Eurographics Conference on Visualization (EuroVis) 2020
M. Gleicher, T. Landesberger von Antburg, and I. Viola
(Guest Editors)

Volume 39 (2020), Number 3

VA-TRAC: Geospatial Trajectory Analysis for Monitoring, Identification, and Verification in Fishing Vessel Operations

S. Storm-Furru and S. Bruckner

University of Bergen, Norway

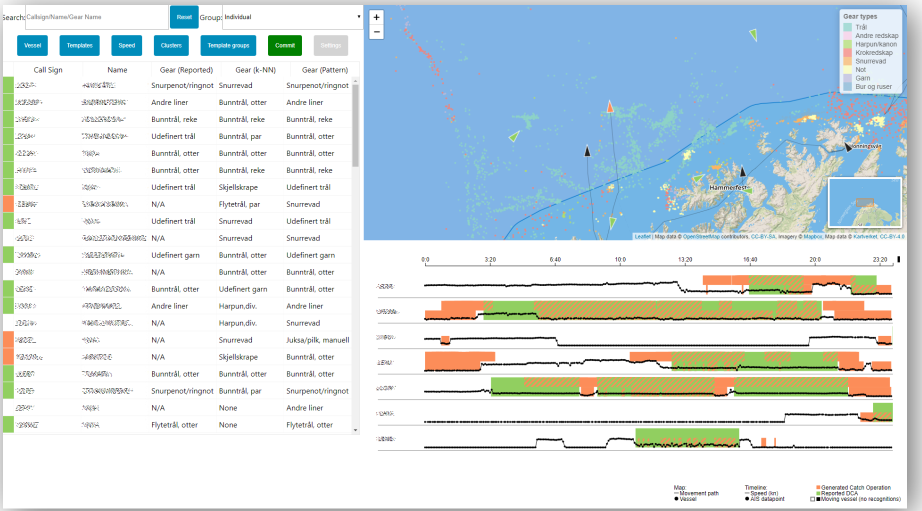

Figure 1: *Overview of the VA-TRAC interface. Vessel identifiers have been anonymized by a blurring filter.*

Abstract

In order to ensure sustainability, fishing operations are governed by many rules and regulations that restrict the use of certain techniques and equipment, specify the species and size of fish that can be harvested, and regulate commercial activities based on licensing schemes. As the world's second largest exporter of fish and seafood products, Norway invests a significant amount of effort into maintaining natural ecosystem dynamics by ensuring compliance with its constantly evolving science-based regulatory body. This paper introduces VA-TRAC, a geovisual analytics application developed in collaboration with the Norwegian Directorate of Fisheries in order to address this complex task. Our approach uses automatic methods to identify possible catch operations based on fishing vessel trajectories, embedded in an interactive web-based visual interface used to explore the results, compare them with licensing information, and incorporate the analysts' domain knowledge into the decision making process. We present a data and task analysis based on a close collaboration with domain experts, and the design and implementation of VA-TRAC to address the identified requirements.

CCS Concepts
• ***Human-centered computing*** → ***Visual analytics; Geographic visualization;***

1. Introduction

Seafood is an important global industry and Norway is the second largest producer of seafood internationally. It serves over 10% of the global fish market and the industry consists of many different domains, including fishery, processing and aquaculture, each employing thousands of people. The Directorate of Fisheries is the main body fully dedicated to its governance, and provides control and analysis of most aspects of the industry. Among its central tasks is the enforcement of a strict regulatory system, designed to preserve a maintainable seafood industry while minimizing negative impacts to the ecosystem by ensuring healthy stocks and environmentally sound processes.

The Norwegian fishing fleet currently consists of several hundreds of vessels considered active by the Directorate of Fisheries [Fis]. These vessels are obliged to report their operations using a system known as the Electronic Reporting System (ERS), as well as their position using the Vessel Monitoring System (VMS). Many vessels are also tracked using Automatic Identification System (AIS) position tracking. ERS reporting is done manually by vessels' crews, and situations commonly arise where the vessels forget to report daily, or send delayed reports. These situations are by themselves breaches of regulation. Additionally, since ERS reporting is done manually, opportunities arise for more nefarious activities, and vessel crews can attempt to circumvent the regulations in various manners. It is up to the employees of the Directorate of Fisheries to ensure that reports have been sent and contain valid data.

The employees working with such report validation primarily include inspectors and control analysts. The control analysts mainly consist of people with regulatory and judicial expertise, and develop the regulations governing reporting. Inspectors are the operative unit, and physically verify that reporting is present and correct by visiting vessels and assessing that the reports are in accordance with the vessel's actual state. Their work therefore requires intimate knowledge of the regulations. Additionally, they perform investigative tasks to identify vessels that have performed illegal activities.

There are four areas of regulations that determine whether a vessel's fishing operations are illegal or not. One of these is the duty to report ERS and VMS messages correctly and on time. In addition, the vessel must have the correct license and quota for the declared species of an operation, it must have acceptable quantities of bycatch, and it must stay within legal geographical regions. Currently, inspections are performed on a mostly case-by-case basis, meaning only vessels that are under suspicion of infractions are inspected. Such suspicions may be caused by tips from the general public or results from earlier inspections. Sometimes, inspectors will also perform random inspections of vessels in an area by boarding them while at sea.

Vessels required to report are projected to increase tenfold during the coming years, partly due to stricter regulations [Fis18a]. Inspectors and analysts working to validate the vessel reports are therefore facing an order of magnitude increase in their workload unless their work is augmented by automatic data analysis systems. As part of the effort to keep the inspectors' work efficient during the coming years, the Directorate has created a working group called the analysis group, consisting of statisticians, the aforementioned control analysts and data analysts dedicated to supporting and augmenting the work of inspectors through data analysis and automatic tools.

In this design study, we present the result of a close collaboration between visualization researchers and the Directorate of Fisheries to develop VA-TRAC (Visual Analytics for TRajectory Analysis and Classification), a geovisual analytics application for identifying irregularities in the fishing fleet's reporting of catch operations. The system has three operational goals: First, it should identify potential fishing operations automatically, let a user compare these with real reports, and enable domain experts to inspect vessel trajectories for suspect activities that may indicate breaches of regulations. Second, it should allow analysts to refine the identification algorithms. Third, it should be scalable to thousands of vessels, meeting the demands faced by the Directorate over the coming years.

2. Related Work

Perhaps unsurprisingly, fisheries management as a scientific field is dominated by articles about fish biology and fishing stocks [Els], and works concerning data, automatic analysis, and visualization in the field are largely related to that. However, the integration of spatial data into models is recognized as an important component of fisheries analysis [CYT14]. For instance, the work of Booth presents a good example of such an integration [Boo00].

More recently, work that focuses on the use of vessel movement data for stock assessment has become more commonplace. For instance, Duchame-Barth and Ahrens have worked on the subject, such as on uncertainty in automatic classification of fishing vessel movement for the purpose of determining caught species [DBA17], and using automatic data analysis of VMS data for evaluating fish abundance in the Gulf of Mexico [DBSA18]. Chang and Yuan [CYT14] discuss the combination of observations and VMS data for estimation of fishing efforts. In a more foundational work, Bonham et al. [BNTW18] describe the standard AIS movement data format and discuss how one can use such data for analyzing vessel operations in ports.

VA-TRAC is situated within the domain of geovisual analytics, which is focused on the analysis of data with a geographic component [AAJ*07]. Data of this type has become increasingly common, as exemplified by the collection of VMS, ERS and AIS data from fishing vessels. A well-known academic exploration of geovisual analytics by Andrienko et al. [AAD*10] sets a research agenda and argues for understanding and taming the complexity of spatiotemporal data by visualization and interaction methods, including the use of multiple views and space-time cubes, addressing scalability issues by path clustering, and employing automatic location extraction. Another theme is the application of geovisual analytics to satisfy the needs of various user groups, and geovisual analytics for movement data analysis [AA13]. As elsewhere in visual analytics, domain-specific research and design studies have been commonplace in geovisual analytics, with application domains ranging from ecology [XD14], over sports [JSS*15], to taxi movement [HZY*16]. Maritime vessel movement is also represented. For example, the Ph.D. thesis of Scheepens [Sch15] presents research into techniques for visualization and geovisual analytics for maritime situational awareness, such as analysis of vessel traffic flows [SHvv16] and glyph design for uncertainty in vessel data [Svv14].

Previous works also present solutions for identification of patterns in vessel trajectory data. For instance, Enguehard et al. [EHD12] present a system to uncover distinct and interesting patterns in fishing vessel trajectories. Their approach uses the fractal dimension and velocity of trajectories to filter interesting regions in vessel movement in time periods of up to a month. The approach is generic across different types of patterns, and can be used for other applications than the identification of catch operations. However, due to how the model generalizes, it does not take into account domain-specific data useful to catch operation analysis, such as the specific movement patterns associated with different types of fishing operations and earlier catch reports. Another method, based on the classification of so-called motifs found in the vessels' movement patterns, is presented by Li et al. [LHK06], though their approach does not include any visualization techniques. Wang et al. [WMY*17] use visual analytics to investigate vessel movements in coastal areas. Their system presents visual summaries of trajectory clusters based on AIS as well SAR (Search and Rescue) data, while our approach is focused on identifying individual licensing infractions.

A commonality among existing systems for the classification of vessel movement is that they are focused on the vessel itself, not individual catch operations, and thus do not take into account data specific to the catch operations. When performing catch operation analysis, this makes evaluation at the level of individual operations difficult. Therefore, the approach of VA-TRAC may be advantageous for the specific domain of catch operation analysis, as it includes available data on catch operations and vessel metadata in the analysis, and focuses the visualization on operations instead of vessels.

3. Data and Tasks

Analyses performed using VA-TRAC are based on sets of electronic catch report (ERS) and automatic identification system (AIS) data. ERS reports contain a multitude of attributes, such as the name of the vessel's master and the vessel's call sign, gear type, reported species caught and quantities on board, which port the vessel came from and is going to, the report ID, the coordinated universal time (UTC) the report was sent, and the vessel's longitude and latitude. The included attributes for each report are dependent on the type of report the item contains. For instance, DCA reports, sent to describe the vessel's daily catch activity (hence the name) may or may not include catch operations. If they do, they have a set of mandatory attributes such as start and stop of the catch operations that have occurred during the day. Similarly, port arrival reports must include estimated time of arrival and a code identifying the port [Fis17]. Some attributes' meanings change depending on the report type. For example, in port departure reports the mandatory attributes longitude and latitude indicate the position the vessel intends to commence fishing, while in DCA reports they indicate the position where each catch operation was started [Fis17].

In addition to the ERS dataset, VA-TRAC uses position data in the form of AIS messages. In contrast to ERS data, AIS data contains a more limited set of fields, with no semantic differences depending on context. The AIS data available to VA-TRAC contains information such as the MMSI (a unique identifier of each vessel), the vessel's navigation status, its current location, velocity, course, and a timestamp. Two additional datasets are used by VA-TRAC. One consists of metadata for the vessels, including the vessel's engine size, length, and license. The other contains a hierarchical structure of equipment categories, referred to as the gear hierarchy.

The AIS dataset available for initial use in VA-TRAC consisted of all AIS messages sent by Norwegian fishing vessels in and around Norwegian maritime borders during 2017 and the first seven months of 2018. The reports are sent in five minute intervals as long as the AIS transmitter on board the vessel is activated. In terms of a framework for understanding movement data quality presented by Andrienkio et al. [AAF16], the AIS data is free-form, two-dimensional movement data with high spatial precision, a high sampling rate and without abrupt changes in position, meaning one can have relative confidence in the accuracy of the spatial dimension of the data. However, AIS data inherently contains a degree of missing positions due to drops in signal strength, commonly caused by mountainous coastlines or irregularities such as bad weather conditions. When the vessels are in open waters, they generally have few to no missing positions. The AIS dataset contains approximately 24 million items in total. The DCA dataset totals approximately 1.8 million items, derived from DCA reports collected between 2014 and 2019.

VA-TRAC is focused on the analysis of data in 24-hour windows, which is the maximal legal duration between two ERS reports. The number of daily active vessels that are required to report varies significantly depending on season and weather conditions, and may range from around 100 to more than 400. This number is expected to increase by up to an order of magnitude over the coming years, as more and more vessels will be required to report.

3.1. Specifics of the Application Domain

Since both the MMSI of the AIS dataset and the call sign of the ERS dataset uniquely identify a vessel, a mapping between the two can be established, allowing the use of AIS data to get the locations between each DCA start and stop for a vessel. By interpolating between these locations, one can derive the trajectories of individual catch operations, each trajectory thus becoming a continuous, non-branching movement path. Catch operation trajectories have various constraints that determine their location and spatiotemporal developments. The following specific features of fishing vessel operations were identified through several discussions with members of the analysis group at the Directorate of Fisheries:

- Certain catch operations cause distinct spatial patterns, often based on the gear type used for the operation. An example is line fishery, which results in straight trajectories with sharp, approximately 90 degree turns.
- As vessels must slow down to perform fishing operations, the development of a vessel's speed during a fishing operation causes distinct patterns when fishery occurs.
- The habitats of various fish species are often located in specific regions, which causes vessels' fishing operation trajectories to be located in these regions. The canonical example is the spring fishery for cod in Lofoten. Likewise, quotas and licenses may contain limitations to where the vessel is allowed to fish, also limiting the locations where the vessel's fishing operations are performed.
- Fishery for certain species is limited in duration, commonly to

a period of a few months. This is due to the movement of fish throughout the year, and for many species there are additional regulations stating that fishery is only legal within a certain timespan. For example, Greenland halibut fishery is only legal in the spring and fall [Fis18b].

- Catch operations' durations are influenced by the gear type used, and real catch operations have a duration of tens of minutes to hours.
- Some fisheries occur more frequently at certain times during the day. For instance, shrimp fishery often occurs early in the morning, as the vessels want to deliver fresh shrimp to customers during the day.
- The vessel's spatiotemporal patterns during a fishing vessel operation are to some degree dependent on the vessel's size, engine, hull, and other features of the vessel's construction.
- Fishery requires at least some form of movement from the vessel, even if that movement is only the vessel drifting due to sea currents and wind.

Vessels with similar characteristics, such as the aforementioned vessel size and gear type, often have similar licenses and quotas. Licenses can therefore be used to group the vessels based on the characteristics of their catch operations. Due to these characteristics being similar across similar licenses, the statistics department at the Directorate has created license groups, allowing vessels to be classified based on other vessels with similar attributes.

3.2. Indicator Algorithms

The analysis group is currently building a set of so-called *indicators* for inspectors. Indicators are values based on analysis of the data collected by the Directorate that provide indications of whether a vessel may have performed illegal activities and may be worth investigating. The goal is to alleviate some of the work the inspectors have to do by filtering down the possible candidates for inspections and to identify repeat offenders. Since it is difficult to categorize the vessels that perform illegal activities, inspectors must verify the reports of individual vessels. Therefore, tools that aim to make the inspector more efficient cannot use techniques to aggregate vessels while still letting inspectors do their job. VA-TRAC instead uses indicators to scale the inspectors' workloads, focusing on the assessment of individual vessels.

The indicators developed for VA-TRAC consist of three algorithms classifying whether fishing operations have occurred for a vessel during the last 24 hours. The algorithms are as follows: A threshold-based pattern recognition algorithm classifies whether a vessel's movement matches its movement during previously received DCA reports, and returns both the gear and the duration of a similar catch operation. A k-nearest neighbors model based on previous fishing operations provides the gear type, and a classifier based on the vessel's speed finds durations where the vessel was moving at a speed that may indicate it was fishing. The operations found by the indicators resemble a pair of start and stop DCA reports, and can therefore be combined with the AIS data to derive the trajectory of the operation. The indicator-based operations contain three attributes: The duration of the operation, the movement path of the vessel during the operation, and the gear used. Since vessels with similar licenses often have similar characteristics for their catch operations, the parameters of the algorithms are modifiable for each license group.

Speed Classification: The speed classifier identifies catch operations based on the AIS dataset, in an approach similar to the classification of fishing operations done by Enguehard et al. [EHD12]. An operation is derived by identifying all reports from the vessel's AIS data whose speed falls below a set threshold. For each group of sequential reports, the timestamp of the first and last element forms the duration of a catch operation. The speed classifier thus returns a set of durations, each indicating a potential fishing operation.

Gear Classification: A k-nearest neighbors classifier uses DCA messages reported over the last 5 years to provide an indication of a vessel's potential gear type. The k-NN classifier returns the most commonly occurring gear type of DCA reports within a given distance threshold of a vessel. If no operations fall within the radius, the classifier returns nothing. The k-NN algorithm was run early in the development phase to look for good initial settings for VA-TRAC, which resulted in two domain-specific constraints being added. Since different gear types for areas change over the course of a year, only the DCA messages that occurred in the current month are used. In addition to increasing the relative number of correct classifications, the number of data points are reduced to $1/12^{\text{th}}$ of the initial size, so the algorithm runs significantly faster. The other constraint is an upper limit to the number of elements within distance the algorithm considers.

Movement Pattern Classification: The final classification algorithm is used for recognition of movement patterns. The algorithm used for this is a modification of the *1$ recognizer* [WWL07], originally created to recognize finger gestures on smartphone screens. Since single-finger gesture patterns on a phone screen have similar spatial characteristics to the trajectory of a single vessel (a single start and endpoint, no branching paths, and continuity), the algorithm was considered to be a good fit. According to the original authors, the 1$ recognizer's accuracy is on par with other template-based pattern recognition algorithms for paths, while the algorithm is easy to understand and computationally cheap. The modified 1$ recognizer works by comparing a set of path templates representing the catch operations of various gears with the trajectory of each vessel. By convolving the template along the vessel's trajectory, one can compare the operation to the vessel's movement throughout its trajectory. At each point of comparison, a piece of the trajectory is selected based on the duration of the catch operation in the template. The piece is then normalized, scaled, and rotated to match the template as closely as possible, after which a score is calculated based on the distance from each point in the trajectory piece to the corresponding point in the template. This is done for all templates, and each comparison score that passed a threshold is marked as a potential catch operation. The pattern recognition classifier returns a set containing both the duration and gear type of all recognitions that score above the threshold.

3.3. Task Abstraction

The task framework of Brehmer and Munzer [BM13] was used for task abstraction. As common in software engineering [McC04], a traditional approach to requirement analysis based on functional and non-functional requirements [CNYM00] was chosen for VA-TRAC, with the functional requirements serving as the basis for task abstraction. The functional requirements were collected during the earliest days of the development process, through a group interview with eight employees of the Directorate, coming from the analysis group, the statistics department, the control department, and the IT department. The interview was conducted in an open manner, and focused on the question of how the participants envisioned either real or imagined workflows for identifying illegal catch operations. Talking points repeated by multiple participants were noted, and became the basis for the functional requirements. The requirements were gradually refined in conjunction with the identification of the indicator algorithms, through individual interviews and informal discussions with members of the analysis group. The following sections describe each task abstraction. The first two tasks are based on existing workflows for catch operation analysis, while the last three are based on requirements to VA-TRAC's workflow.

T1 – Identify Illegal Catch Operations: A catch operation is considered illegal if it isn't reported in a DCA report, if the DCA report lacks some of its required fields, if the operation takes place using non-licensed equipment or in a non-licensed location, or if it occurs outside the allowed period of fishery for a species. DCA reports that lack required fields are not accepted, and are therefore easy to identify using the existing systems at the Directorate. For operations that are contained in a DCA report, the equipment, location, and period of an operation can be analyzed by deriving the trajectory of individual operations in the manner discussed earlier, and the content of the report can be considered as more or less plausible based on the features of fishing operation trajectories. Inspectors will currently analyze vessels in this manner to look for infractions. However, for cases when the DCA report is missing, the inspector must explore the reports and trajectories of vessels looking for the features specific to fishing operation trajectories. If one is found, the inspector must evaluate that the operation plausibly took place by using tacit knowledge such as whether the location of the operation is popular for fishing. Only then do they know which vessels to target for inspections. This is a time-consuming and tedious process, and inspectors are therefore often reduced to either patrolling the ocean or responding to tips from the public to find illegal catch operations from vessels that do not report. The system needs to *derive* what it thinks is a fishing operation from the data by applying the previously discussed indicator algorithms to the ERS and AIS datasets. The result is, for each vessel, a set of trajectories that have been identified as a potential catch operation. Based on this derived data, the user wants to *explore* the operations as described above, to *identify* the *features* (location, spatiotemporal patterns, gear) of each potential operation and compare it to tacit and explicit knowledge of the vessel, to assess whether the operation is likely to actually have happened.

T2 – Explore Vessels' DCA Reports: To find vessels that lack reporting, inspectors explore the reporting state of vessels to find those that lack DCA reports for a given day, and then assess whether the movement of those vessels correspond to a catch operation. Vessels that have sent a DCA report in a given day have, at first glance, performed their required reporting, and the inspector can use this as a filter to find vessels that seem to be fishing without sending reports. In the initial group interview discussions, the first thing that was brought up was the possibility of seeing which vessels have reported catch operations and which have not, to be able to filter vessels in this manner. It was also mentioned that an overview of the vessels in a specific geographical region would be useful as well, for instance if an inspector is in the area and wants to find the reporting state of all vessels in their vicinity. Since these two types of overview have different characteristics when abstracted into tasks, they have been separated into two subtasks.

T2.1 – Overview Based on Reporting State: When getting an overview based on reporting state, the user wants to *discover* which vessels have reported catch operations and which have not, in order to filter the set of vessels for which to further analyze catch operations. The user is in this case looking for a known target, a vessel with absent DCA reports, though the reporting state of each vessel is unknown, and must be *located*. The goal is to reduce the number of vessels that are considered for further analysis by creating a *summary* of the *outlying* vessels where no reports have been received and where the indicators have found a potential catch operation. The task may also involve filtering away vessels that have not reported, in cases where the user wants to assess the legality of catch operations performed by vessels that have sent DCA reports.

T2.2 – Overview Based on Location: The task of seeing vessels in a specific region is somewhat different. In this case, the user still wants to *discover* vessels' reporting states, but they know the geographical location of the vessel and are looking for vessels in the region that have not reported. They are therefore *looking up* specific vessels that they consider suspicious based on the fact that the vessels are not reporting. The goal of the task is to *identify* a specific vessel. The targets are once again outliers, i.e., vessels that haven't reported catch operations.

T3 – Evaluate Indicated Operations: When a possible catch operation is found in **T1**, the user must evaluate the likelihood of it being a real catch operation or not. In practice, this means comparing the results of the indicators to knowledge kept internally and externally by the user (such as knowledge of good fishing areas and active regulations), and will therefore necessarily involve varying queries and searches depending on the user's knowledge. For every potential catch operation found by the indicators, all the data features discussed earlier can be evaluated by the user in an attempt to verify whether the operation actually took place.

1. Distinct spatial patterns in a catch operation trajectory are evaluated by *identifying* the *shape* of the catch operation as a plausible shape for the gear type reported by the indicator. Such plausible shapes are found as tacit knowledge among several members of the analysis group and elsewhere in the Directorate. Equivalently, temporal patterns are *identified* by the *shape* of the changes in the vessel's speed during the course of the operation.
2. The likelihood that fishery has occurred in the location of the catch operation is evaluated by *comparing* the catch operation to earlier reports *located* close to the vessel, or by *comparing* the position to its geographical surroundings. The likelihood that fishery has occurred based on date is assessed by *comparing* the date with knowledge of regulations active at that date.
3. Assessing whether the duration of the operation is plausible is

done by *looking up* and *identifying* its duration. Similarly, the likelihood of an operation being performed at a certain timepoint can be assessed by *looking up* and *comparing* the start and end timepoint of the duration to the real catch operations of *similar* vessels or real reports from the same vessel.

The plausible methods of comparison differ in both search and query actions, as well as targets. However, they have several commonalities. When evaluating an indication, the vessel in question has already been identified and located, and the user is interested in validating its catch operations. Thus, the target for search actions is known, leaving *locate* and *lookup* as the possible search actions. The user either wants to *identify* or *compare* the operation's attributes (comparison with the user's knowledge is the goal even in identification tasks). The targets are either *shape*, when identifying spatial and temporal patterns, or the vessel's geographical region, *features* (duration, date and timepoint of the operation) when evaluating other temporal aspects of the operation, or *similarity* when comparing spatiotemporal patterns of operations to those of other vessels.

T4 – Evaluate Modifications to Indicator Parameters: Indicators are, as the name suggests, merely an indication of what can potentially be a catch operation. Further, the spatiotemporal trajectory of a correct indication is dependent on various external attributes of the vessels, such as license and size. The indicators take this dependency into consideration by being applied with different parameters depending on the license group. Due to changes (for instance gear changes, number of vessels, fish movement) in active vessels for different fisheries, licenses and quotas, the parameters for a specific group may need to be modified periodically. The user therefore needs to see whether such a modification increases the performance of the indicator for the current vessels. The first step when evaluating changes to indicator parameters is performing aggregation over vessels and calculating the effect on model accuracy caused by a parameter change, using the comparison of DCA reports and algorithms. As with deriving the indicated catch operations, this is a step that needs to happen before the user can get meaningful feedback, and is therefore done without visual representation. Thus, the task is a chain starting with a *derive* task. After deriving new information about model accuracy, the user wants to *compare* the result with previous results to see if the *trends* in model accuracy are positive.

T5 – Filter Vessels: When identifying illegal catch operations, the user is rarely interested in going through potential catch operations for all vessels in the same analysis session. As an example, an inspector may want to go straight to a vessel that previously has been caught performing illegal operations, or they may want to only look at vessels using specific gears that fish in highly regulated fisheries. The goal of all such filtering operations is to *derive* a smaller *summary* of the original dataset that includes the vessels considered as relevant for further analysis. Such a summary can include both a group of vessels with a common attribute, or a single vessel. When filtering by geographical region the goal is to *browse* unknown vessels in a specific region. In the other cases, the goal is *locating* vessels of unknown location based on a known attribute. The *features* the user is interested in are highly-context dependent, but include aspects such as the vessel's licensing group, reported gear types, reporting status, as well as the geographical region.

4. The VA-TRAC Workflow and Interface

Figure 2 illustrates the typical workflow of VA-TRAC in terms of the outlined tasks and Figure 1 shows a screenshot of the application's main interface. The workflow starts after the initial automatic analysis wherein indicators are run and stationary vessels are filtered out, with the user identifying vessels that may have been fishing without sending a DCA reports in the tabular view, possibly filtering the vessels by name, gear or license group. The vessels are then selected sequentially to analyze whether the indicators are correct for that vessel. The currently selected vessel is located in the map, which shows the trajectory the vessel has taken, and on the timeline, which shows the duration of catch operations and the speed of the vessel over time. The time and position where the vessel is categorized as fishing are highlighted using a consistent color scheme in both the map and the timeline, and are superimposed on the depiction of the vessel's trajectory. Seeing the actual events of the vessel together with the results of the automatic analysis, the user can form an opinion on whether the vessel may have avoided reporting a real catch operation, or whether the indicators were mistaken in their assessment. The color coding used throughout the main views is based on a categorical red-green-black colormap to represent the reporting state of a vessel. Red encodes vessels and operations that have been identified by the indicators, green corresponds to vessels that have reported a DCA and operations from DCAs, and black encodes vessels and trajectories where neither of the above is true. Despite concerns regarding accessibility, this choice was based on the comments of the domain experts due to existing conventions at the Directorate.

4.1. The Map View

The map view, seen in Figure 3, shows the spatial dimension of the vessels superimposed over a geographical map. Each vessel is marked by a dynamic glyph, which when zoomed out is a dot, and when zoomed in is a triangle pointing in the direction the vessel is currently going (seen in Figure 4). As shown by (1) in Figure 3, the glyphs are color-coded according to the reporting state of the vessel, with a white outline as a luminance contrast boundary, increasing popout [War13]. The green/red colormap on the glyphs allows the user to visually filter vessels that have reported DCAs, those that have been reported by the indicators, and those that are merely moving. Hovering a glyph shows a popup of the vessel's call sign ((3) in Figure 5), allowing for the identification of vessels. Being able to visually identify the vessels' categories and finding their identity addresses **T2.2** (vessel identification), where the user looks for the reporting state of vessels with known locations.

A challenge when working with spatial data is dealing with navigation between different granularities [AAD*10], and zooming into a map is one type of such navigation. This type of interaction occurs a lot in VA-TRAC. To alleviate occlusion and information overload, different glyphs are used depending on the zoom level. This increases the user's perception of vessels, as the dot glyph is horizontally symmetric regardless of rotation and is therefore easier

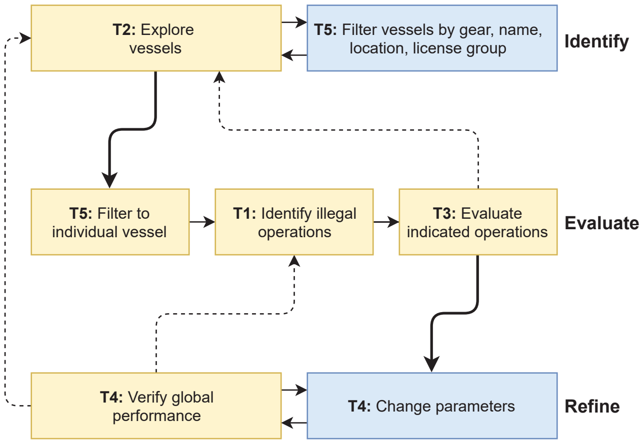

Figure 2: *The user's workflow based on abstract tasks. Yellow tasks are mandatory, blue are optional. The flow of user operations is denoted by arrows. The workflow can be grouped into the following three phases.* ***Identify:*** *After automatic analysis, the user identifies vessels of interest, for instance ones with indicated operations. If not looking for a known vessel, the user filters the dataset to focus on a specific subset of vessels.* ***Evaluate:*** *The user verifies whether a single vessel from the subset may have performed an illegal activity by analyzing its trajectory and reports, and comparing these to the results of the indicators.* ***Refine:*** *If unsatisfied with the results of the indicators, the user tweaks indicator parameters, verifies that the global indicator performance is not reduced, and either keeps refining, explores a new set of vessels, or returns to the vessel currently being evaluated.*

to perceive even at small sizes [BKC*13]. This comes at the cost of not being able to encode the direction of the vessels, but when zoomed out this is not a concern, as one is interested in getting an overview of the vessels' reporting situations based on location (**T2.2**). Additionally, the use of smaller glyphs reduces occlusion, which would hinder the task of getting an overview through the map. As the scale increases, fewer items are visible at the same time and their positions are farther apart on the screen, reducing the need for glyph symmetry, so additional data in the form of the vessel's current course is encoded by switching to an elongated triangle glyph. When zoomed in, each moving vessel's trajectory becomes visible on the map. Overlaid on the trajectories are vessels' movements during fishing operations, also colored according to the reporting state, letting the user identify the spatial component of real and indicated catch operations. The hatched pattern shows overlaps of DCA and indicator-identified trajectories, which is useful when the user wants to evaluate the overlap of indicator classifications and reported DCAs to see how well the indicators match the real world. Other mapping solutions in use at the Directorate encode the movement path as continuous, and the DCA reports as discrete points along that path. Though the data in the DCA dataset consists of discrete datapoints, the underlying activity is continuous, and the catch operations derived from the DCA reports are therefore encoded as such.

Figure 5 shows the map at a very close zoom level. At such a level, the thickness of the trajectory lines is increased to make both real and indicated fishing operations more visible. It thus becomes slightly easier to understand the movement patterns caused by indications of a specific vessel, and to solve the trajectory- and geography-based subtasks of **T3**. In Figure 5, one sees that (1) the vessel's reported catch operation overlaps the operation found by the indicators, and (2) that the indicators have identified more than the reported operation as potential fishing. As illustrated in Figure 2, **T3** is completed after the user has filtered the dataset down to a single vessel, and the user therefore avoids occluded paths even when the line thickness increases. The map view uses publicly available map data from the Norwegian Mapping Authority to display borders of Norwegian waters, annotated as (2) in Figure 3. A blue line represents the border between the territorial coast waters of Norway and the ocean regions controlled by Norway. The ocean regions are overlaid with a translucent geometric shape, which is shaded slightly orange. The border view was added at the request of domain experts at the Directorate of Fisheries, as many licenses and quotas are very dependent on these borders. For example, some vessels are not allowed to fish outside coastal waters, and many licenses disallow fishing in foreign waters.

Akin to the display of vessel glyphs, the display of trajectories depends on the zoom level. When too many vessels are displayed at the same time, the trajectories stop being useful as they occlude each other. Likewise, when evaluating an indicated catch operation in the map, one is interested in the shape of the path, which becomes difficult to discern when zoomed far out [BMMS91]. The choice of displaying trajectories or not is domain-driven: The interesting parts of a vessel's trajectory is when it is fishing, which occurs in limited geographical regions. Furthermore, the user is most interested in seeing the vessels that have not reported their daily catch. Therefore, trajectories of vessels that have reported a DCA are hidden by default and must be explicitly enabled, to not occlude the trajectories for indicated operations that are evaluated in task **T3**. As evident by comparing Figure 6 to Figure 7, occlusion from the trajectories of even two additional vessels can be quite high, and seeing the trajectory of the vessel of interest would become more difficult without the default filtering by DCA.

4.2. The Timeline

The timeline (Figure 8) shows the duration of every vessel's reported and indicated catch operations in the 24-hour window being analyzed. It also shows the vessel's speed over time, and the timestamps of each AIS report received during the period. It has two dimensions: The Y-axis is categorical, each category being a vessel (annotation (5) in Figure 8). The X-axis is quantitative, displaying the time of day. When the user wants to evaluate the temporal aspects of the indicated reports as part of **T3**, they already know what vessel and attribute they are looking for and are thus performing a lookup with the goal of comparison between different vessels or between different points in time. According to Munzner, the canonical example of an effective chart that fits such data and lookup tasks is the bar chart, which is used as the base of the visualizations [Mun14]. In addition to using the most effective available channels, the bar chart is ubiquitous and is readily understood by the domain experts at the Directorate of Fisheries. The temporal

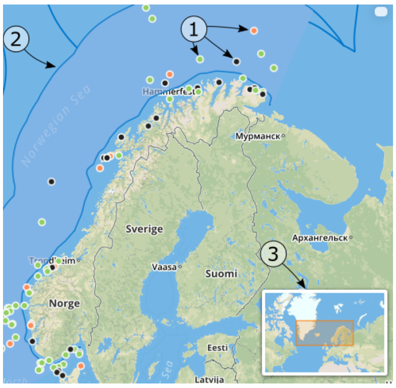

Figure 3: *Zoomed out view of the map. (1) shows vessels of different reporting states. (2) shows the borders of various ocean regions. (3) shows the minimap.*

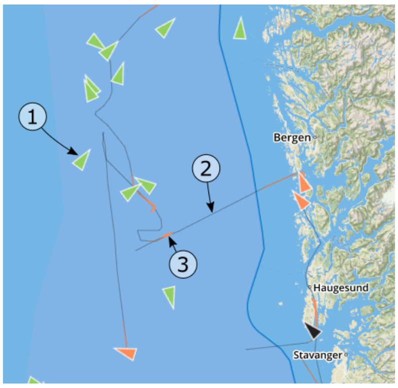

Figure 4: *A zoomed-in view of the map. (1) shows the triangular glyph used for vessels at this zoom level. (2) shows a vessel trajectory. (3) shows a catch operation.*

Figure 5: *A closely zoomed view of a single vessel. (1) shows the overlapping DCA-reported and indicated catch operation trajectories, (2) the indicated catch operation trajectory and (3) the tooltip of the vessel's callsign (blurred to anonymize the vessel).*

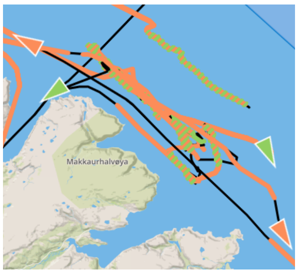

Figure 6: *Displaying all vessel trajectories.*

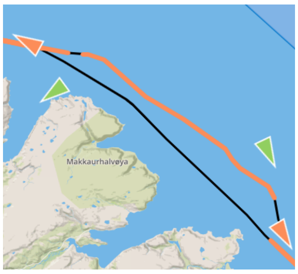

Figure 7: *Hiding trajectories of vessels with DCA reports.*

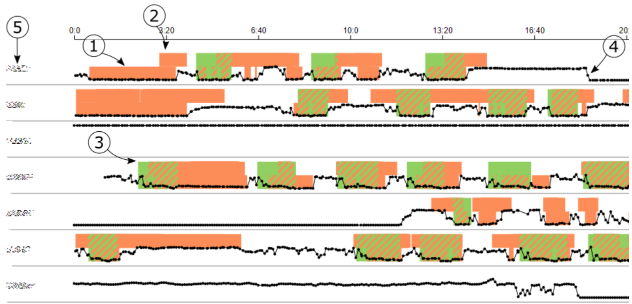

Figure 8: *The timeline view. (1), (2) and (3) show bars of the various catch operations. (4) shows the line chart showing vessel speed. (5) shows the Y axis displaying vessel names (blurred to anonymize the vessels).*

axis in the timeline is horizontally oriented to maximize the screen real estate devoted to it, as identification and analysis of the temporal aspects of catch operations are more important tasks than comparisons between vessels.

Annotations (1) and (2) in Figure 8 show bars displaying indicated catch operations. In contrast to a regular bar chart, the timeline positions each bar according to the duration of a catch operation. The bars use the same colormap as elsewhere in the application, and like the trajectories in the map view, use the red and green hatched pattern to display periods where both algorithmic and reported catch operations have occurred, as shown by annotation (3). Also, operations found by the speed indicator and the 1$ pattern indicator are separated spatially, to allow the user to discern how the system has identified the vessel as fishing. Annotation (1) shows results from speed indications, which are placed in the bottom part of the vessel's timeline. Annotation (2) shows results from 1$ pattern indication, which is placed in the top part. DCA reports cover both regions, as they are not tied to either indicator. Annotation (4) in Figure 8 shows the line chart of the vessel's speed during the day, superimposed on the bar chart. The dots encode the timepoint and speed of the vessel's individual AIS data points. Time is encoded as position along the X-axis using the same scale and labels as the catch operation bars, while speed is encoded as position along the Y-axis. The goal of the line chart is threefold. First, reductions in speed can be caused by fishing operations, so seeing the changes in speed during the day can be used to verify that the indicated fishing operations are within reasonable bounds. Second, it allows the user to get a deeper understanding of the vessel's entire trajectory, as it shows the rate of change in the vessel's position over time. Finally, the density of dots allows the user to understand the uncertainty of the automatic analysis, as they can see the number of data points used as inputs to the algorithms.

4.3. The Tabular View

The tabular view is a searchable text table where each row is a vessel. The attributes shown are the vessel's reporting state, name and call sign, the gear the vessel has reported using, and the ones recognized by the k-NN and 1$ pattern indicators. The vessel's states are encoded using the same colormap as in the previous views. The view is alphabetically sortable on a per-column basis. This allows for flexible on-the-fly pseudo-filtering, aiding task **T5**. For instance, the user can sort by reported gear type if vessels using a specific gear are of interest, or they may sort by reporting state. The use of a table to list the vessels was suggested by a domain expert as an intuitive way to quickly identify vessels that have been recognized by the automatic analysis (**T1**). Additionally, the table is used for **T2.1**, as the reporting state of each vessel is one attribute visible in the chart. Text tables are good at precise information lookup, and are familiar to many audiences [Sha16]. A text table to encode the textual data was therefore found suitable for vessel identification tasks where the user wishes to locate the vessel spatially, but where the vessel's reporting state is known. Further, the analysis group members have an intuitive understanding of table views, having used Excel as part of their daily work for a long time.

4.4. Indicator Views

The main views are focused on the core use case of VA-TRAC: investigating and analyzing trajectories and the results of automatic classification algorithms. There are several secondary views that provide users with fine-grained control over the parameters of the indicator algorithms. The template viewer (shown in Figure 12), for instance, is responsible for handling the manipulation of templates and parameters for the 1$ pattern indicator. The view shows all templates the indicator is comparing to the vessel's trajectory, and highlights the templates that have been recognized. It also allows the user to select individual templates to build new, averaged templates, and commit these to a database. Similar pop-up views are available for the other indicators. Every modification of indicator parameters is tracked and triggers a calculation of sensitivity and specificity measures based on previously inspected vessels. A corresponding chart allows for reviewing these values over time and enables users to revert to previous states, making sure that the global performance of the indicators does not decrease.

5. Implementation

VA-TRAC was implemented as a web application based on React (user interface), D3.js (visualization), and Leaflet.js (maps) on the client side, as well as Node.js and SQLite on the server side. The data API only performs the initial data transformation by querying the values and fetching those that are relevant. All following stages, including analytical stage operations, occur within the client, with the user providing view stage operations. The reasoning for such a "fat client" approach is twofold: First, since multiple users may want to modify the indicator parameters at the same time, initially computing changes locally reduces the potential for conflicts where two users modify the algorithm parameters on the server simultaneously, which could cause confusion in live operations. By keeping such analytical stage operations local by default, one can be sure that one user's parameter modifications don't interfere with those of others. Second, VA-TRAC may be used by inspectors while on board vessels, where the internet connection can be sub-par. Using a fat client allows the inspectors to open the application while in a location with a good connection, and as long as the application is kept running, it will continue working even if the connection falters.

6. Case Studies

In the following, we show two typical use cases of VA-TRAC. For both cases, we use an ERS dataset that contains DCA reports collected in the last five years (between 2014 and 2019). Vessels over 15 meters are obliged to send DCA reports at least once every 24 hours while they are active on a fishing trip (meaning between a port departure and a port arrival report), resulting in the dataset containing approximately 1.8 million items. As these are interactive workflows, it is difficult to faithfully represent them here in the form of static images and text. We therefore encourage the reader to refer to the supplementary videos for a more accurate representation of these case studies.

6.1. Identifying Unreported Catch Operations

The following case study follows the workflow of an inspector as they look for vessels to investigate. The inspector has launched the application and loaded all the data, and is about to start going through the vessels. Today, they are interested in analyzing the coast fishery vessels, and have therefore filtered the vessels to the first license group of coast fishery vessels: "Combined conventional and pelagic coast fishery". Their analysis starts by looking for vessels that haven't reported at all during the last day, and where the program has flagged a potential fishing operation. They sort the list view by reporting status.

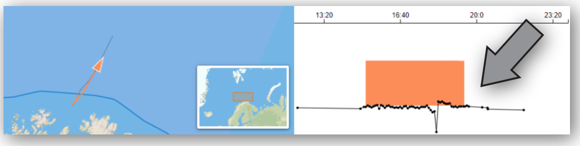

Figure 9: *The vessel's timeline shows that there are lots of missing AIS datapoints.*

One of the combined conventional and pelagic coast fishery vessels has an interesting movement pattern. The vessel has moved close to the coast, then looped around and moved away again. The system has flagged this movement as a potential fishing operation. The scarcity and irregularity of dots on the speed graph shows that the AIS reports have been sporadic, except for the recognized pattern, as illustrated in Figure 9. Due to the irregularity in the vessel's AIS reporting, the inspector decides to disregard the case, as the vessel may have technical difficulties and therefore has a valid reason for not sending its daily DCA report.

Having iterated through each vessel in the first group, the inspector chooses the next relevant license group: Conventional coast fishery vessels. Here, they spot another interesting vessel. The vessel, seen in Figure 10, has moved around an island over a period of a few hours. During the movement, the vessel has slowed down for a period of time, and has therefore gotten flagged by the speed classifier, identified by red marks in both the map and timeline.

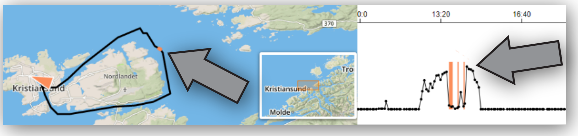

Figure 10: *The vessel has performed a fishing operation-like movement. Due to the proximity to Kristiansund and the airport, the inspector decides this is a false positive.*

The inspector sees that the speed changes of the vessel at the time could plausibly be a fishing operation. However, it would be very uncommon for a vessel to fish that close to a city, and there are plenty of other possible explanations for the behavior, such as the vessel washing its equipment. The inspector decides not to further investigate the vessel.

The inspector finishes verifying all vessels without DCA reports, and moves to verifying the DCA reports of the coast fishery vessels. In the pelagic coast fishery group, they find a vessel that has been flagged by both the speed and 1$ recognition classifiers for a duration, and where no DCA was reported at the time. Figure 11 (left top) shows the point in the timeline where this occurs. They see that blocks for both speed and pattern indications are present. In Figure 11 (bottom), they check the template in question and verify that it could plausibly be part of the pattern of a fishing operation using a seine (a vertical fishing net placed to surround a shoal of fish). The inspector finally checks the k-NN classifier to verify that the vessel probably has been using a seine. They enable the display of fishing operations in the area. In Figure 11 (right top), dots representing earlier fishing operations (a cluster is outlined with a circle) show that fishing using seines ("not" in Norwegian) is common in the area. The inspector chooses to flag the vessel for further investigation.

In the described use case, the inspector found two vessels that had performed potentially illegal operations. Both cases would have been very difficult to identify by looking at traditional maps of the vessels' trajectories, and benefited from being identified by automatic algorithms and verified using multiple views to explain the vessel's behavior. Due to the inspector's knowledge of the area and different forms of vessel behavior, the first vessel was deemed unworthy of further investigation. This type of knowledge is hard to encode in automatic models, and the advantage of having a human in the loop becomes apparent.

6.2. Building New Templates and Tuning Parameters

The other use of VA-TRAC is refining the indicators used for identifying vessels. In the following section, the typical workflow for indicator enhancement is presented. The analyst's goal for the session is building a recognition template.

The analyst starts their session by opening the template group editor and looks for license group and gear combinations that currently have no patterns assigned to them. The analyst sees that no groups contain templates for line fishery gear, and decides to add one. By sorting the vessel list by reported gear, the analyst finds

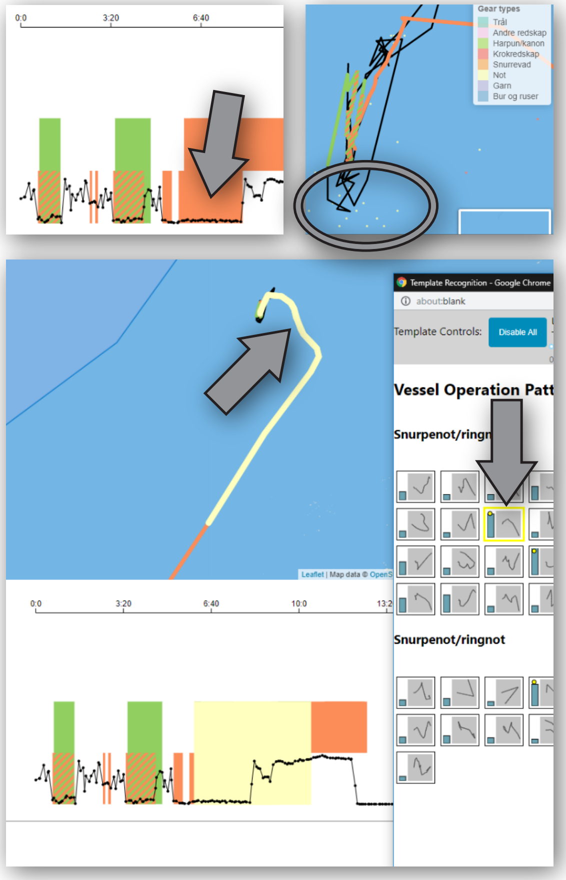

Figure 11: *Left top: A change in the vessel's speed suggests that an unreported fishing operation has occurred. Bottom: Hovering the template in the template viewer highlights the template in the timeline and map. Right top: The grey circle shows a cluster of seine operations in the area the vessel has been moving. Other seine operations are found scattered around the immediate surroundings*

vessels involved in line fishery. They then open the template editor that displays a list of all the templates available for the vessel, as shown in Figure 12. The analyst scrolls through the list, and searches for similar templates that describe the movement during line fishery well. Based on the analyst's knowledge of line fishery, they avoid the templates that are very generic (in this case the short ones with a single angle), and the ones with a very high fractal dimension. They find two patterns that describe a typical line fishery operation. The movement patterns of the operations are characterized by sharp, 90-degree angles and a blocky trajectory. They click on the two candidates to add them to the current average, as illustrated in Figure 12, and save them by pressing commit. From the next line vessel in the list, they decide to add a second template. They identify two patterns they want to use. This time, the average path gets too rounded, and some of the 90-degree angles disappear. To rectify this, they decide to add extra weight to one of the templates by clicking on it an additional time. The weighting causes

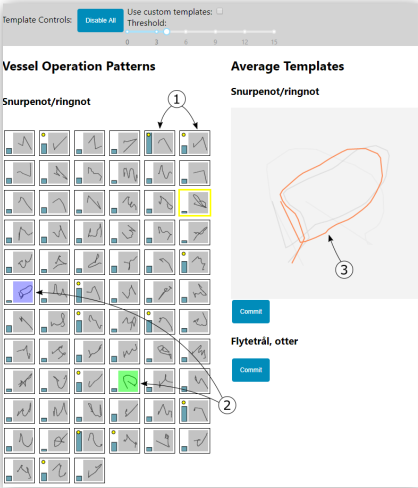

Figure 12: *The template editor. (1) shows two example templates. (2) shows templates used to build the average. (3) shows the average template.*

corners to get less rounded, and the sharp angles get preserved better for some regions of the average path.

Satisfied with the new patterns, the analyst moves to enhance the parameters of the speed classifier. They notice a vessel in the timeline that has sent a DCA report for a period where the speed classifier looks too strict. They select the vessel in the vessel list by searching for its call sign, then open the speed viewer. The vessel's license group is identified by finding the red bar in the histogram list. The analyst notices that many catch operations in the group occur at between 3.0 and 4.0 knots, so they increase the threshold from 3.1 knots. They can see the recognitions update in real-time in the timeline, and due to the result decide that 3.9 knots is a good threshold. We have observed this type of workflow many times when working with the analysis department at the Directorate, as the indicator parameters are still being tuned and evaluated.

7. Process and Validation

Early Stages: The first meetings with the Directorate of Fisheries were conducted as informal group interviews attended by employees from several different areas of expertise, including the statistics department, the IT department, the fisheries monitoring centre, the department of regulations, and the analysis group. Based on these discussions, identification and verification of illegal fishing vessel operations was identified as a primary challenge that could benefit form a visual analytics solution. The design phase initially consisted of detailed discussions with the domain experts, which culminated in expert reviews of several designs, developed using the "Five Design-Sheet" methodology [RHR16]. The reviews consisted of presenting the sheets separately to two people from the analysis group, who gave their comments. The comments were then compared, and used as a basis for choosing the most suitable design. In terms of Elmqvist et al.'s [EY15] patterns, the "Five Design-Sheet" methodology is a *paper baseline* pattern. The researcher creates the designs on paper, letting the users study and provide feedback on the design. An additional presentation of the system's design, this time focused on architecture and algorithm choices, was held for members of the analysis group and IT department before development started. Each of the chosen algorithms was discussed, with the goal of verifying that the algorithms were sound choices for fishing operation identification. Additionally, the meeting attendees discussed ways the system could be integrated into the data pipeline at the Directorate.

Development Stage: Validation in the development phase consisted of continuous justifications of each part of VA-TRAC's design, during which – based on constant interaction with the domain experts – several aspects of the design were adapted. Once the system's design started to settle and initial development had taken place, the system was presented again. The complete IT department and the most of the statistics department each attended a walkthrough of each component of the tool, and were encouraged to comment on the tool's functionality and utility. Most attendees did not have prior experience with the tool, though they possessed at least some familiarity with the ongoing development efforts. Such a review by a larger audience is akin to the *complementary participants* pattern [EY15], especially since the intent of the presentations was to understand how a generalized user, for instance an inspector, would react to the system.

Post-Development Stage: Towards the end of development, a data scientist from the analysis group was appointed as the initial main user of VA-TRAC. They were proficient with data analysis using digital tools and programming, but were relatively new to the job at the time and therefore less experienced in the intricacies of the domain. Various discussions with this person resulted in extra measures taken to ease the transition. For instance, the user engaged in *pair analytics* sessions, where they controlled VA-TRAC under guidance, and attempted a regular workflow. Similar sessions were held with another analysis group member, a statistician, who was less experienced with the use of digital analysis tools, but had much more experience with the domain due to working on statistics based on vessel reports. Each session culminated in an interview where the experts were asked about their thoughts on the system and the development process. During these in-depth sessions we uncovered weaknesses in both the visual encoding and the interaction idioms. For example, it was found that there was a need for more filtering options, as seeing all vessels at the same time was unnecessary for some tasks, cluttered the views, and reduced performance, leading to the addition of license groups not just as metadata for the indicators, but also as a user-enabled filter. Iterating through the vessel list using textual search was also found insufficient, so additional modes of navigation, through clicks on the vessel list, were added. Similarly, legends for the line charts in the timeline view were added based on these sessions.

At the conclusion of the project, a final presentation of VA-TRAC was held for stakeholders and members of the analysis group. The presentation was very hands-on, and sought to accurately convey a typical workflow using VA-TRAC, along the lines of the *once upon a time* pattern. The attendees also had the opportunity to guide and direct the workflow through questions and comments, thus providing a close-to-interactive experience without the system needing to be deployed. The Software Usability Scale (SUS) [Bro96], a quantitative scale of software usability, was used to summarize the impressions by the participants that had not been involved in the tool's design. While the number of participants is insufficient for any meaningful statistical assessment, the SUS is useful to provide an additional indication of the perceived utility of the system. VA-TRAC received an average total score of 73.125 from four respondents. The statements with the best results were "I thought there was too much inconsistency in the system" (indicating strong disagreement with this statement) and "I found the various functions in this system were well integrated" (indicating strong agreement in this case), with respective average scores of 9.375 and 8.75. The lowest scores were given to "I needed to learn a lot of things before I could get going with this system", average score 5.625, followed by "I felt very confident using the system" and "Most people would learn to use this system very quickly", average score 6.25. These lower scores were not unexpected, as the participants had very limited prior experience with visual analytics systems. Nonetheless, there was generally positive feedback on the benefits of the application, in particular in light of future increases in the number of reporting vessels.

Deployment: At present the solution is primarily being used by members of the analysis group, mainly consisting of people with expertise in statistics, data analysis, and development, regulations development, and judicial work. It is further developed for full integration with the Directorate's facilities. Currently, the Directorate is still in the process of adapting its IT infrastructure to support live use of VA-TRAC by inspectors. However, due to the positive feedback to the tool, and to visual analytics solutions (which few members of the Directorate were familiar with prior to our collaboration) in general, we are confident that the system will see wider-scale deployment in the future. Furthermore, there are several additional aspects of the Directorate's work that could equally benefit from visual analytics and we are currently investigating avenues for further cooperation.

8. Discussion and Lessons Learned

Based on the feedback by the domain experts at the Directorate and its plans for adoption of the tool, the approach to catch operation analysis taken in VA-TRAC is effective, and can help counteract the large increase in the number of reporting vessels that is projected to occur in the coming years. The speed of the indicator algorithms allows for the approach to scale to accommodate the increase in reporting vessels.

Based on the feedback from our domain experts, the currently integrated mechanisms for simplifying vessel glyphs and trajectories discussed in Section 4, in combination with filtering operations based on the DCA state and other attributes have proven to be sufficient at limiting the visual complexity also taking into account the projected increases in reporting. However, for scenarios with significantly higher vessel densities and an extended set of tasks, further aggregation and summarization strategies, such as those proposed by Wang et al. [WMY*17], could be a valuable addition.

One current limitation of the VA-TRAC is that it is designed to use data particular to the Norwegian regulatory framework, and although AIS is an internationally used data format for position tracking of vessels, the ERS reports are specific to Norway. Implementing a similar system in other countries therefore depends on the existence of similar reporting regulations. Similarly, the indicator algorithms partially depend on specific features in vessel trajectories during catch operations using gear common for Norwegian fishing vessels. Vessels of other nations may use different gears or otherwise behave differently, and the choice of indicators may therefore not be suitable for all occasions. Even so, many nations, including all EU nations, implement reporting regulations similar to Norway's, and vessels fishing in the North Sea (including vessels from the EU and Russia) generally use the same gear types in the same manner. The VA-TRAC approach can therefore be of value at the very least for these nations.

During the development and validation of VA-TRAC, several impressions were formed about the process of building a visual analytics tool for domain experts in a large governmental organization. As it is common in large organizations [Smi01], there are vast amounts of tacit knowledge the Directorate, but it was initially difficult to get in contact with the right experts for this project. As recommended by Sedlmair et al. [SIBB11], we spent considerable time on identifying the correct stakeholders in the early phases of the project and made sure that they were intensively involved in all aspects of the design by engaging them in frequent meetings and presentations. In order to facilitate this, one of our team members spent a substantial amount of time embedded at the Directorate [SF19]. Using a validation framework such as the nested model of visualization validation was essential in verifying the effectiveness of the proposed approaches and was of great utility in the communication with domain experts. Nonetheless, external factors such as the Directorate's delays in migrating its IT infrastructure, can never be fully mitigated and it is important to not lose sight of such aspects.

9. Conclusion

We have presented the visual analytics tool VA-TRAC, which is used for analysis of fishing vessel operations. The tool lets inspectors identify vessels that may be worth inspecting by running modifiable classification algorithms on AIS and catch operation data to identify unreported catch operations. Through design and validation performed collaboratively with domain experts at the Norwegian Directorate of Fisheries, VA-TRAC has shown promising results as a tool to augment regulation enforcement.

Acknowledgements

The research presented in this paper was supported by the MetaVis project (#250133) funded by the Research Council of Norway. The authors would like to thank the Directorate of Fisheries, particularly the members of the analysis group, for supporting this research, the insightful feedback, and the enthusiasm for this work throughout the process. In particular, we would like to thank John Erik Ågotnes and Kine Iversen for taking so much time out of their day to answer questions and their invaluable help in validating the system, and Per Finne for the help in acquiring the data and for discussions about design and indicators suitable for the system.

References

[AA13] Andrienko N., Andrienko G.: Visual analytics of movement: An overview of methods, tools and procedures. *Information Visualization 12*, 1 (2013), 3–24. doi:10.1177/1473871612457601. 2

[AAD*10] Andrienko G., Andrienko N., Demšar U., Dransch D., Dykes J., Fabrikant S. I., Jern M., Kraak M.-J., Schumann H., Tominski C.: Space, time and visual analytics. *International Journal of Geographical Information Science 24*, 10 (2010), 1577–1600. doi:10.1080/13658816.2010.508043. 2, 6

[AAF16] Andrienko G., Andrienko N., Fuchs G.: Understanding movement data quality. *Journal of Location Based Services 10*, 1 (2016), 31–46. doi:10.1080/17489725.2016.1169322. 3

[AAJ*07] Andrienko G., Andrienko N., Jankowski P., Keim D. A., Kraak M.-J., MacEachren A. M., Wrobel S.: Geovisual analytics for spatial decision support: Setting the research agenda. *International Journal of Geographical Information Science 21*, 8 (2007), 839–857. doi:10.1080/13658810701349011. 2

[BKC*13] Borgo R., Kehrer J., Chung D. H. S., Laramee B. S., Hauser H., Ward M., Chen M.: Glyph-based visualization: Foundations, design guidelines, techniques and applications. In *Proceedings of Eurographics – State of the Art Reports* (2013). doi:10.2312/conf/EG2013/stars/039-063. 7

[BM13] Brehmer M., Munzner T.: A multi-level typology of abstract visualization tasks. *IEEE Transactions on Visualization and Computer Graphics 19*, 12 (2013), 2376–2385. doi:10.1109/TVCG.2013.124. 4

[BMMS91] Buja A., McDonald J., Michalak J., Stuetzle W.: Interactive data visualization using focusing and linking. In *Proceedings of IEEE Visualization* (1991), pp. 156–163. doi:10.1109/VISUAL.1991.175794. 7

[BNTW18] Bonham C., Noyvirt A., Tsalamanis I., Williams S.: *Analysing port and shipping operations using big data*. Tech. rep., 2018. URL: https://datasciencecampus.ons.gov.uk/wp-content/uploads/sites/10/2018/06/here.pdf. 2

[Boo00] Booth A. J.: Incorporating the spatial component of fisheries data into stock assessment models. *ICES Journal of Marine Science 57*, 4 (2000), 858–865. doi:10.1006/jmsc.2000.0816. 2

[Bro96] Brooke J.: SUS – a quick and dirty usability scale. In *Usability evaluation in industry*, Jordan P. W., Thomas B., McClelland I. L., Weerdmeester B., (Eds.). CRC Press, 1996. 12

[CNYM00] Chung L., Nixon B. A., Yu E., Mylopoulos J.: *Non-Functional Requirements in Software Engineering*. Springer, 2000. doi:10.1007/978-1-4615-5269-7. 5

[CYT14] Chang S.-K., Yuan T.-L., Trenkel V.: Deriving high-resolution spatiotemporal fishing effort of large-scale longline fishery from vessel monitoring system (VMS) data and validated by observer data. *Canadian Journal of Fisheries and Aquatic Sciences 71*, 9 (2014), 1363–1370. doi:10.1139/cjfas-2013-0552. 2

[DBA17] Ducharme-Barth N. D., Ahrens R. N. M.: Classification and analysis of VMS data in vertical line fisheries: Incorporating uncertainty into spatial distributions. *Canadian Journal of Fisheries and Aquatic Sciences 74*, 11 (2017), 1–51. doi:10.1139/cjfas-2016-0181. 2

[DBSA18] Ducharme-Barth N. D., Shertzer K. W., Ahrens R. N. M.: Indices of abundance in the Gulf of Mexico reef fish complex: A comparative approach using spatial data from vessel monitoring systems. *Fisheries Research 198* (2018), 1–13. doi:10.1016/J.FISHRES.2017.10.020. 2

[EHD12] Enguehard R. A., Hoeber O., Devillers R.: Interactive exploration of movement data: A case study of geovisual analytics for fishing vessel analysis. *Information Visualization 12*, 1 (2012), 65–84. doi:10.1177/1473871612456121. 2, 4

[Els] Elsevier: Most cited fisheries research articles. URL: https://www.journals.elsevier.com/fisheries-research/most-cited-articles. 2

[EY15] Elmqvist N., Yi J. S.: Patterns for visualization evaluation. *Information Visualization 14*, 3 (2015), 250–269. doi:10.1177/1473871613513228. 11

[Fis] Fiskeridirektoratet: Den aktive fiskeflåten. URL: https://fiskeridir.no/Yrkesfiske/Statistikk-yrkesfiske/Fiskere-fartoey-og-tillatelser/Den-aktive-fiskeflaaten. 2

[Fis17] Fiskeridirektoratet: J-208-2017: Forskrift om posisjonsrapportering og elektronisk rapportering for norske fiske- og fangstfartøy, 2017. URL: https://fiskeridirektoratet.no/Yrkesfiske/Regelverk-og-reguleringer/J-meldinger/Gjeldende-J-meldinger/J-208-2017. 3

[Fis18a] Fiskeridirektoratet: Forslag til plan for utvidet datainnsamling fra fiskeflåten, 2018. URL: https://www.fiskeridir.no/content/download/23171/323126/version/14/file/elektronisk-rapportering-fartoy-under-15-meter-oktober2018.pdf. 2

[Fis18b] Fiskeridirektoratet: J-244-2018: Forskrift om regulering av fisket etter blåkveite nord for 62°N i 2019, 2018. URL: https://www.fiskeridir.no/Yrkesfiske/Regelverk-og-reguleringer/J-meldinger/Gjeldende-J-meldinger/J-244-2018. 4

[HZY*16] Huang X., Zhao Y., Yang J., Zhang C., Ma C., Ye X.: TrajGraph: A graph-based visual analytics approach to studying urban network centralities using taxi trajectory data. *IEEE Transactions on Visualization and Computer Graphics 22*, 1 (2016), 160–169. doi:10.1109/TVCG.2015.2467771. 2

[JSS*15] Janetzko H., Sacha D., Stein M., Schreck T., Keim D. A., Deussen O.: Feature-driven visual analytics of soccer data. In *Proceedings of the IEEE Conference on Visual Analytics Science and Technology* (2015), pp. 13–22. doi:10.1109/VAST.2014.7042477. 2

[LHK06] Li X., Han J., Kim S.: Motion-Alert: Automatic anomaly detection in massive moving objects. In *Proceedings of the International Conference on Intelligence and Security Informatics* (2006), pp. 166–177. doi:10.1007/11760146_15. 3

[McC04] McConnell S.: *Code Complete*, 2 ed. Microsoft Press, 2004. 4

[Mun14] Munzner T.: *Visualization Analysis and Design*. CRC Press, 2014. 7

[RHR16] Roberts J. C., Headleand C., Ritsos P. D.: Sketching designs using the Five Design-Sheet methodology. *IEEE Transactions on Visualization and Computer Graphics 22*, 1 (2016), 419 – 428. doi:10.1109/TVCG.2015.2467271. 11

[Sch15] Scheepens R.: *Visualization for Maritime Situational Awareness*. PhD thesis, Eindhoven University of Technology, 2015. 2

[SF19] Storm-Furru S.: *Visual Analytics for Fishing Vessel Operations*. Master's thesis, University of Bergen, 2019. 12

[Sha16] Shaffer J.: 5 tips on designing text tables, 2016. URL: https://www.tableau.com/about/blog/2016/4/if-you-must-use-text-table-follow-these-5-design-tips-53766. 9

[SHvv16] Scheepens R., Hurter C., van de Wetering H., van Wijk J. J.: Visualization, selection, and analysis of traffic flows. *IEEE Transactions on Visualization and Computer Graphics 22*, 1 (2016), 379–388. doi:10.1109/TVCG.2015.2467112. 2

[SIBB11] Sedlmair M., Isenberg P., Baur D., Butz A.: Information visualization evaluation in large companies: Challenges, experiences and recommendations. *Information Visualization 10*, 3 (2011), 248–266. doi:10.1177/1473871611413099. 12

[Smi01] SMITH E. A.: The role of tacit and explicit knowledge in the workplace. *Journal of Knowledge Management 5*, 4 (2001), 311–321. doi:10.1108/13673270110411733. 12

[Svv14] SCHEEPENS R., VAN DE WETERING H., VAN WIJK J. J.: Non-overlapping aggregated multivariate glyphs for moving objects. In *Proceedings of the IEEE Pacific Visualization Symposium* (2014), pp. 17–24. doi:10.1109/PacificVis.2014.13. 2

[War13] WARE C.: *Information Visualization: Perception for Design*, 3rd ed. Morgan Kaufmann, 2013. 6

[WMY*17] WANG G., MALIK A., YAU C., SURAKITBANHARN C., EBERT D. S.: TraSeer: A visual analytics tool for vessel movements in the coastal areas. In *Proceedings of the IEEE International Symposium on Technologies for Homeland Security* (2017), pp. 1–6. doi:10.1109/THS.2017.7943473. 3, 12

[WWL07] WOBBROCK J. O., WILSON A. D., LI Y.: Gestures without libraries, toolkits or training: a $1 recognizer for user interface prototypes. In *Proceedings of the ACM Symposium on User Interface Software and Technology* (2007), pp. 159–168. doi:10.1145/1294211.1294238. 4

[XD14] XAVIER G., DODGE S.: An exploratory visualization tool for mapping the relationships between animal movement and the environment. In *Proceedings of the ACM SIGSPATIAL Workshop on Interacting with Maps* (2014), pp. 36–42. doi:10.1145/2677068.2677071. 2