Visual Analysis of Multivariate Movement Data using Interactive Difference Views

Ove Daae Lampe[1,2], Johannes Kehrer[2], and Helwig Hauser[2]

[1]Chr. Michelsen Research (CMR), Norway, http://cmr.no
[2]Department of Informatics, University of Bergen, Norway, http://www.ii.UiB.no/vis

Abstract
Movement data consisting of a large number of spatio-temporal agent trajectories is challenging to visualize, especially when all trajectories are attributed with multiple variates. In this paper, we demonstrate the visual exploration of such movement data through the concept of interactive difference views. By reconfiguring the difference views in a fast and flexible way, we enable temporal trend discovery. We are able to analyze large amounts of such movement data through the use of a frequency-based visualization based on kernel density estimates (KDE), where it is also possible to quantify differences in terms of the units of the visualized data. Using the proposed techniques, we show how the user can produce quantifiable movement differences and compare different categorical attributes (such as weekdays, ship-type, or the general wind direction), or a range of a quantitative attribute (such as how two hours' traffic compares to the average). We present results from the exploration of vessel movement data from the Norwegian Coastal Administration, collected by the Automatic Identification System (AIS) coastal tracking. There are many interacting patterns in such movement data, both temporal and other more intricate, such as weather conditions, wave heights, or sunlight. In this work we study these movement patterns, answering specific questions posed by Norwegian Coastal Administration on potential shipping lane optimizations.

1. Introduction

Massive streams of complex time-dependent data arise in various areas of business, science, and engineering (resulting from large-scale measurements, modeling, or the simulation of dynamic processes). Being able to understand time-related developments allows one "to learn from the past to predict, plan, and build the future" [AMM*07]. This can play a major role in scenarios such as the analysis of critical process workflows and developments, project planning or process simulation, and to develop alternative scenarios if required. In our case, the Norwegian Coastal Administration (NCA) has been asked by the Norwegian government to perform an analysis of whether a sea tunnel should be made on Stad. Most of the Norwegian coastline allows vessels to safely travel inshore, protected from the harsh weather of the North Sea by a large number of bigger and smaller islands (see Figure 1). At Stad, however, vessel traffic is forced out in the open sea. This presents a problem since there are demanding wave conditions 90 to 110 days a year in this area.

The tunnel in question would traverse underneath the peninsula below Stad near Selje, and be 1.8 km long, 23 meters wide, 45 m high including 12 m water depth, producing an excavated mass equal to 3/4 of the Giza pyramid. Building this tunnel would amount to a large national endeavor due to its size, so a careful economic rationale is needed in the first place. Part of the rationale would consist of a decreased risk for the vessel traffic in this area. Another one would be saved costs by having vessels not needing to wait for good weather north and south of Stad. The questions of interest for NCA therefore were, how significant the correlation of waiting periods and bad weather is, and whether we can quantify the amount of (lost) hours that could be saved by having the tunnel as a weather-safe short-cut. Another line of questions addresses the potential risk reduction when having such a tunnel. Accordingly, we were interested in how the weather affects the vessels' choice of paths, e.g., do they go closer to shore, or not, when the weather is bad.

Faced with these questions, we engaged with an analysis of the Automatic Identification System (AIS) data for a large historical collection of vessel movements. AIS is a radio system, broadcasting vessel ID and position at regular intervals,
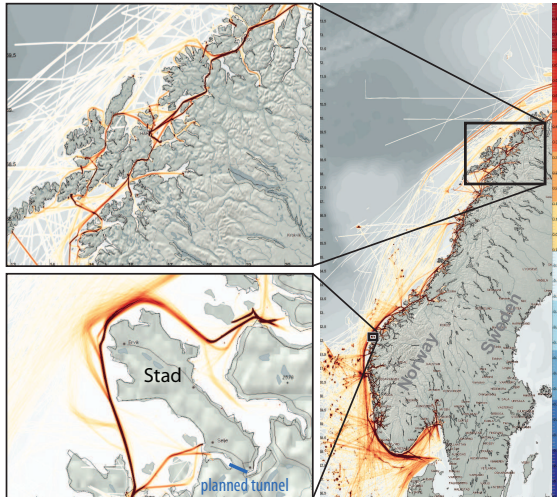
Figure 1: *Vessel movements around the coast of Norway. At Stad (lower inset) the traffic is forced into the open sea, while usually most (local) traffic is within the outer islands protected from the weather (exemplified by the top inset).*

that all vessels above a certain size must have in these waters. Coupled with historical weather observations from stations in the vicinity of Stad, we should have the required data to answer these questions. According to Andrienko and Andrienko [AA10], AIS data contains three attributes characterizing agent movement data, i.e., agent identifier, time, and spatial position. Selecting all data by identifier, and ordering it by time, makes a trajectory, and many of these trajectories then makes a movement dataset. Furthermore, AIS contains a varying number of attributes per vessel, e.g., vessel length, vessel type, and nation, and attributes per journey such as persons on board, destination, and cargo. When we further extend this dataset by spatiotemporal attributes, such as wind direction, wind speed, and wave height, we indeed have a multivariate movement data visualization challenge at hand. Another consideration that we have to take into account is that the bigger the datasets get, the better the statistical confidence of our findings can get. This means that if we have sufficiently many trajectories, we can consider the data as a probability density estimation, as opposed to a set of just a few samples. AIS is collected by a huge network of radio transponders along the coast; on the other end, it is also dependent on the transponder on the individual vessels. Because of this complexity the raw data is prone to several errors. Usually AIS data is filtered to remove erroneous paths or ID conflicts. This paper utilizes the raw data, and the paths that cross land is interactively removed by filtering all the line-segments longer than a given tolerance distance.

In the current workflow of scientists and practitioners, the analysis of trajectory data is done by reducing the size of the problem, and/or aggregating it to a single pass-line, which then is statistically analyzed in greater detail. Such an approach gives a very good quantitative result, a single yes or no with respect to the considered hypothesis. This procedure, however, does not allow for a more flexible exploration of the data, aiding the forming of perhaps new and unexpected hypotheses that then are further analyzed.

In this paper, we demonstrate how a flexible visual analysis is utilized in this challenging application. We contribute a novel way of performing differential analysis of trajectory data and a new work-flow of iterating through these difference views. The presented solution was designed to quickly iterate through a sequence of difference views that are utilized to compare different categorical and quantitative attributes (such as different timespans, vessel-types, wind speeds), and to analyze a set of hypotheses as they emerge during the visual analysis. By using a visualization based on kernel density estimation to visualize the movement data, and difference views representing quantitative differences between the categories, the user can drill-down into the information. Possible correlations between waiting periods of vessels and bad weather conditions can be investigated, moreover, if the vessels' choice of a route is affected by the weather. Analyzing these and other questions supports the decision makers when evaluating whether or not to build the tunnel. In the next sections we first discuss related work, then we describe our application and the techniques employed here, then we analyze the domain questions, before we sum up, and provide a conclusion.

2. Related Work

A large number of publications deals with the visualization and analysis of time-dependent and multi-variate data (see Aigner et al. [AMM*07] and Fuchs and Hauser [FH09] for comprehensive surveys). Common analysis approaches for movement data include the visualization of raw data, computed summaries, or extracted patterns [AAD*08]. Spatial and/or temporal aggregation is often used in order to reduce the data complexity or visual cluttering. With such an approach, data items sharing the same spatiotemporal domain are summarized and depicted instead of the individual data values. According to Andrienko and Andrienko [AA06], data aggregation can be done either by calculating data characteristics (e.g., the sum, arithmetic mean, variance) or by grouping techniques such as clustering or binning.

BinX [BM04] visualizes long time series by binning along the time axis at different levels of aggregation and then displays mean, minimum, maximum value, and standard deviation per bin. Hao et al. [HKDS07] use pixel-based techniques to visualize time-dependent data at multiple resolutions based on importance values per data interval. Andrienko and Andrienko [AA10] visualize movement data as flow maps where the spatial domain is subdivided into appropriate areas (based on significant points in the movement) and aggregated trajectories with common start and

end points are visualized as arrows. Janoos et al. [JSI*07] analyze pedestrian movement data using a wavelet-based feature descriptor in order to detect anomalies. Grundy et al. [GJL*09] propose spherical scatterplots and histograms as an alternative representation of movement data.

A thorough overview on the usage of kernel and other density estimates in visualization is given by Scott [Sco92]. Fisher visualizes the usage frequency of map tiles in his Hotmap [Fis07], which is similar to a density estimate. Willems et al. [WvdWvW09] propose a visualization approach based on the convolution of dynamic movement data with a kernel, where the resulting density field is visualized as an illuminated height map. A combination of overview and details is provided by combining two fields, one computed with a small and one with a large kernel. While their approach provides very good results for presentation, it takes approximately 10 minutes to compute the data from one day (100.000 line segments). It is thus less suitable for a visual analysis where interaction is a key issue. Our approach, on the other hand, performs in real-time for even larger amounts of data using a GPU-based implementation. It is integrated in a framework of multiple views (with linking and brushing) and supports algebraic operations such as computing differences. Our approach also provides quantitative visualization where the value of a single pixel/cell shares the same unit as the depicted data.

Several applications support the visual analysis of temporal trends and patterns using interactive brushing or querying techniques. Interesting data subsets are interactively selected (brushed) directly on the screen, the relations are investigated in other linked views (compare to the Xmdv-Tool [War94]). Feature visualization and specification via brushing in multiple views (including histograms, scatterplots, and 3D views) is an integral part of the SimVis framework [DGH03]. Jern and Franzén [JF06] propose a coordinated multiple views system for exploring spatio-temporal multivariate data. Hurter et al. [HTC09] extract complex features in aircraft trajectories by brushing in juxtaposed views. The brushed trajectories are spread across views with a pick and drop operation. The views can be rapidly configured by connecting data attributes to visual variables such as color or size. Kehrer et al. [KFH10] recently demonstrate how the iterative reconfiguration of depicted view attributes enable a powerful analysis process. In our system, we explicitly represent such transformations of views that support the visual analysis of a set of hypotheses that emerge during the visual analysis (e.g., comparing traffic at different workdays).

According to Verma and Pang [VP04], different data sets can be compared at the image level, the data level, or the feature level. Image-level comparisons include side-by-side visualization and the visualization of differences between the images per pixel. For such approaches, also the selection of an appropriate color map is very important (e.g., using a diverging map to visualize differ-

ences [Bre99]). Polaris/Tableau [STH02] supports the visual exploration of hierarchically organized multi-variate data using a table-based layout of views (commonly called small multiples [Tuf83]). For side-by-side comparison, user-specified categories/hierarchies are opposed such as time (year, quarter, month), products, or spatial locations (town, state, country). Data attributes can, moreover, be interactively transformed (e.g., by aggregation or grouping), filtered, and/or brushed. Woodering and Shen [WS06] propose volume shaders to compare and combine multiple time-dependent volumes by consecutive algebraic set operators and numerical operators. For interaction and visualization of the resulting volume tree they utilize image spreadsheets (compare also to Jankun-Kelly and Ma [JKM01]).

3. Interactive Difference Views

The interactive visual analysis and exploration of the movement data is carried out in a setup of coordinated multiple views with linking and brushing (see Figure 2). The views include histograms, scatterplots, and frequency-based views based on Kernel Density Estimates [Sil86] (KDE). The latter views are computed by convolving the movement data with a kernel (usually a Gaussian) for each sample, resulting in a density estimate that can also be extended to cope with trajectories (using a line kernel instead of a point spread function). The questions of our application partners were answered by developing an iterative workflow for creating quantitative difference views, with the aim of facilitating the fast and flexible investigation of large amounts of movement data. A difference view results from subtracting one KDE plot from another one, which then shows the quantitative difference between them. While animations and side-by-side views often provide good means to answer qualitative questions (e.g., "where" and "when"), they are less suitable for answering quantitative questions (e.g., "how much/many"). Such quantitative differences are explicitly represented in our difference views, for instance, using a diverging color-map [Bre99] (see Figure 2 B).

In the following, we describe our interactive and iterative analysis, our quantitative difference visualizations, and how we can handle large datasets at interactive frame-rates.

3.1. Interactive and Iterative Visual Analysis

From the domain questions, we derived a couple of requirements for our solution. As often in the context of hypothesis testing and analysis, every new finding leads to new questions as well. Accordingly, we shaped our application in an as iterative and interactive fashion as possible. This iterative workflow enables the user to search for one answer, and then further investigate unexpected trends, or to search for multiple indicators forming a single answer.
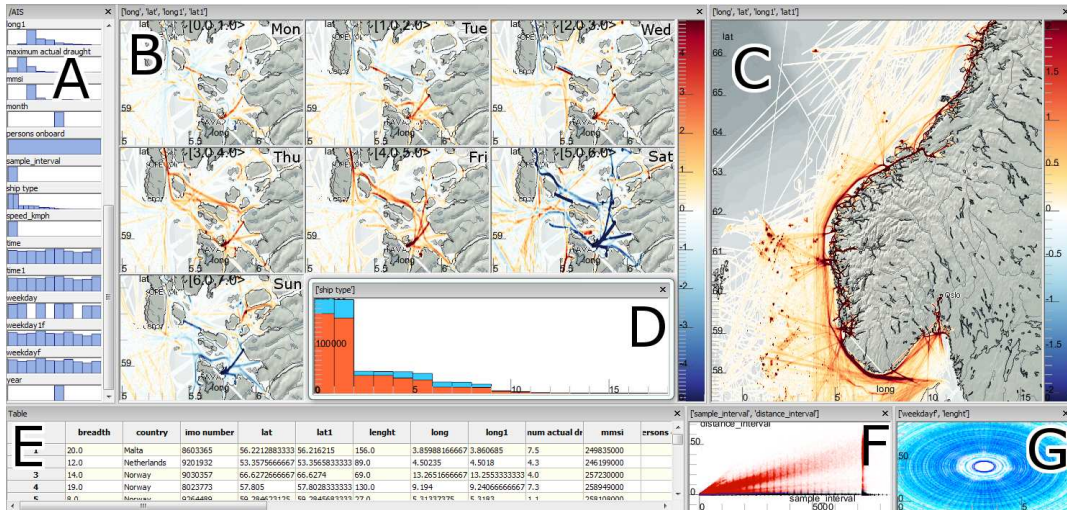
When the user first loads a multivariate dataset, an

Figure 2: *Overview of the described application. A shows the available attributes from the dataset, displayed as histograms. The 7 multiples in B show a close-up to Stavanger, split by weekdays, showing the differences, in traffic volume, from the average. C is an overview and D is a histogram of different ship types. E is a table view of all samples, F a scatterplot, and G a radial plot displaying vessel activity during the week.*

overview of the attributes (variates) is automatically displayed in the dataset window (see Figure 2 A). Every attribute is represented with its own small histogram. These histograms acts as drag-sources, in a drag and drop sense. To construct a visualization, the user drags an attribute onto an empty view. While still dragging, a context frame appears over the current view, with multiple possible drop targets. Each of these drop targets represents a possible binding between the dragged attribute and a property of the current visualization, e.g., a spatial binding to either the x or the y axis of the view, or binding to size or color. In this manner the user quickly creates one or several compound views.

The next step is to relate these views by brushing, and specify features across multiple variates by constructing a set of rules. As an example, the user brushes all northbound vessels with a speed of 5 knots or more in one view, and selects the category of ship-type equal to Tankers in another view. This ruleset is then reflected on all views, using a focus+context style in sample-based views (e.g., scatterplots), and filtering in the KDE plots. As another step on top of these relate-and-filter techniques, we have added a *compare over* expansion. When the user drags an attribute to a frequency-based view, he or she has the option of dropping this to the compare-over option (available from a context menu). This expansion splits the current view into one difference view for each of the categories (or bins if the attribute is continuous). Each of these difference views then displays all the samples matching the given category subtracted by the average. In areas where this category is greater than the average, the result will have a positive sign (red in our figures),

and negative (blue) in areas where there is less than average. Further below, we discuss in more detail what we mean by computing the average with respect to categorical attributes such as weekdays, ship-type, or wind direction. These difference views build on top of the existing ruleset from the previous step, and thus form a two-level rule hierarchy. So as in the above mentioned example, when expanding vessel traffic over weekdays in a map plot, this would show how the northbound tankers with a speed of 5 knots or more, on one weekday, would compare to the average weekday of northbound tankers.

After creating several difference views, the user can select one particular difference view. This view then replaces the previous reference view, and its second-level rule on a category will be added to the level-one ruleset. In Figure 3 we describe this iterative creation of difference views, where categories are selected through a series of difference views. Returning to our previous example, where we had seven difference views, one for each weekday (category), we can select one day, e.g., Sunday, and then all views only show northbound tankers, with 5 knots or more on Sundays. This cycle can be repeated, enabling a deep drill-down into the data.

3.2. Quantitative Difference Visualizations

The concept of difference views, and their ability to display quantitative differences between two comparable views, has been utilized in several other works, yet there is no clear workflow for the flexible configuration of exactly what to create difference views between. To facilitate the creation of
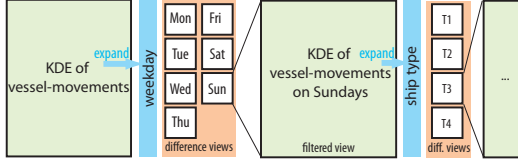
Figure 3: *Iterative data exploration via difference views.*
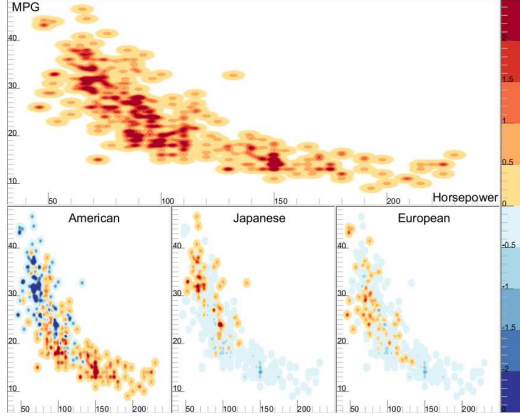

Figure 4: *Miles per gallon (MPG) over horsepower for 406 cars [RD] shows an inverse correlation in the top view. The top view can be expanded in the three bottom views, where American, Japanese, and European cars are compared to the average. We can see that American cars have many more cars with high horsepower. Compared to European cars, they have more horsepower for an equally rated MPG.*

meaningful difference views, we defined the *compare over* functionality, which splits up a current view into several, one for each category. For example, the top view of Figure 4 shows a frequency view of horsepower vs. miles per gallon of the 406 cars in the Car dataset [RD]. After this view has been customized or optionally filtered, the user then drags the origin column (denoting which continent the cars are produced in) into this view, and then drops it on the expand icon in the in-screen menu that pops up. The whole of Figure 4 is the automatic result of this operation. Since the column dragged is a categorical attribute the user is presented with one additional view per category. These views present how samples in this category compare to the average over all categories. The average in this case is achieved by dividing by the number of categories. The compared difference views show the sum of each sample's kernel, where those samples from the current category are given a positive sign, all others a negative sign, and all scaled for averaging. A 2D KDE [Sco92] is defined by

$$\widehat{f}_{\mathbf{H}}(\mathbf{x}) = \frac{1}{n}\sum_{i=1}^{n} K_{\mathbf{H}}(\mathbf{x} - \mathbf{x}_i)$$

with $\mathbf{H}^{-\frac{1}{2}}$ being a symmetric and positive definite bandwidth matrix and $K_{\mathbf{H}}$ being defined as

$$K_{\mathbf{H}}(\mathbf{x}) = |\mathbf{H}|^{-\frac{1}{2}} \mathcal{K}(\mathbf{H}^{-\frac{1}{2}}\mathbf{x}).$$

$\mathcal{K}$ is a multi-variate kernel function that integrates to 1. By defining result of two KDEs, $f(\mathbf{x})$ as the average view, which includes all samples, and $g(\mathbf{x})$ for the subset of only those samples within the given category, we can define our difference view as:

$$d(\mathbf{x}) = g(\mathbf{x}) - f(\mathbf{x}) \tag{1}$$

Instead of first creating the full KDE $f(\mathbf{x})$, and then subtract a subset of samples from that KDE, we can do this in a single step. Since the set of those samples matching the category is a subset of all samples, we can simplify Eq. 1 to a single summation pass over the samples, and scale the samples in the category by $\frac{1-n}{n}$, and the rest by $\frac{1}{n}$.

As an another example, considering temporal ranges, we look at how traffic differs on the different days of the month. To establish the KDE representing an average day, we calculate the temporal range of all the samples, i.e., the number of days our sampleset spans, and divide by this range. Next the temporal range of the samples in the current category needs to be calculated, which is not quite as trivial as the example above. If we have a set of samples spanning over several months, and would like to compare weekdays against weekends. Since there are more days belonging to the weekday category than that of the weekend category, we cannot normalize by the total number of days. Instead, we need to count the number of days matching "weekday" that actually contributes to this subset.

In our solution we have implemented an automatic technique that iterates over the samples and can calculate the sum of smaller temporal ranges that match the current category, e.g., days, hours with particularly strong wind, or weekdays. When this is established, the temporal difference view can also be calculated in a single step using the above equation.

3.3. Large Datasets

A requirement on the application was to enable the analysis of statistical significance of the results. Significance, here, is determined by the amount of noise, the signal and sample size. Since we do not have any influence on the contained noise and the signal size (after data acquisition), we attempt to optimize the quantitative significance of our analysis through the third factor, i.e., the sample size. Allowing for larger datasets to be interactively analyzed, helps to increase the confidence in the extracted findings. If we had chosen to support a sample size just large enough for the task at hand, this would not allow for any flexibility with respect to further drill-down steps or alternative comparisons.

In visualization we often deal with three levels of limitations on dataset size, (1) when the dataset fits into graphics memory, (2) when it is too large for graphics memory, but

still fits in main memory; and (3), when it is too large to fit in main memory, but reside on a file level. Our implementation supports the third category, but in order to keep interactivity, employs a three level data handling scheme. If a file is larger than what the application can hold in main memory, only a subset of the file is loaded. A yet smaller subset is then kept on the graphics card, and displayed at all times in the application. The size of this smallest subset is selected such that interactive speeds can give quick response when brushing, even though there are many views. Immediately when interaction ceases, the application starts rendering, in batches, to the now stationary views from the rest of data in main memory. And again, when interaction starts again, all views fall back to only display the GPU-resident data in the first place. This interaction is shown in the supplementary material. These two top levels gives quick and interactive access to what should be a representative sub-sampled portion of the data. Due to the nature of interactive visual analysis, including filtering and refining, we do not stop there. We allow the application to keep a second file of query results, that allows the user to apply the visual brushes to the file level. This new query results file can then be used for further analysis, and then perhaps have all its data fit in main memory, and thus have it all shown in the visualization views.

4. Answering the Application Questions

The main question that the government wants answered is whether or not to build a tunnel through Stad, and such an answer should include reasons as to why, backed up by quantitative indicators. In this analysis, the domain expert investigates potential decision criteria, and then investigates whether those are significant or not. In this work, the domain expert contacted us with an interest in visualizing the AIS data, and to look at two such indicators. First, it was interesting to look at how many vessels are actually waiting when there is bad weather, and second, whether vessels go closer to the shore when the wind picks up, which increases the risk of accidents. To compare our AIS data with weather data, we obtained meteorological measurements from two stations, Kråkenes fyr, a prominent light house south of Stad, and Svinøy fyr, another light house on a small island north of Stad. These measurements contain wind direction and wind speed, which we then applied to all samples based on their spatio-temporal proximity.

To compare how wind speed affects the amount of stationary vessels, we brush vessels with speed close to zero. We then zoom into the area around Stad on the map view, which now shows areas where vessels are stationary. To compare how this view changes with respect to different wind speeds, we take the wind speed attribute and perform a compare over action. The result is shown in Figure 5. The top-left category (weak winds) shows a greater than average amount of stationary vessels. The bottom-right view with the strongest winds (strong gale and worse) shows a significant drop (7%).
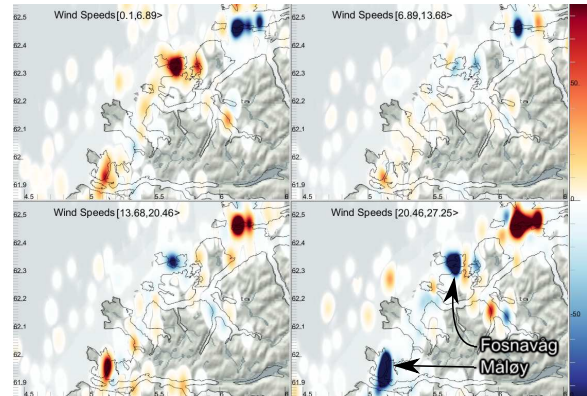

Figure 5: *Close to Stad, brushed to only include stationary vessels. The views show how many (compared to all) vessels are stationary, given different wind speeds.*

The top right view is 3.2% below average and the bottom left 5% above average. Accordingly, there is no trend that either confirms nor reject our hypothesis yet. The explanation for the increase in stationary vessels when the weather is good, is that there is an overall more vessels out at sea when the opportunity calls for it, and the opposite for the strongest winds. To show this, we then include all vessels, stationary or not, but keep our wind categories, and calculate the integrals. This reveals that there is a drop in overall traffic of 20% in the strongest wind category, and an increase in traffic in the lowest category. An overall drop in 20% traffic, but only a drop in 5% stationary vessels, indicates that our hypothesis holds, and that there is indeed an increasing amount of stationary vessels when the wind is bad.

Another approach is to compare the actual traffic past Stad, and to compare how this volume is changing with respect to different weather conditions. If our earlier assumption that vessels need to wait when there is bad weather is true, we should see a decrease in traffic. Figure 6 shows the traffic around Stad expanded into four categories of wind speed. By computing the integral for a selection, we can see that the first category (no/little wind), shows 8.6% more traffic than the average, the next 3.0% more than the average, the third 5.6% less and the last, with winds from strong gale and up, show a significant decrease by 24%. A finding that further strengthens our inital hypothesis, however, we can even investigate further.

The third approach is a more item-based one; we can read from the weather data that 21st of November 2008 had particularly bad weather, and that the following day the weather calmed to a breeze. By counting the number of vessels that passed Stad on these two days, we can see if together they stay within the average passes, and how many have been delayed by one day. By brushing an area around Stad and one single day, the table view (see Figure 2 E), will display that
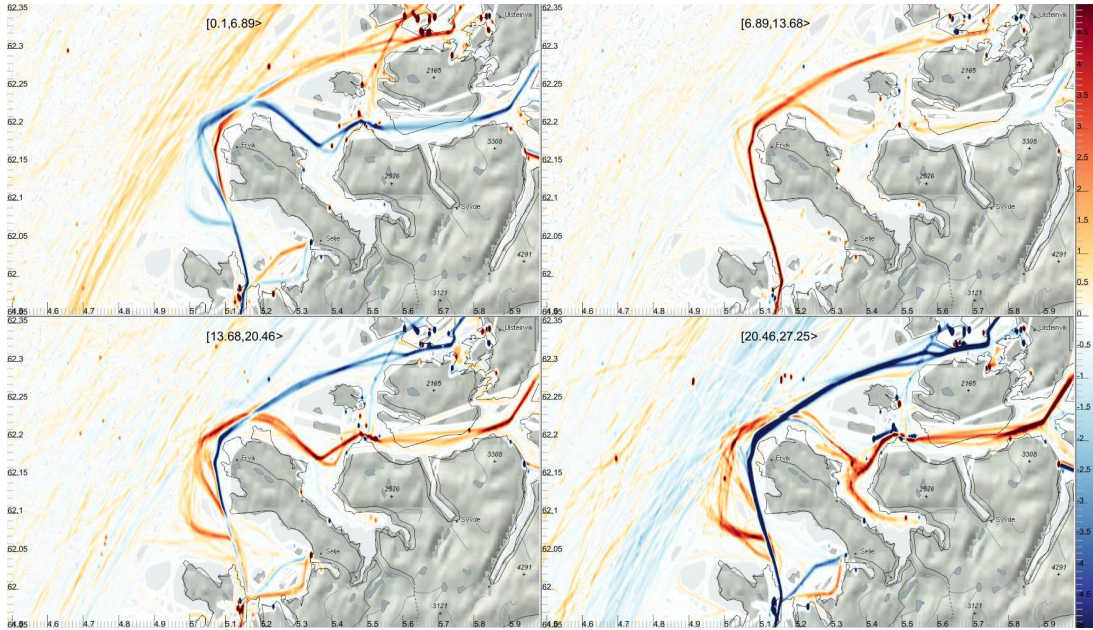
Figure 6: *Passing vessels outside the Stad peninsula, and their changed movement pattern given stronger winds. Red colors indicate more than average traffic in that interval of winds, and blue colors indicate less than average.*
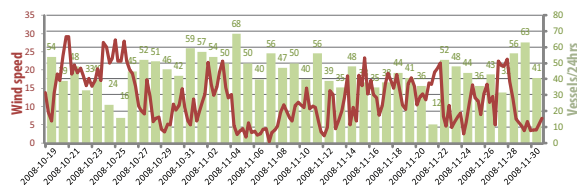

Figure 7: *Wind speed in m/s and passes by Stad, peaks in wind speed forces vessels to wait, and then when the weather gets better, there is an increase of vessels passages.*

the selected area has an average of 38 unique vessel IDs registered per day, on the 21st there were 12 and on the next day 47. Figure 7 show this as well, where the day following the storm had higher traffic. Using these averages, over more than just this case, we could calculate on average how many vessel hours are lost during a season. Comparing to the average in just this one case, however, one can estimate that around ten vessels would have a delay of 24 hours, or, 240 hours lost on a storm lasting less than a day.

The other investigated indicator is whether vessels draws closer to shore when the weather gets bad. In the previous paragraph, we discussed Figure 6 which computed the integral of a selection to see quantitative differences. Answering this question of vessel paths can also be done by studying the same figure. In the first of these four figures, the one with the lowest wind speeds (top-left) there is a red curve

going close to shore, which means that there is a greater than average amount of vessels taking this route when the weather is good. Additionally, in this same figure, the clearly defined blue route further out, defines that there is a much less than average amounts of vessels taking this route. In the next wind category, top-right, this outer route is now "invisible", which means that there is an exact average amount of vessels taking this route. The route close to shore still contains a greater than average amount of vessels. In the third wind category the route close to shore is now clearly defined blue, and those who pass Stad does so selecting the route further away from shore. Similarly with the fourth category, with winds of strong gale or stronger, the route close to shore contains close to zero vessels. Moreover in the two last categories more vessels go straight towards the safety of inshore, where in the two first categories, vessels take the more exposed "shortcut" straight over to Herøy (the island in the top right corner of this map). So in conclusion, Figure 6 clearly shows an opposite effect than the original question, meaning that the stronger the winds, the more distant routes from the shore are selected.

In our discussion with domain scientists, they stated that our application gave them an improved insight into the complexity of their original questions; an insight that later also strengthened their value on AIS as an asset for analysis, which was not fully realized before. Our use of AIS as a probability density estimate, enabled both a non-parametric exploration of the entire dataset, and an in depth analysis of selected details. Furthermore, they found the interac-

tive analysis of AIS as a frequency view "groundbreaking". The application was both flexible and understandable for the users, and showed a great potential for further analysis. Previous analysis required extensive manual labor, and provided statistical analysis for a few chosen pass-lines; this application would alleviate this labor, by providing similar details for every pixel/cell, with a simplified analysis work-flow.

5. Summary and Conclusions

In this paper, we presented an application to investigate particular questions presented by the Norwegian Coastal Administration (NCA). NCA will use conclusive answers to these questions as indicators in their recommendation to the Norwegian government, as to whether or not build a tunnel through Stad. On the first question, concerning the correlation of waiting periods and bad weather conditions, we showed that even with a total reduction, by 24%, in the traffic when there is strong winds; the proportion of the traffic that is stationary vs. the traffic passing Stad is increasing with increasing wind speeds. Another conclusion on this case is found by a sample-based approach, which shows that there is a temporary increase of passings by Stad, after periods of strong winds. On the next question, on whether bad weather affects the vessels to choose a route closer to shore, Figure 6 shows an opposite effect, meaning that more distant routes from the shore are chosen when there are stronger winds. We have demonstrated how this application, using the techniques of iterative creation of difference views and through the use of quantitative visualizations, reached conclusions to the questions posed, and the flexibility to search for several alternative indicators, and thus also meet future demands.

6. Acknowledgments

References

[AA06] ANDRIENKO N., ANDRIENKO G.: *Exploratory Analysis of Spatial and Temporal Data – A Systematic Approach*. Springer, 2006.

[AA10] ANDRIENKO N., ANDRIENKO G.: Spatial generalization and aggregation of massive movement data. *IEEE Trans. Vis. Comput. Graph.* (2010). (RapidPost).

[AAD*08] ANDRIENKO G., ANDRIENKO N., DYKES J., FABRIKANT S., WACHOWICZ M.: Geovisualization of dynamics, movement and change: key issues and developing approaches in visualization research. *Inf. Visualization 7* (2008), 173–180.

[AMM*07] AIGNER W., MIKSCH S., MÜLLER W., SCHUMANN H., TOMINSKI C.: Visualizing time-oriented data: A systematic view. *Computers & Graphics 31*, 3 (2007), 401–409.

[BM04] BERRY L., MUNZNER T.: BinX: Dynamic exploration of time series datasets across aggregation levels. In *Proc. IEEE InfoVis 2004* (2004), pp. 215–216.

[Bre99] BREWER C.: Color use guidelines for data representation. In *Proc. Section on Statistical Graphics* (1999), pp. 55–60.

[DGH03] DOLEISCH H., GASSER M., HAUSER H.: Interactive feature specification for focus+context visualization of complex simulation data. In *Proc. VisSym 2003* (2003), pp. 239–248.

[FH09] FUCHS R., HAUSER H.: Visualization of multi-variate scientific data. *Comput. Graph. Forum 28*, 6 (2009), 1670–1690.

[Fis07] FISHER D.: Hotmap: Looking at geographic attention. *IEEE Trans. Vis. Comput. Graph. 13*, 6 (2007), 1184–1191.

[GJL*09] GRUNDY E., JONES M., LARAMEE R., WILSON R., SHEPARD E.: Visualisation of sensor data from animal movement. *Comput. Graph. Forum 28*, 3 (2009), 815–822.

[HKDS07] HAO M., KEIM D., DAYAL U., SCHRECK T.: Multiresolution techniques for visual exploration of large time-series data. In *EuroVis 2007* (2007), pp. 27–34.

[HTC09] HURTER C., TISSOIRES B., CONVERSY S.: FromDaDy: spreading aircraft trajectories across views to support iterative queries. *IEEE Trans. Vis. Comput. Graph. 15* (2009), 1017–1024.

[JF06] JERN M., FRANZÉN J.: GeoAnalytics—exploring spatio-temporal and multivariate data. In *Proc. Intl. Conf. Information Visualisation (IV '06)* (2006), pp. 25–31.

[JKM01] JANKUN-KELLY T., MA K.-L.: Visualization exploration and encapsulation via a spreadsheet-like interface. *IEEE Trans. Vis. Comput. Graph. 7* (2001), 275–287.

[JSI*07] JANOOS F., SINGH S., IRFANOGLU O., MACHIRAJU R., PARENT R.: Activity analysis using spatio-temporal trajectory volumes in surveillance applications. In *Proc. IEEE VAST* (2007), pp. 3–10.

[KFH10] KEHRER J., FILZMOSER P., HAUSER H.: Brushing moments in interactive visual analysis. *Comput. Graph. Forum 29*, 3 (2010), 813–822.

[RD] RAMOS E., DONOHO D.: 1983 ASA data exposition dataset. lib.stat.cmu.edu/datasets.

[Sco92] SCOTT D. W.: *Multivariate density estimation: theory, practice, and visualization*. Wiley, 1992.

[Sil86] SILVERMAN B.: *Density Estimation for Statistics and Data Analysis*. Chapman & Hall/CRC, 1986.

[STH02] STOLTE C., TANG D., HANRAHAN P.: Polaris: A system for query, analysis, and visualization of multidimensional relational databases. *IEEE Trans. Vis. Comput. Graph. 8*, 1 (2002), 52–65.

[Tuf83] TUFTE E. R.: *The Visual Display of Quantitative Information*. Graphics Press, 1983.

[VP04] VERMA V., PANG A.: Comparative flow visualization. *IEEE Trans. Vis. Comput. Graph. 10*, 6 (2004), 609–624.

[War94] WARD M.: XmdvTool: Integrating multiple methods for visualizing multivariate data. In *Proc. IEEE Visualization* (1994), pp. 326–336.

[WS06] WOODRING J., SHEN H.-W.: Multi-variate, time-varying, and comparative visualization with contextual cues. *IEEE Trans. Vis. Comput. Graph. 12* (2006), 909–916.

[WvdWvW09] WILLEMS N., VAN DE WETERING H., VAN WIJK J.: Visualization of vessel movements. *Comput. Graph. Forum* (2009), 959–966.