed to be displayed in a web browser.

IMTS Information Management and Technology Services is the IT branch of DFO.

javascript Javascript is a programming language that is one of the core technologies of the World Wide Web, alongside HTML and CSS.

JSON JavaScript Object Notation. The JSON format is a popular way to serialize data and is an alternative to the CSV (comma-separated values) and XML (eXtensible Markup Language) formats.

LAN A series of computers connected to each other and capable of communicating with each other over wired or wireless connections.

Mission Same as Cruise .

Open Set A Set that has not been closed. All new Sets will start in an open state and will remain so untill closed. Sets that are activated will automatically be opened. The Current Set is always open.

ORM Object Relational Mapping are an abstraction of relational entities (database tables) as objects.

Port Sampling Scientific program collecting samples from commercial fishing activities.

REST-API Representational state transfer (REST) Application programming interface (API) is an API that follows the REST software architectural style, created to guide the design and development of the architecture for the World Wide Web.

Sampling Requirements The species-specific requirements for samples to be collected during a scientific cruise.

Sampling Protocol The detailed description of what data is to be collected during a scientific cruise.

Set A Set contains all fishing activity and sampling results for a particular Station .

Set Flag A flag used to indicate that the data contained within the set has failed to pass a specific validation tests. Closing an Open Set with active flags can be done by overriding the validation mechanism.

Species An identifiable taxon that can be assigned to a species code.

Station A target location specificied by coordinates where a scientific activity is to take place.

VB Pre-.NET Visual Basic for Applications is the early version of Microsoft VBA and is no longer supported or updated by Microsoft.

VBA Visual Basic for Applications is a programming language built into most desktop Microsoft Office applications. More details can be found on the Wikipedia page for this programming language and from Microsoft.

VCS A Version Control System records changes to a file or set of files over time so that specific versions can be recalled later. For example, git is a VSC.

Vue.js A JavaScript framework for building reactive user interfaces. https://vuejs.org/.

WAN A telecommunications network that extends over a large geographic area, such as the internet.

APPENDIX A Tasks for administrative user

Prior to a survey, a number of configurations must be done for Andes to function as intended. At a minimum, the following information must be ingested by Andes in order for a mission to take place:

1. Strata polygons
2. List of stations
3. Species list
4. Sampling requirements

APPENDIX B Further resources about Andes

The main git repository supporting the development of the application is hosted by IMTS. The following resources can be found there:

- deployment checklist
- species lists for Gulf, Maritimes and Quebec
- strata polygons
- sampling requirements from previous surveys