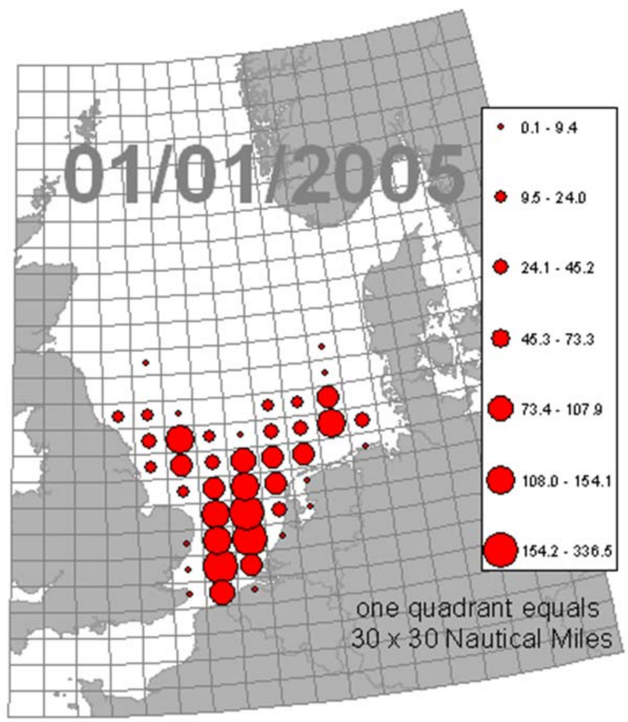
Centre for Geo-Information

Thesis Report GIRS-2009-10

Geo-visualization of North Sea plaice fisheries: a study in usability and stakeholder comprehension

Colin B. Gold



WAGENINGEN UNIVERSITY
WAGENINGEN UR

Geo-visualization of North Sea plaice fisheries: a study in usability and stakeholder comprehension

Colin B. Gold

Registration number 82-10-21-268-080

Supervisors:

Ron van Lammeren
Wim van Densen

A thesis submitted in partial fulfilment of the degree of Master of Science
at Wageningen University and Research Centre,
The Netherlands.

19/05/09
Wageningen, The Netherlands

Thesis code number: GRS-80436
Thesis Report: GIRS-2009-10
Wageningen University and Research Centre
Laboratory of Geo-Information Science and Remote Sensing

Foreword

Over the course of this thesis project I was, at times, challenged and at other times, overcome by the banality of certain functions. In the end, I came to realize the importance of all stages of the research and am overall happy with the result. Indeed, a project which could have gone on for much longer, attempting to incorporate the feedback of many more stakeholders, I do believe that it, in its current state, is a good starting point for any further research on the subject of usability studies of fishery visualizations and animations, in particular.

I benefited from the assistance of many individuals who selflessly performed many different tasks and I wish to thank them wholeheartedly for their efforts. They were: Ron van Lammeren and Wim van Densen, foremost, for supervising me throughout this whole process; Floor Quirijns and and Edwin van Helmond for being so kind as to lend me their datasets so I had something to actually study; Arend Ligtenberg, Philip Wenting, Aldo Bergsma and Willy ten Haaf for providing technical and administrative support at the times when it was most needed; and lastly Marcel Machiels and Paul van Zwieten for also lending invaluable support and knowledge. There were surely others who allowed me to succeed by offering me assistance in other ways, and to them I am grateful.

It was truly an unforgettable time in the Master of Geo-Information program at Wageningen University, and I value all people whom, and experiences which I encountered. Thank you to everyone.

TABLE OF CONTENTS

Tables and Figures

Summary

In the North Sea today, there exists a major problem related to management of plaice resources. Recently, plaice habitat has shifted North, while the fishing industry has largely continued to exploit the areas of the southern North Sea due to fuel prices and legislative restrictions. The result of this interplay of actors is the enormous bycatch of plaice individuals which are under the Minimum Landing Size, since juvenile plaice are still present in the South, while adults are located for most of the year in the central and northern regions.

If migration and life history patterns of plaice can be mutually understood through active and cooperative participation by fishermen, scientists and policymakers, the industry's management and therefore sustainable future, can be better assured. The idea for accomplishing this is to create a GIS animation which has been constructed with input from all stakeholder groups and to compare its usability against traditional static maps. To undertake this, four different steps must be followed: the identification of stakeholder group competencies and needs; an inventory of available data; a review of visualization options for the data; and the testing of the usability of the completed visualization.

The identification of stakeholder competencies was accomplished largely through the review of available literature and already established fishery GIS applications. Data was found to be quite bountiful but the usefulness of it was questioned, as it may not suit the needs of this project. After having consulted other fishery GIS, the data to use was chosen and processed into what seemed most usable in terms of the specifications of the project. This data included Effort, Discards, and Catch. All datasets included appropriate ICES quadrants and dates.

The investigation into available visualization techniques was based on the assumption that users would be at the forefront of our focus, as this is a study in usability. Forms (cartographic variables), purpose, technology and data were all considered, always focusing on simple visualization strategies for maximum understanding. Specifications were made on the cartographic choices such as layout and data representation, as well as the decision to employ ArcGIS as the development software of our animation.

The usability analysis drew upon social and cognitive sciences and the User-Centered Design Process. Cognitive limitations were taken into account and a simple visualization which encouraged focused attention became the first prototype. It was evaluated on four criteria: users' performance, visualization learnability, design flexibility and subjective satisfaction. The usability of the two visualizations (static maps and animation) was tested, each with a questionnaire of objective and subjective questions. The results of the tests were evaluated on time, errors made and completion percentage for objective questions, while the subjective answers were compiled and summarized. Since the stakeholders were so few in number, determining which visualization type was better was difficult.

The Discussion section of the report noted interesting conclusions. Key among them was that both visualization types were effective at answering certain types of questions but not the same. Users expressed a desire for more interactivity in the animation because as it was presented, it was not considered as usable as the static maps, particularly for learnability. Overall, the process of usability testing seemed to be a success, if understood that this is only the first iteration of the process. The flexibility of the design seemed appropriate for the limited number of participants but this must be increased in order to make firm conclusions. Another must for the visualizations is the improvement of the fishery data which was used, as the first choice dataset was not complete enough to yield a usable visualization.

1. <u>Introduction</u>

1.1 Research Setting

The North Sea is a large body of water located in Northern Europe. With an ocean current from the North Atlantic in the Northwest, the North Sea is on the frontier of the boreal and Lusitanian ecosystems. As such, it is home to some 230 species of marine life (ICES 2007a). Currently, 31 out of 113 commercially harvested species are listed as overexploited by ICES, the International Council for the Exploration of the Seas (Daan and van der Mheen 2004).

These species are concentrated on the continental shelf in the North and sandy shoals in the South, with relatively low species richness in the deep waters of the center (ICES 2007b). Consequently, fisheries are nowadays locating their activities particularly in the accessible southern regions close to their ports of call (ICES 2006). This area is becoming overfished, as 99% of fish are caught within the 200-mile Exclusive Economic Zones, which is only one percent of all water volume of the North Sea (Valavanis 2002).

Fishermen are not the only ones to blame though. The fragmented nature of the jurisdiction of fishery activity worldwide has also contributed to the unsustainability of the situation. The lack of regional cooperation in management of these resources has forced the United Nations and European Union to enact regulations to combat this "tragedy of the commons"(Daan and van der Mheen 2004).

For European Plaice, *Pleuronectes platessa*, the species with which this report is most concerned, various restrictions have been placed on the fishing industry. Among these, include: a Total Allowable Catch, or TAC; minimum size of landable fish; areas where fishing is limited or forbidden for certain categories of boats; technical/mechanical limitations; and seasonal restrictions (Daan and van der Mheen 2004)

Governing bodies have done more than just put a band-aid on the problem of over fishing though. The European Union has taken note of the potential usefulness of Geographic Information Systems in contributing to the solution, by making inventories, analyzing, and monitoring fishery industries (Valavanis 2002). Indeed, problems such as these can more easily be addressed if some kind of pattern can be recognized (Visvalingam 1994). If migration and life history patterns of plaice can be mutually understood through active and cooperative participation by fishermen, scientists and policymakers, the industry's management and therefore sustainable future, can be better assured.

1.2 Human – Species Interaction

The one key element to better managing plaice is the migration and life history patterns of this demersal flatfish. Adult plaice spawn in early spring in the eastern English Channel, the southeast North Sea, and the Kattegat, between Denmark and Sweden (Wennhage et al. 2007). There are several criteria that contribute to the spawning location: water temperature, water salinity level, water depth, and substratum. Typically, plaice prefer warmer waters, lower salinity, a depth of less than 30 meters, and a sandy or fine-grain, soft-bottom substratum (Eastwood and Meaden 2000). Ocean currents are capable of moving the larvae quite some distance (Wennhage et al. 2007), as many juveniles spend their first summer in the shallow waters of the Wadden Sea, the German Bight, as well as other favorably-conditioned places in the southern North Sea (ICES 2007c). In the autumn and early winter, the juvenile plaice move to deeper water, and in the spring they return again to the shallows of the coast, exhibiting a fidelity behavior that will last their lifetimes (ICES 2007c). As the juvenile plaice age, they will habitually venture into the deeper water of the central and northern North Sea (van Keeken et al. 2007). As such, plaice year classes are quite spatially stratified. An example of this can be seen in Figure 1, where smaller (younger) place are concentrated close to the shores and in the South.

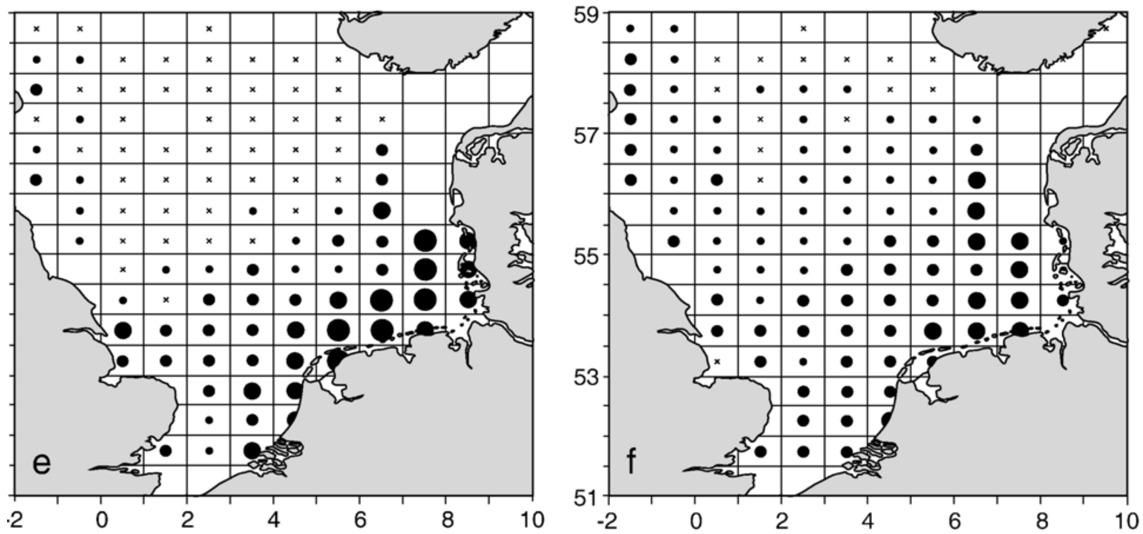


Figure 1: Plaice catch of 15-19 cm individuals (left) and 20-29 cm individuals (right) between 1999 – 2003. From (van Keeken et al. 2007).

It is primarily for this reason that plaice fall victim to a high level of bycatch and discarding by fisheries. Sole, from family *Solidae*, generally are worth 12 Euros per kilogram, while plaice are worth a mere 2 Euros for the same quantity (van Densen 2007). Adult sole are located in the southern and southeastern North Sea, in the same areas as juvenile plaice. Additionally, the mesh size of the nets used to catch sole is 8 centimeters, while that for plaice is 12 cm. The minimum landing size, MLS, for sole is 24 cm, and 27cm for plaice (van Densen 2007). It so happens that the small mesh for sole also catches plaice, and since the MLS for plaice is larger than that for sole, many undersized plaice are inadvertently trapped.

Normally, an obvious solution to this would be to go to an area where sole are plentiful and plaice are less prevalent or at least occur in landable sizes. Unfortunately, no such place exists because juvenile plaice feeding grounds are shifting to deeper water due to availability of food (van Keeken et al. 2007) Complicating this phenomenon further are the Total Allowable Catch, TAC, and other such regulations. Logically, fishermen flock to the areas where they can earn more money by landing less fish, in this case, the southern North Sea. Unfortunately, in order to catch more sole, they also have a large bycatch of juvenile plaice. According to van Densen (2007) nearly three out of four plaice that are caught are subsequently discarded. In addition to the TAC, the days at sea are also regulated in the Netherlands. This also acts to keep boats closer to shore and therefore catch unlandable fish (ICES 2007c). Another aspect that drives fishermen to stay close to the shore is the current price of fuel (ICES 2006). Venturing further offshore means spending more money on fuel to get there.

Still though, the technology that is used in the demersal, mixed-species flatfish industry is advancing at a faster rate than the fish can reproduce and regulations can be imposed (Valavanis 2002). It is currently estimated that fish are two times more likely to die from being caught by fishermen than of natural causes (Daan and van der Mheen 2004). This clearly suggests that the regulations in place are not sufficient. In fact, Rijnsdorp et al (2006) suggest that the Common Fisheries Policy, an EU initiative, as well as others, have failed outright at maintaining the sustainability of the fishing stock and industry. Although there is some disagreement about the usefulness of TAC, Rijnsdorp et al (2006) maintain that the TAC is not an effective strategy and that new regulations are desperately needed to reduce the discard rate of fisheries.

1.3 Problem Definition

Daan and van der Mheen (2004) argue that setting a new TAC and other restrictions do not resolve the overfishing problem, they merely reduce it. Fishermen, scientists and policymakers really need to be working more closely with each other. All of the information that is available to one group should also be made available to the others. For instance, scientists use many different graphical displays to represent their data, but fishermen are generally only familiar with simple, two-dimensional plots of their catch data (van Densen 2007). This lack of a common understanding between groups leads to negative results and even distrust. Fishermen do not trust scientific data because they are not accustomed to viewing it, while scientists aren't willing to accept fishermen's data because it is often corrupted by misreported data or excessive discard data (Rijnsdorp et al. 2006). Fortunately, the intentional misreporting of data has recently not been a grave problem. Still though, discarding is perfectly legal and statutorily, fishermen do not have to report their bycatch (van Densen 2008). In fact, scientific data is not flawless at all. In addition to the problem of being corrupted, this data is also not representative of the situation in the sea because statistics are generated from very limited observations, from both survey vessels and commercial fishery catches (ICES 2006).

A start at ameliorating this trust and reliability problem could be the use of a new visualization strategy which communicates the phenomena more effectively, such as a GIS. The display system must be simple and understandable, especially by non-GIS professionals (Fortunati et al. 2002). Of course, the specific needs of the GIS will determine what is necessary to create (Valavanis 2002), but at the very least, it should be presented using a well known software package, preferably compatible with other formats (Garces et al. 2006).

Unfortunately, there is not an abundance of literature on this subject because oceanography GIS is a relatively new field of study. Luckily though, many marine GIS and datasets have swiftly been put in place around the world, including a GIS for the North Sea area (Valavanis 2002), so there are case studies to use during the process of creating one for plaice. As Wood (2000) explains, human cognition is visual, thus using life history data in an animated migration pattern will already be a good start.

1.4 Research objectives and questions

As has been discussed previously, plaice fisheries in the North Sea are not sustainable (ICES 2007b). In order to allow the fish to return to healthy ecological levels, some sort of intervention is necessary. The CFP regulations such as TAC, MLS and the mechanical restrictions have not done enough to ensure the future of this species (Rijnsdorp et al. 2006). When scientific data and recommendations do become available, they are dismissed by fishermen as not representative of reality (van Densen 2007). Therefore, we will be exploring, through the use of geovisualization, how to improve communication between fisheries stakeholders and trying to ensure the mutual understanding and acceptance of their varied spatial-temporal data.

In order to accomplish this task, we have to first identify a stakeholder target group. Since this is loosely set out already, it should not be a terribly drawn-out process. We also need to know more about the stakeholders in order to better craft a visualization tailored to their needs and desires. Their competencies and past experience need to be chronicled and understood, so that they may be worked in to the chosen visualizations. Most of this work can be deduced from the available literature and from communications with Wim van Densen.

The second part of this research is the requirement for an inventory of the numerous datasets available to us. As mentioned, there are numerous datasets already collected, compiled and created (Valavanis 2002), it is now a matter of sorting through this vast body and determining which datasets are most appropriate for this study. Whether datasets are publicly available or otherwise restricted shall also be considered. In which format and the how up-to-date the data are is also of particular interest. Conversely, the manner in which the datasets were collected and their accuracy are not the primary concerns of this study however, as long as the data suits the needs of the visualization tool and catalyzes discussion among stakeholder groups.

The primary concern of this research is how to model the data in a GIS so that it is understandable and usability is maximized. The objectives of this study, therefore, will be techno-social, focusing on participant design and cooperative understanding. From the outset of this project, it has been understood that one design solution should be an animation, or "filmpje," so the usability issues relating to such kinds of media will be focused upon. Different cartographic techniques will be employed separately and applied to the animated visualizations. Which options for interaction with the GIS that are selected will be considered and decided upon based on the ultimate needs of the users.

These technical specifications will be tailor-made to the needs of those people using the GIS, drawing upon the fields of cognitive science, usability engineering, human-computer analysis and others. Although an attempt will be made to make the visualizations usable from the outset of the design process, we will require that the usability be tested. The usability of the visualization will be tested during an interview / evaluation session with stakeholders. A questionnaire will be employed in order to judge the usability.

According to the past several paragraphs, the following general research questions will be addressed in this study, with particular emphasis placed on the bulleted items.

1. Which stakeholders will be involved and what do they require of the visualizations?
    - Past experiences
    - Competencies
    - Needs and desires

2. What data may best be used for the visualization and what are their characteristics?
    - Specific datasets available
    - Specific datasets chosen
    - Format, accessibility, currency

3. How can the selected data be visualized to best support understanding and acceptance?
    - Data processing
    - Cartographic techniques
    - Animation specifics

4. How do stakeholders understand and accept the selected visualization?
    - Evaluation criteria
    - Test methodology
    - Static maps versus animated

1.5 Contents

Following this initial Introduction, five more chapters will be presented. In Chapter 2, the fishery data will be discussed, including some fisheries data background information, an overview of the sources utilized, analysis and discussion of the chosen data, and possible constraints which the data may have on our work. Chapter 3 is on the subject of visualization theory and different ways to incorporate it with our fishery data. Chapter 4 will describe test methodologies for usability and how our questionnaire has been created. Chapter 5 highlights notable results obtained from the usability tests. Finally, Chapter 6 is the conclusion, discussion, and recommendation for future testing in this domain.

2. Fishery Data

It should be clear up to this point that one of the main factors inhibiting people from understanding the problem of unsustainable activity in North Sea plaice fisheries actually is the state of the data, itself. This chapter is thus devoted to fishery data, with the hope of clearing up any misunderstandings. The first few sections will attempt to define more precisely which phenomena the data represents and how available it is. There will then be several sections dealing with the actual data utilized in this study, including an in-depth analysis of its usefulness for the proposed visualization.

2.1 Context

As explained in the Introduction, there is a conflict between scientists and fishermen when it comes to setting limits for Total Allowable Catch (TAC) because of the ongoing debate on the evaluation of stock size (Daan and van der Mheen 2004). Historically, the European Commission sets the TAC based on information provided by the regional fisheries agencies (Rijnsdorp et al. 2006). What is already clear is that neither scientific survey data, nor data collected by commercial fishermen (as limited as it is) are representative enough of the actual ecological situation. To better illustrate this, let us consider the seasonal life history patterns of plaice in the North Sea and how (where) the fishermen exploit them.

As was already discussed, plaice distribution and migratory behavior is tightly spatially correlated and overlaps with the habitat of sole in the North Sea (Eastwood and Meaden 2000). Spawning of the adult plaice occurs in the shallow coastal waters in January and February, after which a 3-4 month larval period occurs, where the young, unhatched recruits follow the tidal drift (ICES 2007c). In the summer months, juvenile plaice congregate in coastal estuaries such as the Wadden Zee and German Bight in areas with around 1-5 meters of depth (Wennhage et al. 2007). Then in early winter, plaice juveniles move off the coast into deeper and colder water (ICES 2007c), returning in the spring. Thus, the general pattern of plaice migration is summer feeding in deeper regions, winter spawning in the shallows (van Keeken et al. 2007). The result of this is that plaice individuals are moving into deeper water as they age, but additionally, later generations of plaice are moving into deeper water as a result of the changing environment (ICES 2007c). (Figure 1).

Uncertainty in the exact location of plaice during the year is problematic for the fishing industry for a number of reasons. Coastal areas, which are only 1% of the total sea area, account for 90% of the total catch of fish. Similarly, the 200 mile Exclusive Economic Zones (EEZ) account for the same percentage of the fish catch (Valavanis 2002). This means that fishermen are sticking close to shore when fishing, and taking into account this shift North of plaice, this should result in catching more undersized individuals. One estimate (Aarts and van Helmond 2007) shows that 60 percent of the catch close to shore is discarded, whereas in the regions farther North, discards only account for around five percent of catch. Similarly, the so-called Plaice Box, instituted in 1989 in order to preserve juvenile plaice nurseries by cordoning off from fishing an area in the German Bight, has now become insignificant, given the displacement of plaice (van Keeken et al. 2007). Furthermore, drastic seasonal variations in discard prevalence (twice as many discards in September as in December (Aarts and van Helmond 2007)) shows that fishermen are clearly not exploiting the resource as effectively as would be hoped.

Given these challenges in the assessment of fish stocks, data on the stocks can be adversely affected which necessarily results in management difficulties. The Common Fisheries Policy

(CFP) is blamed for a biased stock assessment and the setting of the TAC above recommended levels (Rijnsdorp et al. 2006). This is partly due to discard sampling being so expensive, thus only very limited survey voyages can be undertaken (Daan and van der Mheen 2004), which leads to questions on the efficiency and reliability of data. Additionally, there is a one year lag in data (Rijnsdorp et al. 2006).

Various other attempts at assessing the fish stocks have been undertaken worldwide with varying degrees of accuracy and reliability. Biological surveys and samples are also undertaken in the United States, coupled with interviews of fishermen. Federal logbooks are also kept based on the catch of the fishermen (Valavanis 2002). Other more advanced techniques such as Vessel Monitoring Systems and E-logging systems have been implemented in Greece and the South Pacific region, as well as onboard observer programs, where scientists accompany commercial vessels and record their catch (Valavanis 2002). This just goes to show that it is hard to collect standardized, reliable data across international boundaries such as the North Sea. Still, using any of these methods, it is clear that as size of fish increases, so too does the sea depth (and usually distance from shore) that the fish is found. This, at least, minimizes the need for a 3-D habitat model (Zeller and Pauly 2001).

2.2 Existing data

The sheer number of datasets that are available is impressive and daunting. As Valavanis (2002) says though, relative simplicity should be maintained while creating a fisheries GIS, especially since the users in this project are likely to be novices with this software. Many of these datasets can be acquired for free from the organizations which have already established a marine GIS (Valavanis 2002).

The common GIS fisheries datasets are rather independent of collection method. They include two different categories of data: predator-related datasets (catch per tool, landings per tool, location, tool pressure), and prey-related datasets (species areas / fishing grounds and biological data on the species such as sex, weight and length) (Valavanis 2002). Data on prey and predator is commonly collected by national or regional fishery organizations on periodic survey voyages.

Although the spatial component is important for fishery management (Booth 2000), unfortunately, another major variable has been historically not commonly collected: date of catch. This, along with the location of catch, can help to develop a better understanding of the presence of fish at certain times of the year, which relates directly to our current study. It also needs regular update, as has been explained previously; plaice distribution is in flux.

To make the data needs of this project more precise, we are looking for datasets which allow us to animate the data on a monthly scale with records for each ICES quad of the North Sea. We are particularly interested in discards, Catch (landings and/or total catch), effort and mesh size. This data should also be easily transformed into static, paper maps.

After trimming down possible data sources a bit, four different organizations seemed to offer the best possibility of finding usable data for this visualization research. They were as follows: the Marine Life Information Network (MarLIN); Eurostat, the statistical branch of the European Commission; the Marine and Fisheries Agency; and the International Council for the Exploration of the Seas (ICES). All of these sources provide data in varying forms and formats, which will be highlighted hereafter.

The Marine Life Information Network (MarLIN) has some data available for online download. MarLIN is a UK-based organization and only keeps information for UK waters. For plaice count data, the study area covers many areas of interest on the eastern and southern coasts of the England and Scotland (North Sea). Unfortunately, the trawl surveys which have made their data available were only conducted in 1992, 1994, 1998, 1999 and 2002. Of these, ’92 and ’94 are extremely geographically limited, while ’98, ’99 and ’02 are quite complete for UK waters. This data is not nearly complete enough for use in this project, also because the count numbers of plaice are not divided per age class, nor market category.

Eurostat possibly possesses the most standardized data, meaning that the dataset fields are identical and complete over each year, but that does not mean it is the most reliable. They do offer plaice landings per country, monthly for the “North Atlantic” and annually for individual ICES areas. Unfortunately, this spatial breakdown is not sufficient for the purposes of this study, as we would like to plot the data more precisely.

The Marine and Fisheries Agency, of the UK, provides exhaustive landings data from 1995 – 2005, in the form of PDF reports. Monthly reports are available of the “Return of Sea Fisheries Statistics.” Several interesting subcategories of data exist, chiefly: landings into UK by UK vessels, landings into UK by foreign vessels, and landings abroad by UK vessels. Unfortunately, there is not data for landings abroad by foreign vessels. Count data is also devoid of a location indication, making it, unfortunately, useless. There also exists a dataset of landings by ICES area of capture, but, unfortunately, this data is annual, and not monthly, as was hoped for.

The International Council for the Exploration of the Seas (ICES) has numerous datasets with differing availabilities. Currently, there are data on Spawning Stock Biomass (SSB) per year and ICES quad, available online for free download. The more useful datasets are protected by password-secured download. These datasets come in three types: standard maps and graphs, aggregated data, and raw data. The standard data is per survey area, concentrating on abundance of “relevant” ages of fish. Aggregated data is more interactive and can be manipulated more easily. Raw data is, of course, the most useful because it can be mined for its most relevant components. Components include: abundance per age/length, and SMALK (Sex, Maturity, Age-Length Keys) per individual. The data which ICES provides originally comes from its member country fisheries agency; so in this case, IMARES, the Dutch Institute for Marine Resources & Ecosystem Studies. Thus, the requested data was obtained from several researchers at this institution.

2.3 Used data

Two different datasets were sought from IMARES: one detailing catch and effort, and another for discards. The first one was obtained from Floor Quirijns and the second from Edwin van Helmond. For the initial evaluation of the data suitability, please refer to Appendix 1. This section is devoted to further detail about the two data sources.

Catch and Effort data were obtained from a dataset which originally provided Catch Per Unit Effort (CPUE). Catch is normally the total amount of fish taken by a fishery (ICES 2009), but in this case, since it is separated into market categories, it can be assumed to actually be the landings (or total amount of fish taken over the Minimum Landing Size (ICES 2009)). The definition for Effort is given by Aarts and van Helmond (2007) and van Densen (2008) and is understood as the amount of horsepower of fishing vessels multiplied by the number of days fishing. Catch Per Unit Effort (CPUE) is the amount of fish caught divided by the Effort, so as to standardize it a bit.

This dataset is just a small fraction of a much larger database including other commercial species. It was created by IMARES through sampling of fish markets (van Densen 2008). In this dataset, a number of fields existed. They were:

- Year
- Month
- ICES quad
- Mesh size
- CPUE 1
- CPUE 2
- CPUE 3
- CPUE 4
- Effort

Two years were selected for the data: 2005 and 2006. Data from each month was preserved. ICES quads did not cover the entirety of the North Sea but a large enough portion was available. The mesh sizes denoted were somewhat already classified how we envisioned them being. Separating them into several categories was originally envisioned so as to compare discards of different mesh sizes. CPUE (in kg/hpday) was broken down into plaice market category sizes, 1 being the largest fish and 4 the smallest. Lastly, effort, expressed in 2000 hpdays, makes up the most important variable.

This dataset is relatively complete. Spatial incompleteness already having been mentioned, the temporal completeness of the dataset should also be considered. Indeed, there is not data for every ICES quad for every month in the proposed time frame of two years. Nonetheless, we will continue with the visualization and decide afterward whether the dataset still shows enough to be useful.

The Discards dataset is part of a voluntary data collection scheme developed by the Dutch Fish Product Board, Productschap Vis (PV). Since IMARES scientific sampling surveys are conducted on less than one percent of the total fishing effort (Aarts and van Helmond 2007), a need exists to try to collect more fishery data. Anyway, it is not obligatory for commercial fleets to report their discards (Rijnsdorp et al. 2006), so we need this voluntary data source to get an idea of the phenomenon of discarding of plaice.

There were many extraneous fields which were included in this dataset that were of no use to this research but the following ones were helpful:

- Year
- Week
- ICES quad
- Mesh size
- Date
- Latitude, Longitude In
- Latitude, Longitude Out
- Volume over MLS
- Volume under MLS
- Discard percentage

This dataset was not particularly refined or standardized. Of note were the notation of time, spatial description and mesh sizes. Some records in the dataset were denoted by week and year, while some were by date. Thus, the weeks and years were converted to dates, with each record receiving the date of the first of the corresponding month (as will be explained later). Some records in the dataset had an associated ICES quad, which was very useful, as they seemed to match up with the ICES quads in the CPUE dataset. Unfortunately, some records only listed latitude, longitude demarcations. It also had to be decided which of these would be used, since one coordinate set represented where the nets were put into the water (Latitude, Longitude In) and the other where the nets came out (Latitude, Longitude Out). Ultimately, the Out was chosen. Still other records (actually about 50%) were missing both spatial identifiers altogether. Mesh sizes were not nearly as regular as those for the CPUE data. Thus, we had to group them into categories, such as the ones we found for the CPUE data. As with the location information, about half of the records are also missing mesh sizes. MLS, as in Volume over MLS and Volume under MLS, stands for Minimum Landing Size. This volume is measured in liters (Aarts and van Helmond 2007). The Discard percentage had been computed by dividing the Volume under MLS by the total landings. Again, nearly half of the records are missing a discard percentage, and it cannot be computed because they are also missing landings information.

There is a serious problem with this dataset, as it is missing so much of the vital information that we need for this research. It was originally envisioned that data should be separated into mesh classes to be able to compare discards per small, medium and large mesh sizes, but given the current state of data incompleteness, this goal seems unlikely. Records for mesh size 80mm are relatively plentiful, while for 80-120mm and above 120mm, there are hardly any records which have survived the necessary deletions described in the last paragraph. As will be attempted for the CPUE and effort data, a visualization will be created and later deemed either sufficient or insufficient for using in the usability test.

2.4 Data Processing

Since the dataset of CPUE and effort was nearly “GIS-ready”, we will begin with that one. As mentioned previously, the dataset’s records and fields were very “clean” and did not require very much preprocessing. The primary change needed was the reclassifying of mesh sizes. Instead of being structured as one mesh size per class, we wanted to have the following classes: 80mm, 80-120mm and 120mm+. 80 was actually already the smallest mesh size, so that required no work, and neither did 120, as it was the largest. Combining those records between 80 and 120mm took slightly more time and effort.

It was also necessary to change the notation of time, to include the month and year in the same field. This had to be done in order to display the data according to the format required for animations in ArcGIS, which was the software chosen to build the visualization. The spatial and temporal resolution does seem to be extensive enough that a visualization of effort could be successfully undertaken and appreciated by viewers.

As may already be expected, the discards dataset caused the most problems. In fact, some of the preprocessing had to be undertaken in ArcGIS and while some could be manipulated directly in the database. Converting Latitude, Longitude to an ICES quad took some work. In fact, since the data represented single instances of hauls, there were potentially numerous records per ICES quad per time slot. Thus, these had to be aggregated and converted to polygons in order to be displayed as ICES quads. Secondly, the time field(s) needed to be edited in order to properly aggregate the data and eventually represent it in a format usable by the animation function of ArcGIS. The data came with a week number and therefore had to be converted to the approximate month that the week corresponded to.

Ideally, discards and catch would have been included in the same dataset, so we could be ensured that the data would be comparable. As it was not from the same source, we had to make an assumption that the discard data that were available were representative of the whole monthly discard percentages per ICES quad. Assuming this, we were able to combine the two datasets in order arrive at a sort of juxtaposition of the data as can be seen in Figure 7. This joining of data was based on the ICES quad, time period of the data and mesh size. That is to say, if there were discard data for July 2005 in quad 35E6, they would be appended to the CPUE and effort data record of the same spatial and temporal occurrence.

Unfortunately, there were very few records of discard which overlapped with those for CPUE and effort. While we would have liked a one-to-one relationship, with one complete record (CPUE and discards) per quad per month, yielding approximately 4800 records per mesh size, we received about 10 each for 80mm and 120mm, and around 100 for 80-120mm. Please refer to Appendix 1 for further explanation. We are therefore left with an extremely paltry number of data records.

In order to be able to compare discard percentages and catch, we will have to make a sum of the catch data in order to know the total catch and not just the catch per market category. Having this total will then allow us to know the percentage that each market category makes up. We must also deal with the fact that discard data is presented in volume (liters) and catch data is in kilograms. To convert volume to kilograms, we must divide the volume by 0.89 (Aarts and van Helmond 2007).

When finally finished with processing the data, we see that it is not even close to a complete dataset. Had we intended to visualize the data for all mesh sizes combined, the result would have been slightly more interesting but still nothing like what we expected or wished for at the beginning of this project (Appendix 1). The geographic coverage is not even complete when looking at all months together, let alone each month individually (Figure 2).

Attributes of mesh80_combined

OBJECTID *	Shape *	Join_Count	ICES	date	MEAN_VOLUME_ABO	MEAN_VOLUME_BEL	YEAR	MONTH
5	Polygon	1	33/F2	01/05/2005	2	2	2005	5
18	Polygon	1	33/F4	01/05/2005	18	11	2005	5
54	Polygon	1	34/F3	01/10/2006	3	6	2006	10
65	Polygon	1	36/F3	01/01/2005	48	13	2005	1
112	Polygon	1	37/F6	01/06/2005	12	3	2005	6
133	Polygon	1	37/F6	01/11/2006	5	1	2006	11
142	Polygon	1	39/F3	01/05/2005	5	0	2005	5
160	Polygon	1	39/F6	01/11/2006	77	30	2006	11
173	Polygon	1	40/F6	01/11/2006	42	19	2006	11

Record: 1 Show: All Selected Records (0 out of 9 Selected) Options

Figure 2: Data records after combining discards and catch for 80mm mesh into a single dataset.

2.5 Conclusion

Mostly in light of the state of the discards dataset, it does not appear wise or useful to employ this visualization in the usability testing. We need a dataset which is complete (or nearly so) in order to test the stakeholders' understanding of animations. Hopefully, we will be able to use the discard data somehow, as we found percentages of discards to be very similar to what van Densen (2008) suggested was likely.

There were some decisions made which could have slightly affected the final results, highlighted in the previous paragraphs, but they would have not eliminated so much data. We are talking of, for example, the decision to use the Latitude and Longitude Out coordinates as the spatial location for the discards and the conversion of weekly dates into monthly ones. While it is likely that some fishing hauls cover more than one ICES quad, there is no way for us (or even the fishermen themselves) to know how much of the haul came from which quad; therefore a decision had to be made. Also, the aggregation of weeks into months did change the discard totals slightly, because a month is not exactly four weeks, so there are cases of as much as three days of discard data being added to the wrong month. This is again though, an unavoidable side effect.

It is clear that in order to be able to work with this data in the future, the completeness of this dataset must be improved. Perhaps it is possible to persuade more commercial fisheries to take part, but at least the quality of the data that is already collected should be improved to ensure the inclusion of time, space and mesh sizes.

For the proposed visualization, we will therefore go on with the effort dataset and try to make something useful out of that. While visualization of effort is nothing new in fisheries science and management, studying how stakeholders interact with visualizations is indeed novel.

3. <u>Visualization of Fishery Data</u>

"A viable data model of the real world that encapsulates the essence of the phenomena under study is a pre-requisite for the efficient management, use and communication of geospatial information" (Fairbairn et al. 2001).

As can be deduced from the above quote, the fundamental premise of visualization, and in this case cartography, is the transformation of non-displayed data into a form of representation that can be projected onto a map. Fairbarn et al (2001) also provide us with a useful schema of five issues related to visualization: users, data, form, purpose and technology.

All of these issues must be addressed when creating a visualization. *Users* must have an adequate understanding of the visualization. The choices of which techniques to use depend largely on cognitive issues of users, which will be summarily explained in this chapter but which are a central focus of Chapter 4. *Form* deals with visual design and user interface. Basic recommendations for graphic (and indeed cartographic) design will be mentioned. *Purpose* addresses the tasks that the user must carry out. *Technology* relates to the new and varied ways in which data can be represented. More developed methods for displaying data, given the new technology, will be highlighted toward the middle of this chapter. *Data* includes issues such as generalization, organization and attributes. Although all of these issues are of importance to the choices related to geovisualization (and will be discussed), particular emphasis in this chapter will be placed on *users, technology* and *data*.

3.1 Users

Although much more will be said on the subject of users in the next chapter, it does justice to the importance of it to mention it now. When dealing with people and especially when dealing with the variation in their choices or understanding, we should take a social approach, in this case, an approach toward cognitive science, since as Plaisant (2005) puts it, there is no "one size fits all" in data visualization. There are many other disciplines which have lent their theory to geovisualization studies, including scientific and information visualization, cartography, image analysis, exploratory data analysis and Geographic Information Science (Kraak 2003). This multidisciplinary approach attempts to "provide theory, methods and tools for visual exploration, analysis, synthesis and presentation of geospatial data" (MacEachren et al. 2004).

One of the most often cited elements in the theories is the need to know the users and especially their competencies (Fairbairn et al. 2001). That way, their strengths can be played to in designing visualizations and give them a better chance of understanding what is presented. Arias et al (1997) also acknowledge stakeholder shortcomings; how their narrow views of the problem can be ameliorated through the effective use of the media. MacEachren et al (2004) even suggest that stakeholders of varied backgrounds and competencies can learn from each other through sharing, negotiating and refining their concepts of the visualization. One way or another, considering users from the very beginning will play a part in how successful a visualization's understandability will be.

In constructing a visualization, if the researcher keeps central the notion that users must find something familiar to them in the representation, then ultimately, the chances of understanding are increased. Furnas (1986) explains in his "Fisheye Lens Model" that people are highly familiar with their own environments (for example the data that they are most often exposed to) but are generally not at all adept at working with other data. This principle can also be applied to the multimedia that users are already familiar with. Staples (1993) echoes Furnas, in promoting the use of representations that contain “metaphors” which are already familiar to users. “Metaphors” in this sense can be thought of as icons or pictures which resemble something that they are meant to represent; for instance, the color blue for water. Although, Fairbarn et al (2001) do caution about using levels of abstraction that are too far-removed. Indeed, this is an important point to consider, especially in light of Plaisant (2005)’s suggestion for using map projections which are commonly used in the everyday life of the users.

Summarily, it is absolutely necessary to have a good understanding of the users since they are a heterogeneous group and individuals should not be assumed to be representatives for their larger group population. Knowing the users will help to design usable systems, by showing us what they already know, so we can design with their competencies in mind.

Norman (1993) cautions that multimedia encourages “sensory experience at the expense of thoughtful reflection.” We do not mean our multimedia representations to at all take the place of reflection or analysis; quite the contrary. Representations are meant to generate hypotheses, develop solutions to problems and construct knowledge (Kraak 2003). In another text, Kraak (2003) puts it another way, acknowledging the challenge that exists for designers to transform complex data into information and then for users of geovisualizations to transform that information into their own knowledge. The task is not as daunting as it may seem however. Maps actually allow users to almost instantly identify patterns in data (Visvalingam 1994). Similarly, Larkin and Simon (1987) explain that indexing information into diagrams leads to “useful and efficient inferences.” It is through the understanding of how this information is processed by the users that we will be able to construct “rules” for making graphics (Lohse 1997).

3.2 Form

Jacques Bertin’s pioneering work on “Fundamental Graphic Variables,” according to MacEachren (2001), seems an appropriate place to begin our exploration of forms that visualizations can take. Although using slightly different terms for his explanation of visual variables, Chang (2006) offers definitions of the terms: hue, value, chroma, size, texture, shape and pattern. Hue, generally used for qualitative data, can be described as the dominant wavelength of color, or what we generally just call color. Value is the lightness or darkness / magnitude of color (Figure 3). Chroma, also called saturation or purity, is the richness or brilliance of a color (Figure 4). Both value and chroma are commonly used to show important or quantitatively larger data. It is common to pair the two in visual displays. Size is the largeness or smallness of an object. Texture is otherwise known as the “spacing of symbol markings” (Chang 2006), or in other words, giving the impression that the graphic actually represents something tangible. Shape can be described as the geometric form of an object. Pattern, although less well-defined, is the repetitive use of objects for fill, such as hatching (Figure 5).

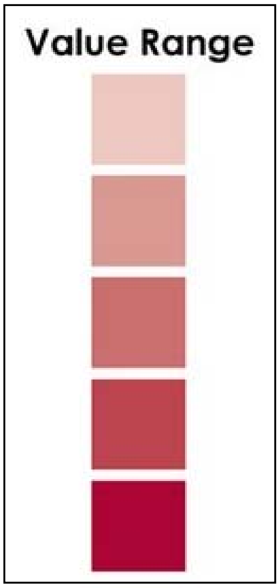


Figure 3: Color values, ranging from light to dark

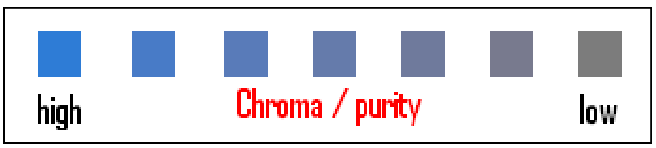


Figure 4: Color chroma range, with low values closest to gray

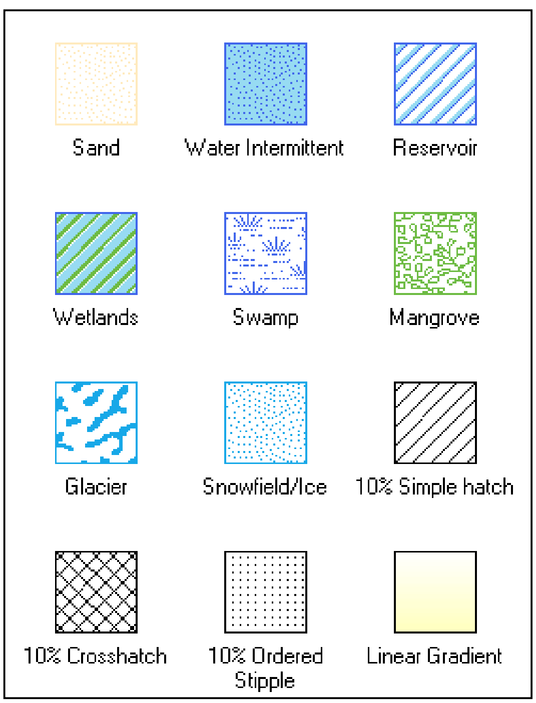


Figure 5: Various patterns available in ArcGIS

While graphic variables may seem confusing in an explanation such as this, once applied in a visualization, they should seem quite logical and the meanings apparent. Other aspects of the form of the visualization that should seem obvious on screen are the labels. Tufte, as explained by Casselman (1999) suggests that labels, whether representing values as Bertin discussed, or otherwise, should be direct and not coded. That is, there should be no question as to the meaning of them. Similarly, Tufte promotes the elimination of all unnecessary parts of a figure (Casselman 1999) to reduce confusion. For it is not only the quality of the content of the visualization but also the quality of the authorship, which makes it usable (Waterworth and Chignell 1997).

3.3 Purpose

The purpose of the visualization will, of course, change with the learning goals set forth for the experiment. Fairbarn et al (2001) proposes six characteristics of a geovisualization: level of interaction, dynamism, sensory channels, environments, media and user access. These six characteristics rely completely on the purpose of the geovisualization and will therefore be laid out here. The level of interaction is the degree to which the user can change the input or processes of the visualization. While usually passive, meaning there is little to no change to be made, there can also be a high degree of interaction. Dynamism is related to interaction. It determines whether the visualization is animated or static. Oftentimes, interaction is included in animated visualizations. Sensory channels are the manners in which the user can receive communication with the visualization; for example visual or acoustic. Environments are the setting in which the user experiences the visualization. Fairbarn et al (2001) lists immersive and remote environments as the most common. Media is, of course, the manner in which the visualization is delivered to the users, either in paper, on screen, or through projections. Lastly, user access is the accessibility of the visualization to the users.

3.4 Technology

What exactly do users get access to though?, may be the next logical question. Do they need special access or Twenty-first Century knowledge to use these visualizations or should they "rely on well-worn procedures and paradigms" (Kraak 2003) that are so common in the field? This section will attempt to delineate some of the current, creative technological trends and recommendations in geovisualization.

Roberts (2005) actually allows that starting with old visualizations is a good idea and that multiple newly updated ones can be offered to see which one is preferred by users who may be used to the old visualization technique. This variety of data representations is also seen as a good method by Waterworth and Chignell (1997), who explain that this will lead to more people having access to the media and will also allow better knowledge discovery and analysis because if one method does not work for them, most likely another one will.

Representations are no longer only static maps though. They are often animated through time (Fairbairn et al. 2001) such as with snapshots, which can be played consecutively to show obvious trends (Andrienko et al. 2005). Some argue that the simple "video-player metaphor" (Fairbairn et al. 2001) is no longer sufficient though. In order for learning and knowledge discovery to occur, some kind of interaction or manipulation of the visualization must be included. MacEachren (2001) even argues that the term "geovisualization", itself implies interaction.

While proactive roles can play an important part in visualization, there still need to be some basic, automated functions which allow novices to also benefit from them (Fairbairn et al. 2001). For instance, the ability that representations have to transcend language barriers (Lohse 1997) would be totally lost if interaction was required for a visualization in another language. This scenario further necessitates the need for effective legends and graphs for the cognition of representations (Fairbairn et al. 2001) because if words and concepts are not clear to the user, they should at least be able to gain some understanding of the data through logical, visual means. Plaisant (2005) adds that speech is also commonly added to animations to unload some of the cognitive work to a sense other than vision. Ultimately though, human cognition is visual (Eastwood and Meaden 2000).

Maximizing the number of users who are able to understand the visualization is the ultimate goal. This can be attained through providing the visualization in several forms and with visually explanatory materials.

Andrienko et al (2003) go on to explain that the visual representation method of animation is particularly effective when slides of the animation are faded into each other because the eye is drawn to the areas where change occurs. Andrienko et al (2003) also provide us with some classifications of techniques and recommendations for effective comparisons of time visualizations. Two different exploratory techniques of theirs are especially interesting: elementary and general exploration.

In order to understand the difference between the two, first consider the following scenario. The visualization is composed of three variables; when, what and where. If "when" is known (visualized), how do the data's characteristics of "what" and "where" appear? Elementary exploration would be concerned with the characteristics of individual objects (what) and locations (where) at the given time (when). General exploration considers the spatial distribution (where) of the objects (what) over the time span (when). Thus, elementary exploration tries to answer questions of detection, for example; did the objects change?, while general exploration takes it a step further into complication by measuring the phenomena and asking how much change occurred. For change detection, simple methods such as map iteration (display of sequential representations of data) and overlaying two different times should suffice for users to understand. For change measurement, however, slightly more complicated methods should be employed; namely, times series of change maps, and particularly in combination with a filter, which removes variables that are not under consideration (Andrienko et al. 2003).

In order to compare two different variables that are active during the same time resolution (for instance comparing fishing effort per mesh size), each animation should be viewed several times for the users to really memorize the state of the variables. During each reviewing of the animation, they should focus their attention on a new spot on the animation. Another option is to display the two animations side-by-side and allow the users to see directly the differences between the two variables in space (Andrienko et al. 2003).

3.5 Data

Now, we pass to the final, although probably most important, section of our discussion of geovisualization theory. The geovisualization products would be nothing without the data behind them. This section will attempt to give a state-of-the-art on the application of visualization theory to fishery data.

Fishery science, indeed, has an important spatial component to it, like no other form of natural resource or agricultural discipline. Its uniqueness is that fish are not seen (usually) and therefore fishermen go to areas where there have historically been fish or to areas where ecologically, they are expected to be. Many fish species are also highly mobile. Geographic Information Science (GIS) has been employed in fishery science since the early 1990's, including for estimation of bycatch, based on life history parameters (Valavanis 2002). Among the fishery science literature consulted for this research project, this genre of application for GIS was, indeed, well represented (see (Stoner et al. 2001; Zeller and Pauly 2001; Valavanis et al. 2004) among others). In the recent past, coastal and ocean management has used GIS as a communication tool for policymakers and has met with great success (Wright and Bartlett 2000; Valavanis 2002). The essential question is now why GIS has not been at the front of fishery science as well.

The lack of integration of GIS into fishery science can be explained in both technical and social terms. Key among the technical reasons is the fact that fishery science, and fish life history patterns, are extremely complex and depend on many environmental parameters (Caddy 2001), bycatch data is limited and conclusions are hard to make due to the dynamic patterns of fish migration (Valavanis 2002). As a consequence, people have been reticent to attempt projects of this sort. If undertaken, the project can quickly become meticulous and drawn-out due to the heterogeneous data types and formats which must be integrated (Devillers et al. 2006). This can lead to extremely intensive computational modeling, which deters even more would-be researchers (Kwan 2004).

Hesitancies related to social issues can be even more daunting to overcome, as feasible solutions may be harder to find. There is a general reluctance of large scale fishing industries to adopt GIS or Remote Sensing (RS) technologies and that means that policy is not informed as precisely as it otherwise could be. That is, indeed, unfortunate because in the few places where geospatial technologies have been introduced, they have been a huge success (Valavanis 2002).

Although GIS is new to fisheries science, it has met with great success in other maritime disciplines. The lag in application is understandable though for technical and social reasons.

One such place is in Greece, where the use of Vessel Monitoring Systems (VMS) has increased in the past few years thanks to an enterprising pilot project (Valavanis 2002). This system is a great improvement over the old system, which is still the norm in the USA and the Netherlands (Valavanis 2002), although there is strong support among Dutch fishery scientists for such an initiative (Daan and van der Mheen 2004). The outdated system relies on manual logging of vessel geographic position and is inherently flawed, but can be validated with the use of GIS (Valavanis 2002). In a different initiative, United Nations Food and Agriculture Organisation (FAO) is working on implementing a worldwide fisheries GIS, called FiGIS, intended for use by policymakers (Valavanis 2002).

Before implementing any GIS-related visualization technique, certain precursory planning steps must be taken in order to increase the chances of success. From the start, the designer should know whether they intend to implement a GIS with interaction (called "direct manipulation") or without interaction (also known as an animation) (Bishop and Lange 2005). The visualization technique to utilize should also be based on the characteristics of the data; in other words, it must match the data type and structure (Fairbairn et al. 2001). This was a particular challenge in this research project and will be further examined later on. The determination of the tools needed for the GIS should also be thought about early on in the product development process (Valavanis 2002) because this decision should be based on the research questions. It should be understood, however, that not all issues can be foreseen when planning in the formative stages of visualization research (Houde and Hill 1997), so there should be some room for change if unforeseen issues arise partway through.

3.6 Fishery data integration

By basing our own research on the principles above, it is hoped that this project can contribute to the better understanding of fisheries stakeholders' understanding and therefore, use of geovisualizations. Ultimately, through the incorporation of all recommendations by fishery stakeholders into the visualizations (Politis 2003), we hope to be able to the create "rules" of graphics-making, as Lohse (1997) put it.

Before reaching that goal however, there are three stages of production which need to be undertaken before the visualization can be delivered to its users: to build the visualizations from the raw data that was highlighted in the previous chapter, to assure the layout properties of the visualization, and to construct the animation which will deliver the visualization to the stakeholders.

3.6.1 Building the visualizations

Building the visualizations was one of the most difficult tasks of this research. While basic ideas of what form the visualization should evolve into were well known (van Densen 2008), the exact manner in which to approach the subject was a daunting task. The first thing was to find appropriate data and then to decide upon the software that should be used to transform and ultimately, to deliver the visualization to the users. This, of course, depended on which tools or functions were necessary for the visualization. Thus, after consulting the literature related to previous examples of fishery information systems, a simple, animated ArcGIS-based application was decided upon, which will be further detailed in the next sections.

One does not have infinite choices though when transforming data into visualizations. Visualization choices must be based on the characteristics of the data available, specifically the structure and type of data (Fairbairn et al. 2001). Begg et al (1999) utilized dot density for fish distributions on static, periodic maps. This is an interesting concept but our data is not available in a format compatible with this method. Drapeau (2001) created temporal visualizations of fishing effort per month, which he implements with raster images. This is an example which is quite close to ours except we plan to use more data than raster format could contain. Therefore, we will branch out in our own direction and use polygons, which seem to satisfy our needs more closely. In this case, there are other important limitations to the data that were discussed in the last chapter, which must be taken into account. Whether these limitations inhibit or prohibit the visualization from being properly understood and absorbed by the users has yet to be seen.

Two important criteria to decide upon are the data and software to be used. These were selected with the help of the literature review and consultation of other fishery GIS examples. Ultimately, ArcGIS was chosen for the software, while it was necessary to take into account the limitations of the data in order to best know how to transform it.

The visualizations that are envisioned can be fit into four themes: comparisons of discards within mesh sizes (a), comparisons of discards between mesh sizes (b), effort (c) and a theme integrating discards, CPUE and effort (d). They are called themes because they encompass the same data and data structure but are each represented in a variety of manners, which should lead to better access and knowledge discovery by the users (Waterworth and Chignell 1997). The different data handling steps undertaken in producing these visualization themes were not homogenous and will now be explained.

It is wise to provide an example in order to explain, in the best manner, how the first two visualization themes can be realized (Figure 6). Imagine that there are three categories of mesh sizes that we would like to segregate the discard data into; under 80 millimeters, 80-120mm and above 120mm. In order to visualize the discard percentages, it is necessary to break them down into classes. For this example, let's assume they are between 0 and 100 percent of the total catch. For the comparison within mesh sizes (a), we can simply break the three datasets down into classes of any limit we choose, because a comparison will not occur *between* the different mesh sizes, but only *within* them. That is, we are not interested in comparing them to each other, only in comparing the spatial relationships of discard percentages within the same mesh size dataset. For instance, we compare discard percentage for 80mm mesh for the German Bight versus that which occurs off East Anglia, in England.

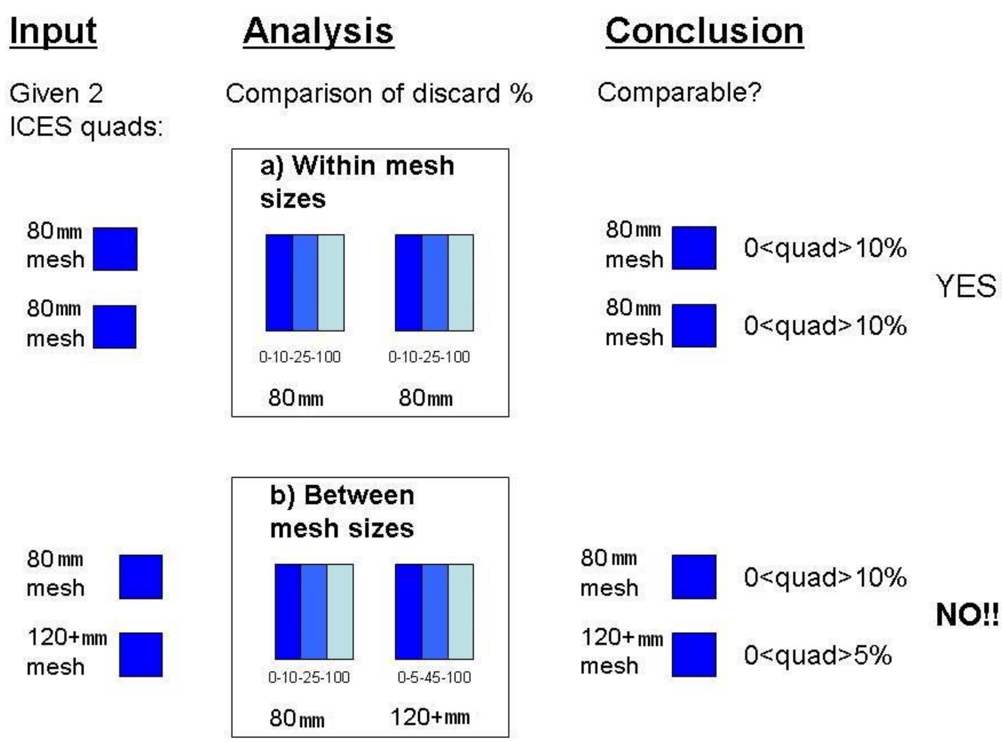


Figure 6: Explanation of need for two different discard datasets

Things get a little more complicated when we want to consider the comparison between mesh sizes (b). Here, it is absolutely necessary to have class limits of the same range and interval in order to visually compare discard percentages *between* mesh sizes. In Figure 6, above, think of this in terms of colors. If dark blue represents 10% and lower discards for <80mm mesh size, the same color should represent 10% and lower for 80-120mm and likewise for the dataset of >120mm. The trouble lies in the fact that the ranges of values (discard%) are not the same for each mesh size. One could even logically conclude this since it is expected that larger mesh sizes will not catch as many fish. Since the total ranges are not identical, neither will be the individual class ranges. The same class ranges and intervals were thus not used, so the colors for class range also do not correspond to the other mesh sizes. Therefore, it would be impossible to visually surmise, for a given quadrant, how the discards compared for the <80mm mesh or 120+ mm if we use the *within* class limits. A set of class limits based on comparisons *between* mesh sizes is necessary to answer questions about comparisons of discard percentage between mesh sizes. For instance, comparing discard percentage using 80mm mesh nets versus 120+mm mesh nets for the ICES quads off the coast of Holland.

The visualizations for effort (c) suffer from the same trouble as the above examples, therefore both *between* and *within* versions were made for effort as well.

Before we really go into choices of color and other graphical representations, the fourth visualization theme (d) should be highlighted. This theme is truly a hybrid because the data come from diverse sources and have possibly never before been combined as they were here. The Catch Per Unit Effort (CPUE) data needed to be transformed in order to be used in collaboration with discard data. In short, the CPUE was divided into merely Catch; disregarding the Effort. The Catch data was thankfully already separated into market categories; 1 being the longest and 4 being the shortest. After the discard data were converted into a comparable percentage of total catch per quadrant, they were "appended" to the catch data (marketable categories) as a sort of fifth class (Figure 7).

There are four themes of data which could be of use:

a. Discards compared within mesh sizes
b. Discards compared between mesh sizes
c. Effort
d. Discards as a percentage of catch

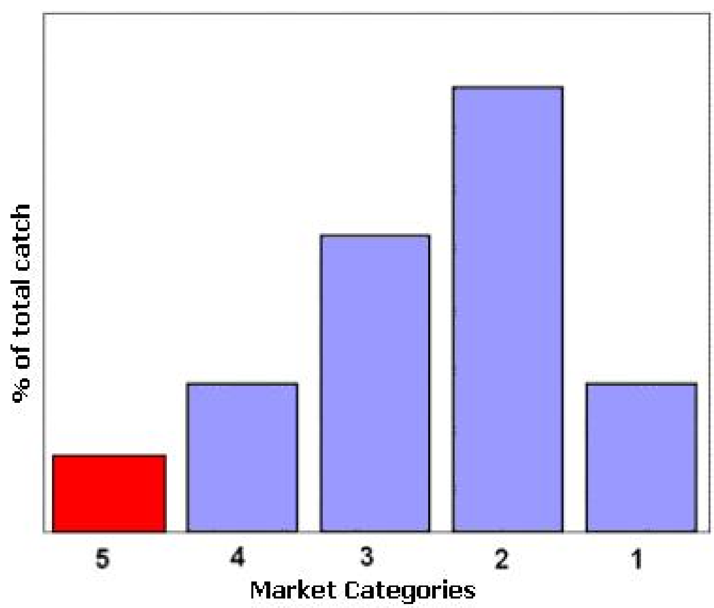


Figure 7: Example graph of discards (red) versus market categories (van Densen 2008).

As was discussed in the last chapter, there are serious limits to the utility of this visualization theme because discard data is so scarce. Nonetheless, the idea is a good one, and we will continue now with a discussion of the color choices and other cartographic techniques employed.

Since color was briefly touched upon earlier, this seems a good place to begin the discussion on visual variables. Initially, in order to represent different discard amounts for visualizations (a) and (b), a gradual variation in hue was used, but it did not sit well in pre-trials, and indeed this principle is supported by the literature (Chang 2006). Hue variations are best used for qualitative data and quantitative data can be represented well by a combination of chroma and value (Chang 2006). Thus, the visualizations are now displayed in a beige to brown scheme, with each ICES quadrant receiving an appropriate color. Effort (c) is also displayed primarily using color for two of the three visualizations; one as above, and another of colored dots. The use of color is also an important aspect of the (d) visualizations, where pie and bar graphs are utilized. Since the large percentage of the total catch that discards occupy is obviously a detriment to the population of plaice and the industry's sustainability, it is represented in red (Figures 8 and 9). This sets it apart from the catch data in a metaphorical way which Staples (1993) would have supported. As a negative aspect, the red discards resemble a stop sign, and therefore a contrast is made with the blue catch data, allowing the user to see how large the discards are compared to the rest of the catch.

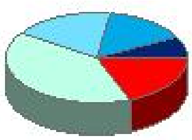


Figure 8: Example pie graph of discards (red) and catch per market category

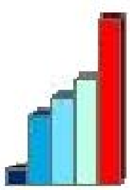


Figure 9: Example bar graph of discards (red) and catch per market category

The different display strategies warrant a closer look. The strategies for (a) and (b) are fairly straightforward and should not arouse too much confusion but can be seen in Figures 10 and 11, respectively below. For effort (c), a graduated symbol map was created (Figure 12), whereby larger efforts are represented as larger dots on the map. There is concern, however, that the sizes of the dots are not varied enough and therefore are not conducive to visual comparison. If the dots are increased in size though, unfortunately, they will cause spatial displacement of other dots, which is also the case with the bars and pies in the next visualization theme. The bar graph strategy of visualization theme (d), where discards are compared to catch per market category can be seen in Figure 9, and pie graphs in Figure 8, above.

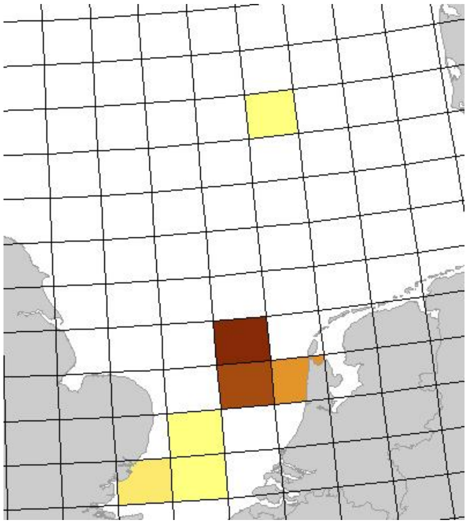


Figure 10: Discards compared within mesh sizes

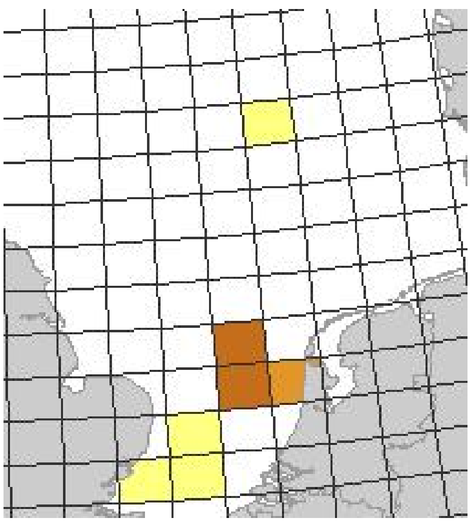


Figure 11: Discards compared between mesh sizes

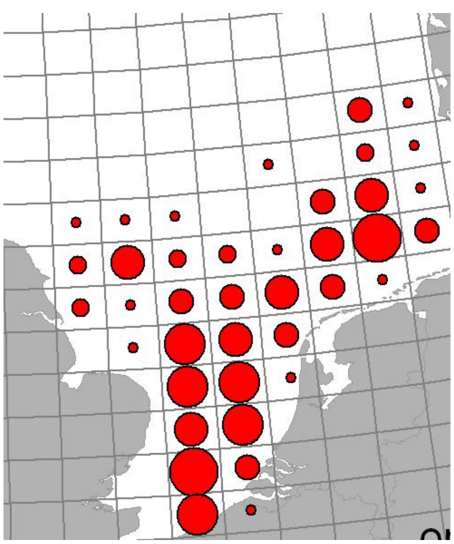


Figure 12: Proportional symbol visualization of effort

The major differences being, as mentioned previously, the color scheme will be different and in many cases, the discards will be much higher. The pie chart may show more effectively the relationship between discards and catch because it displays the classes in terms of a circle and therefore percentages can be more easily estimated.

Color can be used effectively to symbolize quantities. Different display strategies such as graduated symbols and bar or pie charts also show quantity well.

3.6.2 Layout properties

Layout properties include things that are not geovisualization characteristics but are directly related to them. They are concerned with the spatial and geospatial representation. Indeed, these data are geographic in nature, and should therefore be as precisely located as possible. We saw in the previous paragraph that some visualization strategies do not accomplish this precision goal very well. It is all the more important then that when possible, data should appear to be sited in its actual location.

It is for this reason that coordinate systems, projections and scales must be considered. Especially for novices in geovisualization theory and use, projections used should be familiar ones (Plaisant 2005). For fishermen, who are used to referring to nautical charts, it would also be useful to use a coordinate system that they use frequently to aid in their understanding of the geographic location of the data. In nautical charts, ICES quads are always 30 x 30 nautical mile squares (van Densen 2008), so using the Universal Transverse Mercator (UTM) will accomplish this. In light of this, the inclusion of the ICES quads should also be ensured, as they add visual clues for the users to locate themselves and also serve to enclose the data in a defined spatial unit. The scale should also be appropriate for the visualizations. We are somewhat limited in this, as displacement can occur if visual representations are too big for the area they represent, but at least the entirety of the North Sea should be viewable, as this is a study on that region. Taking care to not make things too small shall also be a focus, in order to make viewing (and hopefully understanding) easier.

Understanding will also be made more possible through the effective use of a legend and other visual explanations such as the presence of text denoting units (if necessary) and other information meant to set the data in context. In these visualizations, the depiction of time progression is central.

While it is certainly not preferable to include too much confusing ancillary explanation (Casselman 1999), we should make an attempt to provide the users enough information so that they can view the visualizations and not have any questions. This should also be done in a manner which transcends language barriers (Lohse 1997), especially since we will be dealing with stakeholders whose primary language is not English. The use of common abbreviations or mathematical symbols has been employed for these visualizations. Spatial contiguity of information which have a relation with each other (Harrower 2007) should also be a focus; for instance, placing the legend close to the items which necessitate it. In fact, the simpler the layout is, the better (Politis 2003).

Correct spatial representation must be assured for the purposes of accuracy of data and familiarity to users. Other visual clues about the data should also be made available.

Another important bit of explanatory material that must be included is the explanation of the time lapse. We have chosen to employ a text-based time depiction. That is, the date is displayed in the following format: "01/05/06" and advances by one month each time still. During preliminary testing, the use of a "time slide bar" was suggested. This seems to be an interesting idea and perhaps if it was dynamically integrated into the application, it would be well served. There was some literature (Plaisant 2005) however, which cautioned on the use of time sliders due to possible usability issues.

3.6.3 Animation construction

There are other aspects of visualization that do not come into play until the final stage of delivery to the stakeholders. This stage is the crafting of the visualizations into an animation, an idea first suggested by van Densen (2008). Having chosen to use ArcGIS to undertake this task, we must also note that when working with specific software, it is also necessary to effectively take into account and deal with its limitations. Such was the case in this project, as there were, at times, ideas which could not be integrated into the functionality of the software. Thus, an explanation of the features of the animation, as they are currently, and not ideally, is given below.

The ability to record an animation in ArcGIS is limited in that the options for the presentation of the animation are not spectacular. The aesthetics of the display window are lacking, if the inclusion of the legend and other layout properties is desired. Nonetheless, all the important information is included in the display, as can be seen in Appendix 2. It was difficult to fit in all of the features in a space-conscious manner, wanting, of course, for the focus to be on the visualization itself, which is as large as possible.

There is also very little interactivity allowed in the interface. As the animations are merely AVI movie files, the only interactivity is the ability to pause, play and loop the animation for multiple, consecutive views. While this simplistic automation is especially useful for extreme novices (Fairbairn et al. 2001), more experienced users may desire to manipulate the animation further.

Not directly affected by the limitations of the software is the choice of the length of the stills. More specifically stated; the length of play time that each month of data occupies in the animation can be manipulated to emphasize different properties of the data. Snapshots are, indeed good at showing obvious trends (Andrienko et al. 2005) in temporal data. As was discovered in our pre-trials, stills of a long duration show macro changes the best. In other words, from start to end, general patterns in the data can be easily identified. Short stills, however, allow users to discover changes from month to month more effectively. Both observations are useful, of course, so a compromise has to be made.

Careful attention should be paid to the map layout and interactivity of the visualization, which will be largely dependent on the chosen software. The length of the stills should also be considered according to the goals of the visualization.

The usability of these visualization / animations will be tested according to the plan set out in the next chapter. MacEachren et al (2001) and (Fairbairn et al. 2001) both suggest that it is not only interesting and useful to know how single users interact with visualizations but also how multiple users interact with each other and the visualization because this can tell the researcher a great deal about how visualizations can be used for negotiation and collaboration purposes, as this collaboration is not a goal commonly pursued (MacEachren 2005). While this could be a useful strategy in the future, the present focus will be on ensuring that we evaluate the understanding of users individually, in the hopes of eventually progressing to the strategy above.

3.7 Conclusion

While this chapter is clearly about the issues related to building the geovisualization, it should also be clearly understood that the choices made at this stage directly impact the usability of the visualization. Cognitive issues should be capitalized upon, in designing the visualization. We must know the users well, including their competencies and experience levels. We need to use the knowledge that they already possess in order to introduce new material to them (our visualization). We will use familiar "metaphors," principles of cartography and other characteristics related to form. In fact, the use of the visualization should enhance and increase their understanding of the data itself, thereby doing more than just representing the data graphically, but actually allowing knowledge discovery.

One of the most limiting factors of the visualization of the fishery data is the use of ArcGIS as the software for design and presentation of the visualization and animation. While the choices of cartographic representation (icons, colors, styles, etc.) are quite numerous, other elements of the presentation (time display, progression of stills) are less than optimal. In future iterations, it may be interesting to use another software package (or multiple ones) to gauge which features are most appreciated and why.

The manipulation of the data, choices for layout properties, as well as the construction of the animation will all be better suited to the users if we focus early on cognitive issues. Thus, the usability analysis, which is delineated in Chapter 4, will go much more smoothly.

4. <u>Usability analysis of created geovisualizations</u>

The central goal of any visualization design should be, of course, to "amplify cognition" (Plaisant 2005). Thus, the geovisualization techniques utilized in the construction of the aforementioned animations try to convey the information that they represent in an understandable and "usable" manner, and therefore they rely heavily on usability study theories.

It may seem obvious to the reader what usability is, but it has, in fact, many different facets which can be better understood after studying the term a bit further. "*Usability*," as defined by ISO (1998), is:

"…the extent to which a system can be used by specific users to achieve specified goals with effectiveness [the extent to which a goal is reached], efficiency [the effort to reach goals], and satisfaction [the user's opinion on system performance] in a specified context of use." (Fuhrmann et al. 2005)

In other words, there are several features to take into account when evaluating the usability of the animations; whether it meets its goals, whether the effort expended seems justified, and the user's opinion on the functioning of the animation. It is not a "yes" or "no" answer or something that can be determined easily by any evaluator, and is thus, as Dykes et al (2005) explain, a careful study of users' reactions to visualization possibilities. Ehn and Löwgren (1997) even suggest that usability is a subjective evaluation and will not be the same for everyone.

4.1 Background

Usability studies as a discipline, indeed have their roots in the social and cognitive sciences, including psychology (Fuhrmann et al. 2005). These disciplines have contributed much to the understanding geographers have about how people use their maps. In fact, in Geography, very little research has been done into specifically how users interact with geovisualizations such as interactive maps and other novel representations of spatial data (Harrower 2007). Harrower even calls geographers "inward-looking" because the same old literature is cited over and over again and very little new research is being done into why certain visualizations work (Harrower 2007). Furthermore, Fuhrmann et al (2005) tell us that "usability inspection methods" are not at all common in GIS, despite the fact that the use of this understanding will make geovisualizations all the more powerful in conveying messages and allowing knowledge discovery. It is puzzling, therefore, that behavioral research is not more commonly cited in usability studies (Landauer 1997).

There is certainly a lot to be offered from these disciplines. In addition to the overarching themes of the socio-cognitive sciences and psychology, interesting research has been done in the fields of Human-Computer Interaction (HCI), User-Centered Design (UCD) and Computer-Supported Cooperative Work (CSCW). Fuhrmann and Pike (2005) have done much work in the first of these two. They define HCI as the study of the design, implementation and evaluation of systems which seek to increase efficiency between humans and computers. UCD, which is a subfield of HCI tries to maximize efficiency of HCI methods by taking into account end user work practices, tasks, goals, etc. UCD borrows highly from usability engineering, and can be considered somewhat synonymous with human-factors engineering and software ergonomics (Fuhrmann and Pike 2005). UCD will be explained in more detail in the coming paragraphs. Lastly, CSCW, as Fuhrmann and Pike (2005) tell us, is a highly established field in which systems are evaluated from usually the designer's side and are not user-created. It is, however, a truly multidisciplinary field, drawing on the expertise in the fields of anthropology, sociology, organizational psychology and management science, among others (Olson and Olson 1997).

4.2 Methods

Nearly all literature from across varied fields, suggests that usability should be a priority at the very beginning of the development process (Edsall and Sidney 2005; Fuhrmann et al. 2005; Fuhrmann and Pike 2005; Plaisant 2005). Software developers tend to overestimate their own capabilities in discerning the usability aspects of their products. This is called the Egocentric Intuition Fallacy, and as described by Landauer (1997), can be a pitfall because usability design and analysis only comes at the end of the development process.

An attempt at subverting this practice is explained by Allen (1997), who differentiates between Mental Models and User Models. Mental Models are understood to be the users' ideas about how the software / hardware works, and User Models are the "expectations" that the software / hardware have about the users' capabilities. By trying this approach, it becomes obvious that the first thing we have to do is define the tasks we expect the users to undertake (context of use) and get to know more about the requirements of the users themselves. Fuhrmann et al (2005) suggest the framework below (Figure 13).

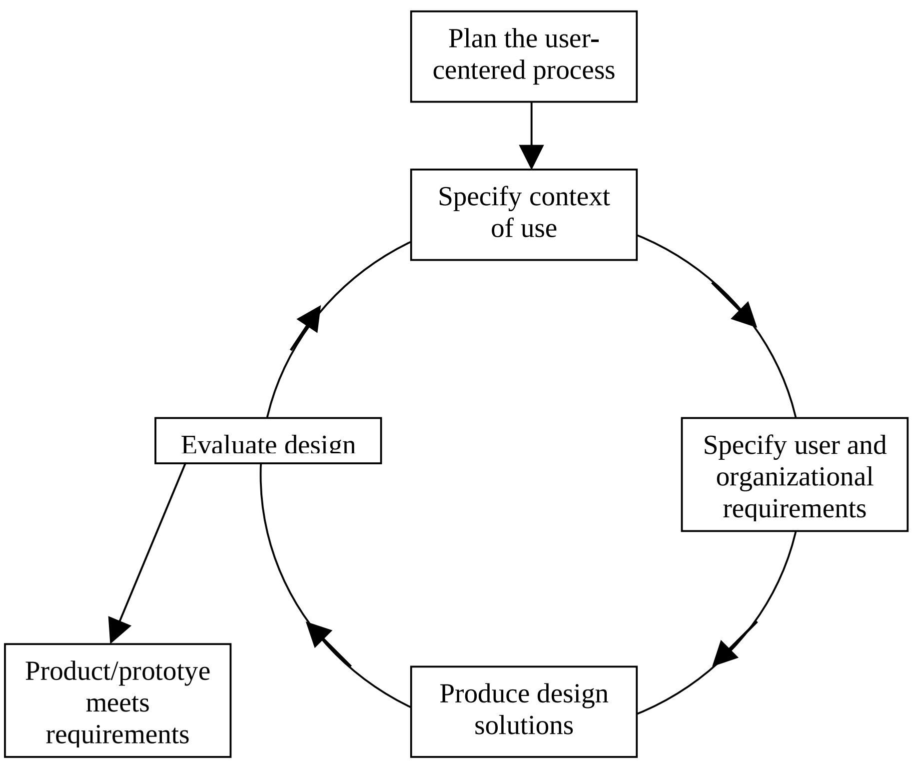


Figure 13: User-Centered Design Process (Fuhrmann et al. 2005)

Once we have figured out the user and organizational requirements, we can attempt to build a solution which responds to the needs of the users and goals set out in the delineation of the context of use. Fuhrmann and Pike (2005) list five “tenets” which should be a priority in the design process (Figure 14). It should come as no surprise that all five of these Tenets address the question of usability.

<u>Tenets of User-Centered Design</u>
1. Easy to Learn
2. Efficient to Use
3. Easy to Remember
4. Limit Users’ Errors
5. Pleasant to Use

Figure 14: Tenets of User-Centered Design (Fuhrmann and Pike 2005)

As may be apparent from these Tenets, the main bottleneck for geovisualization is not technological advancement, but cognition of the information being presented (Harrower 2007). Some argue that visualizations should be made for lay people, as opposed to experts (Dykes et al. 2005) because learning results from an interaction between new visual stimuli and previously learned knowledge (MacEachren 1995). Still, others have shown that there is not always a gap in knowledge between perceived experts and novices (Harrower 2007). Nonetheless, the process of understanding visualizations, particularly animated ones should be further delved into, as well as the many hindrances to learning.

Harrower (2007) writes about cognition of animations, highlighting four different limitations that users encounter: cognitive overload, retroactive inhibition, split attention and the modality principle (which will not be discussed). Cognitive overload occurs when the working memory (WM), which retains limited amounts of information for a matter of only seconds, becomes saturated. This state of being confused or misled has also been called “toxicity of overload” by Atkinson and Mayer (2002). The interplay between the WM and long term memory (LTM) must be improved in order for learning to occur. The ideal state is the Germane Cognitive Load, which can be achieved by reducing extraneous cognitive load, such as poor organization of graphs, visual or other distractions, etc (Harrower 2007).

Confusion is especially likely during the viewing of an animation because the information from one still of the animation must quickly be transferred from the WM to the LTM in order for the next still to be processed, as it appears. The failure to accomplish this action is known as retroactive inhibition. Harrower (2007) suggests that perhaps a looping of the animation can be applied in order that each still is seen multiple times to increase comprehension or a “pre-training” period could be added where the animations are viewed ahead of time or some kind of preview of the information can be provided. Other strategies that can be used to avoid confusion are user control measures, which accomplish more than just passive watching, by adding an interactive component to the animation so details can be absorbed by the LTM.

Another limitation on cognition is the split attention effect which can be either of the genre of spatial or temporal contiguity (Harrower 2007). Data in an animation may be either too far from one another in space, or too spread out in terms of time elapsed. For example, the legend may require as much attention as the animation itself, or there may be too many pauses to make any connection between stills. It is recommended that labels directly on the animation take the place of legends, if at all possible (Harrower 2007).

To summarize, in order for cognition to be maximized, it must "play off", that is, interact with previously held knowledge. Additionally, the visualized information must be able to be digested quickly or shown repetitively and it must be obvious upon first glance with very few necessary deductions being made. If these recommendations are stuck to, confusion should be negligible.

Other specific recommendations abound, taking into account the recent paradigm shift towards User-Centered Design, which has attempted to address the problems associated with understanding and knowledge discovery, not as expected by the developer but as directly informed from user – developer interaction (Edsall and Sidney 2005). Edsall and Sidney (2005) suggest limiting the animation to a portion of the data, as it is usually multifaceted and includes many different variables which could each be visualized separately. Incorporation of interactivity is also encouraged by numerous sources, for the reason that interactivity allows each user to gain a better understanding; in fact, as we have already seen, his own, personalized understanding of the information. These multiple perspectives help to better gauge usability, and as Virzi (1997) explains, even if the usability evaluation is limited in time, multiple evaluators are more useful than a single one for a longer period. Interactivity can be as basic as allowing pausing, rewind or looping options for animations and incorporates the ideas of cognitive science already mentioned in this chapter (Edsall and Sidney 2005).

Cognition can be supported through focusing the user on small bits of information, and interactivity.

Although we have been talking about usability studies which start at the very beginning of the development process, attempting to anticipate usability issues early, it should be understood that this is not the only technique used. Another strategy for ameliorating software (and indeed, geovisualizations) suggests that it is the number of times an evaluation is undertaken, and not necessarily at which phase of the development process that the intervention occurs, is the most important. Gould et al (1997), in their explanation of the design process, extol the use of prototypes as a means to judge usability. This strategy is called Iterative Testing or Formative Evaluation (Landauer 1997). Fuhrmann and Pike (2005) call this process, when testing different visualization options against each other at the end, Summative Design.

Even after the visualizations have been created, the goal of improving communication and understanding between stakeholders has not been accomplished. The GIS must be delivered in a manner which takes into account the competencies and specifications of many stakeholders. To this end, a delivery framework must be designed. There are two concerns which arise when attempting to design this framework: what will its capabilities be? and in which environment will it be delivered?

In order to know exactly which parts of the visualization are problematic to users, a framework has to be set up to collect their concerns and suggestions. The Delphi Method can be especially helpful in accomplishing this goal. The Delphi Method acknowledges that many times, convening in the same place all the stakeholders of a certain problem is impossible. This method therefore consists of structured, asynchronous communications between experts, users and developers, attempting to reach goals, but not necessarily though consensus (Fuhrmann and Pike 2005). Furthermore, anonymity is encouraged so that users feel free to make any suggestions without embarrassment or hindrance. In short, the Delphi Method is an iterative design process which encourages discussion, and synthesis and refinement of ideas (Fuhrmann and Pike 2005).

It appears that the important factor in increasing usability is making multiple usability tests and improving at each stage. The usability test, itself, must be usable also though and allow the users to evaluate the usability of the visualizations most effectively.

Through the framework outlined in this section, users may be able to do a sort of validation of the model. That is, they will be able to judge, based on criteria to be set out in the next section, how the animation performs in comparison with the traditional static maps. Users will be able to express their expectations and say whether the models seem to be representative of the reality that they have experienced at sea. Perhaps the animation or static map will be overly complex and there will be too many variables present. Perhaps the interaction level of the animation will not be sufficient for showing the users what the relationship between the variables is. Perhaps the cartographic techniques used in creating the visualizations will not be effective at communicating the data. If any of these are the case, users' suggestions will be noted and the model could be augmented to include the suggestions.

4.3 Evaluation Criteria

Regardless of the time the intervention of usability testing or engineering takes place, there are certain criteria that are tested and methods to best determine the ease which users will interact with the product. Gould et al (1997) offer a list of the aspects related to usability that can perhaps be of some help in understanding the complexity of the evaluation / determination. Granted, this list is directed more towards software development, but it is nonetheless, useful for us to evaluate geovisualizations produced for users of varied backgrounds. Evaluation criteria include: system performance (reliable and responsive), system functions, user interfaces, reading materials, language translation, outreach / support programs, modifiable / extensible, installation, field maintenance / serviceability, and advertising / promotion (Gould et al. 1997).

This is a rather extensive and all-encompassing list which may be too general for our digestion; therefore, let us discuss some more specific examples of usability evaluation criteria. Ehn and Löwgren (1997) present their "Measurements of Usability" in a four point structure. They talk of user performance on specific tasks (evaluated by the time it takes users to complete the task, the completion percentage and the number of errors), learnability (evaluated through the same means, as well as information retention), design flexibility (whether a heterogeneous group can succeed), and subjective satisfaction.

Ehn and Löwgren (1997) also touch on two other means of evaluating usability in systems: Quality in Use, and Quality Perspective. The first one bases its evaluation on structure (materials and technology used), function (the contextual purpose), and form (the experience of the user). While the first two criteria are objective, the third one is subjective and may therefore offer the most insight. The Quality Perspective is similar to the aforementioned evaluation, in that it offers objective and subjective criteria for the user to base the assessment on. They are: Constructive Quality, or correctness of design (whether the system is maintainable and robust enough); Ethical (whether all users will be treated fairly or equally); and Aesthetical (whether the setup is appropriate, and the interface and interaction with the system are pleasing to the user) (Ehn and Löwgren 1997).

For judging the degree to which usability is attained, we can look to the time it takes users to complete tasks, their proficiency and their efficiency. We can also evaluate the

visualization on whether a varied group of users are accommodated and what features of the visualization are most appreciated.

While the preceding paragraphs explained a bit about criteria to be considered in a usability evaluation, the proceeding one will explain, according to Virzi (1997), the Usability Inspection Methods utilized by evaluators. While these methods are intended for experts, in some cases, to minimize costs, non-experts have been used with effectiveness. For example, in Resource-Constrained Methods such as heuristics analysis, "design principles" (a kind of guideline) are followed by evaluators and then necessarily compiled at the end of the evaluation to produce a (hopefully) definitive ruling on the product's usability. In Usability Expert Reviews, verified experts carry out the usability evaluation using their own methods and their rulings are aggregated at the end to form a single consensus. Group Design Reviews encompass specialists from many fields related to the product who evaluate it as a group and reach a common consensus together. Finally, Cognitive Walkthroughs can be performed for a specific task and only that task, and include on-the-fly learning, a sort of ease-of-use test (Virzi 1997).

Since none of these methods are tailor-made for the evaluation of the usability of fisheries geovisualization animations, it is necessary to craft something specifically for this application. The evaluation method used here will draw heavily upon the literature discussed previously and fill in the gaps, as they arise, to address our own needs. Indeed, there is "no comprehensive and generally accepted model on how to design good human factors into computer systems" according to Gould et al (1997).

4.4 Fishery Animation User Test Methodology

This section will explain how the theories above relate to the current research project on the subject of fisheries data animation. After setting the above literature into context, the methodology envisioned for the usability analysis and evaluation portion of this project will be laid out, including highlights of the stakeholder meeting and usability analysis format.

As has been explained previously, this research project envisions three main stakeholder groups: scientists, policy makers and fishermen. As should also be clear by now, the main goal of the creation of the visualizations, discussed in Chapter 3, is to increase understanding between the stakeholder groups by attempting to "put them on the same page"; that is, to ensure that they have the same understanding of the problem under study.

As the previous chapters have shown, the communication of spatial data is extremely cumbersome, especially when it is not just communication for the sake of conveying an understanding of some data but for actually facilitating new knowledge discovery by the users of the visualization. Therefore, the methodology about how to best go about this task of creating an animated visualization was given much thought and has been based on literature and other examples exhibiting the desired characteristics.

It seems most appropriate, therefore, to begin with an established framework and build off it, as we see fit. According to Fuhrmann et al (2005)'s ***User-Centered Design Process*** (Figure 13), the first step is to specify a ***context of use*** for the animation. This task was largely already completed when the project began because of the work of van Densen (2007), and was further enumerated during the initial thesis meeting (van Densen 2008). ***User and organizational requirements*** were the next task to be tackled, and were also somewhat already determined by van Densen (2008) from the outset. Still though, the development of the understanding of the users and their requirements was undertaken largely through reading related literature and similar studies, as well as through conversations with van Densen and other fishery experts.

After what was deemed a sufficient assessment of user needs was completed, ***design solutions*** for the product were considered and pursued. As much as possible, given limitations and requirements, the integration of the ***Tenets of UCD*** (Fuhrmann and Pike 2005) (Figure 14) was followed. The Tenets include: ease of learning, efficiency of use, ease of remembering, limit user errors and pleasant use. By designing the visualization for novices, and not trained experts in geovisualization, the Tenets seemed to be well addressed. In addition, this design strategy also seemed to limit the chances of cognitive overload (Harrower 2007) which would have led to misunderstanding and confusion among users.

As Harrower (2007) also suggests, interactivity such as pause, rewind and looping functions, as well as a pre-training period have been incorporated to allow users more control of the animation, and hopefully provide them with a better opportunity to digest its meaning. To combat ***split attention*** (Harrower 2007), every month of the two-year span is visualized, while spatially-speaking, all elements are close together in the display screen. The data that is visualized is filtered by mesh size, as limiting animated data is extolled by Edsall and Sidney (2005) as a useful technique to combat confusion. (See Appendix 2 for the visualization layout).

Although Landauer (1997)'s Iterative Design technique is not envisaged for this stage of animation development, the use of ***prototypes*** (Gould et al. 1997) has been largely successful in more precisely determining which animations would be used in the evaluation by the stakeholders. Lastly, as a means of evaluation, it has been assumed from the beginning that several different visualization techniques would be tested against each other in order to determine which one is preferred by the users. This strategy, which has been elaborated by Fuhrmann and Pike (2005) is known as Summative Design.

The last phase of the UCD Process is the ***design evaluation***. There are three different relevant evaluation strategies from the literature: the Measurements of Usability and Quality in Use both from Ehn and Löwgren (1997), and Resource-Constrained Methods, explained by Virzi (1997). There are four facets to the Measurements of Usability which are particularly interesting for our evaluation; user performance, learnability, design flexibility and satisfaction. ***User performance*** is the ultimate objective evaluation criteria, as we can give users a grade of pass or fail. ***Learnability*** is concerned with information retention, asking users to evaluate the animation based on their cognition of its contents. This is the central question we have set out to answer. ***Design flexibility*** takes into account whether the animation was effective for a heterogeneous group of people, such as our expected audience. The ***satisfaction*** criterion is, of course, subjective, and will allow evaluators to give us their opinions and raise their own questions and comments, and for this reason, may be the most interesting of the criteria for our application, especially if it fails in its intended goal, they will be able to hopefully fill in some of the missing links.

Ehn and Löwgren (1997)'s other evaluation strategy, Quality in Use, is also three-tiered. It consists of the following categories of criteria: structure, function and form. ***Structure***, the materials and technologies used in the construction of the animated visualization, is interesting because we will be able to find out if our methods and information delivery have hit the mark or not. ***Function***, which is the contextual purpose of the application, will show the users' opinions of one of the very first determinations which was made into the methodology of the project, whether the attempt to ameliorate communication and understanding among stakeholders was worthwhile. Finally, evaluating the animation based on its ***form*** is also a subjective criterion because it gives the users the chance to comment on how the visualization worked for them personally in its current incarnation.

The last evaluation strategy is not as structured, but it could be if so chosen. It all depends on how the evaluators answer the guidelines, or in this case, questionnaire, provided as part of the ***Resource-Constrained Method***, also known as heuristics analysis. In this method, evaluators, usually experts, but not necessarily, evaluate the visualization based on design principles which will be set forth in the questionnaire accompanying each visualization. Their opinions need to be compiled at the end of the evaluation period, as is intended for all of these strategies, in order to make a definitive conclusion about the visualization's usability.

4.5 Meeting and questionnaire setup

The goals of the meeting and analysis (via questionnaire) were two-fold: to determine if the general setup of the visualizations was sufficient for user understanding of the phenomena under study, and to determine the usability of the visualizations, given the comparison with the static maps.

The evaluation analysis took place during a meeting between stakeholders. The meeting place was a small room with multimedia capabilities, specifically a projector and PC workstation. There was also a waiting area for attendees. The opening of this meeting included an introduction period, so the participants had a better idea of what they would be doing. An explanation of the agenda was offered. Careful attention was paid, however, to not say too much and influence the results of the usability test.

The testing commenced with only one of the participants present, as to not bias our results by allowing the other to have a head start on things. The first part of the questionnaire was the Personal section (Appendix 3). Here we ask general questions about the participant's experience in different subject areas. This should give us an idea of what we are dealing with in terms of knowledge and capacity for acting as a representative sample of the stakeholder population. It lasted five minutes.

The second section includes the questionnaires for both static and animated maps (Appendices 4 and 5). The questionnaires consisted of two sorts of questions: objective and subjective. Objective questions have a right or wrong answer. Opinions or personal feelings will not affect the chosen answer. Most objective questions are multiple choice, closed questions, but there are some open answer, fill in the blanks. The goal of objective questions is to decide whether or not participants are able understand and make conclusions from the data, as it is presented in the visualizations. Subjective questions allow the users to add their own personal feelings about the data and visualizations. Thus, they are completely open-ended and the users can say as much or as little as they wish. Subjective questions allow us to know how the users felt about the visualizations and process of usability testing, in their own words. What we may have missed in the objective questions, should be answerable during the subjective ones.

First, the static maps were presented in paper format on a table in front of the participant (See example in Appendix 2). Questions were presented one-by-one on a paper, one per page. The questions were read directly from a pre-prepared questionnaire (which is elaborated upon in the next section). Some questions are standard for every stakeholder; while there is also the opportunity to personalize them, if the stakeholder has something to add. Each question has its own unique time allotment denoted in the last section of this document. We went through the questions individually and after completed, we switched over to using the animation.

The animation was shown on a computer which the participant had control over, so he could scroll back and forth through it at his leisure. The questions for the animation were presented in the same manner as for the static maps. The participant was able to play the animation after each question was asked.

After the questionnaire for each map, there was a short wrap-up session to tie together any loose ends and allow the participant to add anything he would like. Then, the next participant entered and the process started again from the beginning.

Following the stakeholder meeting, the evaluations were analyzed. Particular attention was given to whether scientists and fishermen had similar answers or comments because this information will offer the best evidence of whether they possess the same understanding of the visualizations.

5. Results of usability tests

The meeting with several stakeholders took place on Tuesday, March 31, 2009. The stakeholders were two renowned Dutch fishery scientists from different institutes. It was decided that the initial stakeholders meeting should be held with a smaller group in order to judge whether it could be useful for the larger population at a later time. The two fishery scientists who were present for the usability tests were the most available of possible participants, so they were included in the first run.

Since the sample size is so small and both individuals come from the same class of stakeholders (fisheries scientists), we will try to retain some anonymity when discussing results. While there were some notable differences in the results per participant, we will mainly be focusing on the similarities in order to get a better idea of how fisheries scientists digest the information presented in the visualizations.

On the questionnaires, there were two types of questions: objective and subjective. Objective question results were analyzed based on whether they were correct or not, the time taken to complete them, and a comparison between the static, paper maps and the animated one. Subjective question results took more time for analysis, as sometimes participants' answers were difficult to aggregate. The answers to both objective and subjective questions are loosely grouped according to the type / complexity of question. Thus, they will be presented roughly in numerical order.

5.1 Answers to objective questions

For the following sections dealing with questions on the questionnaire, please refer to Appendix 4 (statics) and Appendix 5 (animations). Additionally, the participants' answers, along with analysis of these answers can be found in Appendices 6 and 7.

For the first seven questions, participants seemed quite able to digest the basic information shown on the interface of the visualizations; both static and animated, as they are the same in face value. These questions dealt with basic map reading capabilities. Questions 6 and 7 required the users to view the entire animation or the first and last static map. The first five questions could be entirely answered using the first month or any other visualization. The first seven questions (for both static and animated) were easily answered by both participants, almost always under the targeted time for each question. With 28 questions given (7 per person per visualization), there was only one wrong answer. Participants were able to answer the questions using the animations more quickly than with the static maps, generally around ten seconds faster.

Questions 8, 9, and 10 are tuned to determine if users can make simple observations about which month(s) something occurs, if given a geographic location. This kind of question is something that would be central to the use of the visualization if it would be used for reference by the stakeholders in their day-to-day activities. For example, if a fisherman were to consult the visualization to see in which months effort was highest in the southern quads. Users were highly in favor of the static maps for this type of question, as the times for completion using the animated map were sometimes triple those for the statics, and thus, well over the recommended target time. The number of questions wrong was virtually identical however (3 vs. 2, in favor of animated). For 3 of the 5 wrong answers, the problem lay in the fact that questions 8 and 10 were multipart and were only partially answered correctly.

Specifically, the answer should have included several months, and participants left one or more out of their answers.

Question 11 seemed to give the participants a lot of trouble, while it was intended to be one of the easier questions of the test. The participants understood it to ask for them to track the changes from one month to the same month in the next year, while it was intended to just compare the two time periods. Upon further reflection, it could have been misunderstood due to the nature of the language used; therefore the results from it will be disregarded.

Questions 12 and 15 were of the same type as 8,9,10, and very little difficulty was encountered by the participants in answering these correctly within the target time when using the static maps. Again, the users were asked to answer with a month, given a geographic location which something could occur. The animated maps were not as useful, according to the time it took to answer these questions using the animations, but the correct answers were attained.

For questions 13 and 14, users were asked to do some rough mathematical estimates, determining and cumulating amounts of effort over several months. This was not an easy task. There was a large difference in the amount of time the users took for these questions. Curiously for question 13, the animation was much more useful in terms of time, while statics were highly preferred for question 14.

The static map was marginally preferred for question 18, in that it took slightly less time to answer it correctly. In this question, users are given a geographic area and asked to do some basic estimating of effort.

Questions 19-21 are impossible to judge whether the animation or static maps were better, as the participants had contradicting time results for each. Question 19 gave location and time and asked participants to choose from a list of phenomena which were represented in that place and time. Questions 20 and 12 were again, nothing new in terms of content, merely asking for a month when something occurs in a known geographic location. We can conclude though that there seemed to almost never be enough time to answer the questions within the target. Question 20 had a wrong answer for the animation, while question 21 had one for the statics.

5.2 Answers to subjective questions

Subjective answers are also quite interesting. They really allow the users to give their opinions on the usability issues related to the animation and static visualizations that were produced. We will go through the subjective questions one-by-one and find similarities in the answers of the participants. Static and animated visualizations will be compared.

Question 16 focused on the MapTable exercise.

Answers to question 17 show that static maps are appreciated because it is possible to view all the months simultaneously, allowing easy side-by-side comparison, while the animations do provide some benefit in visualizing the movement of the changing effort.

Answers to question 22 conclude that statics show somewhat well the effort distribution and quantity, while animations are particularly good if you want to focus on one ICES quadrant throughout the year.

Question 23 is a multipart question. When checking the intensity of effort for one ICES quad in one specific month, participants concluded that static maps were the preferred medium. For comparing effort distribution in consecutive months, the animation is specifically well-tailored to this task. Comparing effort distribution for the same months in different years (or more generally; non-consecutive months), animations are particularly bad for this, while statics are considerably better. Lastly, animations are strongly preferred for tracking changes in effort of a single ICES quad throughout the year.

General usage of static/ paper versus animated maps was dealt with in question 24. Users both believed paper maps were more useful. It was noted that each format has its pros and cons though, for instance; paper maps could take longer to use than the animated one.

Question 25 asked for elements of the visualizations that should be improved. There were several categories of improvements for both map formats. Participants, for the most part, had the same suggestions for improvement. For the static maps, satisfaction was indeed very high. Two notable comments were that the legend quantities were probably not necessary, as they were never or very seldom used, and that the numerical date (particularly the day) were troublesome and should be changed in favor of the text name. Animated maps garnered more criticism. Although it was due to technical restrictions rather than design ones, the length of the stills of the animation drew negative comments from the participants. They thought that each still should be shorter than the current 2-3 seconds. Participants also expressed a desire for more interactivity in playing the animations. Particularly, they asked for the possibility of immediately being able to “zoom” to a specific month. Although less problematic in the eyes of the users, the choice of ICES quad projection did come under scrutiny.

The desire for a slide bar for changing the month displayed in an animation was also expressed in question 26. For static visualizations, users said that, ideally, a seasonal or bi-monthly map would be available (somewhere between 8 – 12 stills). Except for questions which specifically ask about a particular month, this number of stills would suffice to show the yearly spatial variation of effort.

Lastly, question 27 allowed the users to choose which visualization they preferred. They both acknowledged that static and animated maps each had their own place and would, in some situations, be more useful than the other.

5.3 Conclusion

This section has presented the “raw data” results from the usability testing performed during the stakeholder meeting, by means of two questionnaires. For both the static and animated visualizations, two different types of question were developed: objective ones and subjective. While it is more often than not possible to find useful results, occasionally the responses of the stakeholders do not agree with each other. It is clear, however, that certain kinds of questions are more easily answered by either the static or animated maps. Otherwise, some useful suggestions have been received, which will be further detailed in the following chapter.

6. <u>Discussion, conclusions, and recommendations</u>

The results highlighted above will, in this chapter, be broken down and analyzed further in order to see if they support or hinder our original goal and research objectives. First, though, it is useful to restate the intended objectives of this research and refocus on the ultimate product.

6.1 Recap

Our original goal of making a visualization which was accepted by all the different stakeholder groups (fishermen, scientists and policy makers) is still central to our study. With a multi-dimensional stakeholder group, we also want to take into account the varied backgrounds and levels of familiarity with map displays and information visualization. Designing for novices is often preferable to designing for real experts (Dykes et al. 2005), so it is ensured that everyone will be able to participate and the usability will be maximized.

Several descriptions (Fuhrmann et al. 2005) of “usability” focus on three necessary criteria: meeting the goals, justifying the effort put in and having good user opinions. In this project, the last of these criteria is most important, as, in fact, our main goal is to gain good reviews from our users. Another strategy for evaluating the usability of our visualization is the “Measurements of Usability” of Ehn and Löwgren (1997). This strategy has four criteria: user performance, subjective satisfaction, learnability and design flexibility. User performance is evaluated based on the time it took to complete the exercise, the completion percentage and the number of errors. Subjective satisfaction is how the users responded to the exercise. Learnability relies on the previous methods, as well as information retention. Design flexibility evaluates the exercise based on whether it accommodates a heterogeneous group. Of all the evaluation strategies reviewed in Chapter 4 of this report, the “Measurements of Usability” fits our needs the best, as it bases its assessment on categories which are also relevant to this research. Thus, the discussion section will go about evaluating the results of the usability questionnaires based on this strategy.

We do acknowledge that usability is subjective however (Ehn and Löwgren 1997) and that our sample size, at the moment, is very small, but it is hoped that we can still extract some useful conclusions from our results, which were presented in the previous chapter. Furthermore, we do intend for this research to be elaborated upon and extended to the other two stakeholder groups (fishermen and policy makers) because as Virzi (1997) explains, having multiple short usability evaluations is much more useful than having only a few long-term ones. It is hoped that through this process of constantly improving the animations according to suggestions gathered from successive usability tests, that we may arrive at a product which satisfies all of the needs of the stakeholders. This process has been called “Iterative Testing” by Landauer (1997).

6.2 User performance

Based on the criteria used to evaluate user performance; time taken, number of errors and question completion percentage were considered. The evaluation of the questionnaire answers has been largely carried out with the time it took to complete the question as the main criteria in judging whether the static or animated maps were preferred by the users. There were not many wrong answers at all, so number of errors may not be definitive. Completion percentage is also not very useful because users were often over the target time for each question. It was determined, in fact, during the participant tests that they would not be cut-off from questions when the target time passed. Upon noticing their progress, it was deemed that if this practice were employed, there would be very few results at all, and to actually have answers, even if they were achieved in longer times, would be more useful. This, then accounts for the reliance on completion times for evaluation, and not on number of errors or completion percentage.

Upon evaluating the results of the questionnaires, we are able to make some conclusions about the users' abilities to answer questions based on the two visualization media (static maps and animations). The following conclusions will be explained in their corresponding subsections.

- 6.2.1 Not having to rely on one's own memory was an advantage for using static maps.

- 6.2.2 Questions which could have been answered equally as easily with either the static or animated visualizations were almost always answered more quickly with the animations.

- 6.2.3 Questions which should be easier to answer based on animations were always answered faster with statics.

- 6.2.4 Those questions which we expected to be more easily answered with statics were half of the time faster with animated visualizations. This, along with 6.2.2 and 6.2.3 lead us to believe that there was some sort of bias involved, skewing our results.

- 6.2.5 Some objective questions did not allow us to make conclusions, so they should be further refined or proposed to more test subjects.

6.2.1 Non-reliance on memory

Although we hypothesized that the animation would be useful in conveying some information to stakeholders, there also appear to be some scenarios where the static visualizations are more useful. Such is particularly the case when asked a question in which the user has to remember something from one month and continue viewing the other months.

This can be explained through the work of Harrower (2007), in his explanation of retroactive inhibition and the split attention effect. Retroactive inhibition affects the transfer of information from the Working Memory to the Long Term Memory. In other words, if information has to be quickly processed and stored in memory, it may not be remembered correctly. Split attention can occur when the temporal contiguity of information is not close enough. In this context, if remembering the information from an early still and then being asked to recall it at the end of the animation, a user may have difficulty because of the amount of time that has elapsed.

The remembering that is required takes three different forms in answering the questionnaires: *comparing, consulting* and *computing*. Comparing effort quantities between months is necessary in Question 18. Visually comparing the coast of Denmark in each month was easier using the statics. This is expected, since the users could view all months at the same time and determine which one had the most effort without having to rely on memory.

Consulting other map elements, such as the legend, scale, title, et cetera also necessitates the use of memory, especially as stills are passing and the eye is diverted from the data. The answers to Question 12 conform to this expectation. Visually comparing the size of the graduated circle denoting effort intensity was logically easier using the static maps because the circle did not disappear quickly, allowing the users to compare with the legend for more than the duration of one still of the animation.

Lastly, computing is featured in several questions. Question 14 asks users to roughly aggregate the effort in every single month and then compare the months. Expectedly, it took them less time using the static maps. This way, users did not have to remember the data month-to-month; they could compare it simultaneously with other months' data.

6.2.2 Equal opportunity

There are quite a few questions on the questionnaires which can be answered using either the animation or the static visualizations and using neither one will provide a competitive advantage. That is because these questions deal with elements which are featured on the very first visualization month: 01/01/05. Either looking at the first static visualization which catches the eye, or glancing upon the "default page" (01/01/05) of the animation, should yield the same result.

The first five objective questions on the questionnaire are identical for both static and animated maps. All the information that is necessary to answer these questions is available on the first still (or static paper). Accordingly, we would expect that the answer would have taken the same amount of time for the static question and the animation one. In reality, the animated ones (which were answered second) were nearly always finished in a shorter amount of time and without error by both participants. Question 20 could have been answered just as easily using either visualization formats and that is exactly what happened with the users. One chose animation, the other: static. This was the same question on both questionnaires but asked in a different way.

6.2.3 Expectation: animation

There were some questions which we expected animations to be better suited in answering. Indeed, this is the point of this research. It seemed that there should be some times in which animations can better show data in an understandable manner than static maps could. Our expectations were not supported, however. The animation was the quicker media on several occasions but there was never a question where the animation was more useful (in terms of time taken to answer the question) when we expected it to be.

Question 8 was answered overwhelmingly faster during the static questions, which was surprising. Using the animation should have been easier to note months that a specific ICES quad were free from effort. All it would have taken was to locate the quad, press play on the animation and keep a keen eye to the quad. Using the static maps should have been much more difficult, as the user would have had to locate the quad and physically move his focus to the next month at every paper.

The same can be said for the next question, number 9, which required the users to locate the most geographically extreme occurrence of effort and note the month that it occurred. The animation should have been slightly more useful for this question, but not to the extent of the previous question. In fact, the static was slightly faster. Again, question 10 should have been answerable in slightly less time using the animation, as this question was of the same type as previously. The result was the same, except this time answering using the static maps was overwhelmingly faster, in one case, taking three times shorter.

6.2.4 Bias

As acknowledged in the first conclusion, there are some times when static maps are better suited to answering questions of a specific type. We did notice this on many occasions when we did not expect it but also on some occasions when we did. These were highlighted in the first subsection relating to memory.

We also observed times when statics should have been faster but they were not. These instances are particularly interesting, given the users' apparent preference for the statics in early conclusions. For question number 6, the more common answer was not expected. In fact, for this question, which asked for the timeline of the data, it was expected that the static answers would be much quicker than the animated, since it would have been necessary to view the entire animation (more than one minute in practice, but closer to three in reality), while the answer could have been available almost immediately using the static maps since all stills were laid out on the table and could be seen from beginning to end.

Again, in Question 13, the faster performing visualization was the animation, overwhelmingly. This question asks users to compare effort quantities and distribution for three separate months for both years. This would require them to scan, estimate and remember every single month, much as they were asked to do in the first subsection on memory. It is quite impressive that they were able to do these tasks more effectively with the animation, which changes rather quickly.

The rapidity of the animation stills should also have hindered the users from answering Question 21. Surprisingly, one of the users was able to use this visualization type to correctly answer. Given a specific date and location, users are asked to say what is remarkable about that occurrence of effort. Since no playing of and waiting for the animation would be necessary, as well as the fact that the map would not change, the obvious answer would be for the static map to be more accommodating to users.

Based on our observations of many answers, it seems likely that the users were unfairly influenced (biased) by some of the questions, particularly the ones which appeared on questionnaires for both visualization types. Key among these examples are questions 1-5. As mentioned, we expected that these questions could have been just as easily answered using the statics as the animation, but in fact, animations were nearly always faster. This probably suggests that the users recognized the question and correct answer from the first round of questions and were able to provide the solution more quickly. Question 20 also seems to have been answered in a biased manner for one of the users.

Perhaps even more striking or evident of prior knowledge of the questions was seen in Question 6, where it would have been far more convenient to use the statics to answer the question. This was the question dealing with the timeline of the data. As explained before, it would have taken approximately 150-180 seconds to answer this question using the animation because of the technical slowdown of it. In reality, the users answered the question in an average of 7.5 seconds (Appendix 7).

6.2.5 Inconclusive

Unfortunately, we received many answers which did not allow us to make any firm conclusions. These answers to objective questions had a split preference; that is to say, one user arrived at a faster answer using the animation and the other user with the statics. This was the case for six questions (Appendix 7). Additionally, there was one question which users were clearly confused by and thus it was voided.

Some of these questions seemed obvious initially, but evidently there was misunderstanding. For example, Question 15 should have been easily answered if using the animation, as the "jump" North is quite evident when the data is dynamic. Question 7 also seemed to be relatively easy using the animation, but one user preferred the static maps. Static maps would have been well-suited to answering Question 19 but there was again some disagreement among users which media type was better.

Any of these questions perhaps were worded poorly or otherwise were not clearly presented to the users. There is also a chance that because users were so few in number, we were unfortunate enough to just have several users who misinterpreted the questions and we would have received more conclusive answers had there been more participants in our initial usability study. Whichever is the case, it is necessary to go back and try to sort out what happened and try again with more participants.

Based on the answers to the objective questions, it is difficult to make many firm conclusions. Based on total time, the statics were more useful in answering the questionnaire (Appendix 7). Based on completion percentage (number of questions which were answered under target time), the animations were more useful. The number of errors was greater for static maps. Many of the questions that we expected to be better answered by animations were, in fact, better answered by statics and vice-versa. There are six questions which were answered faster using the animation and six for the static maps, while there are seven which we did not come to a consensus on which format was better. It seems we need to go further and take into account the answers to the subjective questions.

6.3 Subjective satisfaction

Since users' performances were not a conclusive measurement of the usability of the visualizations, their answers to the open-ended, subjective questions will now be considered. Using these criteria is useful because it gives the users a chance to tell us, in their own words, how they thought the visualizations performed and they are not limited to the answer choices provided in the objective section of the questionnaires. It also gives them a chance to not only explain *whether* they could answer questions correctly using the two visualization formats, but also *why* they were able to.

All of the subjective questions asked the users to rate the given statement from 1-5 in terms of their agreement with it, with 5 being the highest level of agreement. Unfortunately, it was not very often that we received answers in this format. Nonetheless, we were able to come up with some conclusions. They will be laid out in the same manner as for the objective conclusions.

6.3.1 There were some cartographic elements which the users particularly found useful.

6.3.2 Visualizations also included elements or other aspects which could have been further improved for the better functioning or understanding of the participants.

6.3.3 Users identified both visualization types as serving some use, but each was more competent than the other in answering specific kinds of questions.

6.3.4 Animations, as they current are, cannot replace the statics.

6.3.5 Time may not be a good indicator of usability.

6.3.6 The interactivity of the animations must be improved in order for them to perform on the same level of usability as the static maps.

6.3.7 The number of still necessary in the animations can be reduced.

6.3.1 Appreciated cartographic elements

While the data may have been somewhat surprising, the manner in which the data was represented and displayed was largely appreciated by the users. On question 25, it was noted that the cartographic elements seemed to be appropriate for the data being presented. Notably, there was appreciation for the placement of elements on screen, and size and colors of elements. Additionally, users noted that they never even once consulted the numbers on the legend, so we may assume that this meant that the choice of graduated circles to represent quantity was a good one. It was an intuitive and logical choice if the users were immediately conscious that the larger circles meant that the intensity of effort was higher (Larkin and Simon 1987). Thus, to avoid including unnecessary information in our visualization, which could lead to misunderstanding (Casselman 1999), we will eliminate the numbers in future iterations.

6.3.2 Needed improvements

Having just mentioned the legend numbers not being utilized, we can now highlight some other recommendations that the users had for the improvement of the cartographic elements of the visualizations. Notably, in Question 25, users expressed a mild desire to have the grid displayed in a different projection, in order that it represented ICES quads in the same manner as maps which they were accustomed to viewing. This is, indeed, an important consideration for the next phase of testing, as making the users as "comfortable" as possible (Plaisant 2005), is one of the main goals of the creation of these visualizations. That is to say, they should be presented with a visualization which is familiar to them, so their attention is not diverted from the actual fishery data.

6.3.3 Concurrent usefulness

Attention and comprehension is also central to the next conclusion. Both participants appreciated the understanding that each format of visualization could offer them. It was clear from their answers to questions 17, 22, 23 and 27 that for some tasks the statics were better, and for others, the animation was more useful. Static maps performed better when the task required viewing all months simultaneously, which, of course, allows the comparison of monthly data without having to remember previous information (Harrower 2007). Statics were also more useful for consulting single month occurrences and especially useful for non-consecutive months. Single month occurrences seem obviously best consulted with statics, as there is no waiting time for previous months to be played, as with the animation, and the month can be viewed for a longer period of time, in order to improve comprehension (Harrower 2007). Non-consecutive months are also largely considered best compared with static animations because there is no chance for temporal contiguity to be split (Harrower 2007). Users do not have to wait through and be distracted by meaningless months of information while waiting for the months which they seek. They will thus, not be affected by an overload of information (Marrs et al. 2002).

Animations are far from useless though. One feature which the statics do not contain is dynamism, which was addressed in Question 17. The participants equally expressed their appreciation for the understanding that could be gained about the spatial patterns of effort through time thanks to the moving data. This progression gave them a sort of visual “boost” as to how the data changed through the months on a general level. More specifically, animations can be effective for viewing consecutive months (Question 23) and tracking one ICES quad through time (questions 22 and 23). These conclusions were already highlighted in the objective questions, so it is reassuring to hear that the users’ subjective responses correlate with their competencies in answering the objective questions.

6.3.4 Animations no replacement

As may have been surmised from the previous two paragraphs, animations and statics are both useful media but they cannot perform all tasks equally well and thus must both be kept as possibilities. That being said, users were explicit, in answering Question 24, that animations could not be used in place of statics. Perhaps if animations included elements which the users thought were lacking, then they would be more widely accepted. In their current form, however, animations were not highly favored. Users acknowledged that static maps could take longer to use than animations, in some cases, but they were still more upbeat about using statics.

6.3.5 Time evaluation

Since it appears that users are still positive on statics, despite the drawbacks that they pointed out in the previous subsections, it may be that using time as a criterion for assessing which visualization media is more effective and efficient for answering questions is not a good one. Stakeholders have said outright that time is not a negative aspect of using statics, so it is clearly not a good indicator of their priorities or preferences for using visualizations. Perhaps then, it was not necessary to spend so much time evaluating the objective questions with time as the main usability criteria. This conclusion though, shows that the inclusion of both objective questions and subjective ones is a valuable tool in the usability assessment of geovisualizations. Had we only employed one type of question, we would have not been privy to this information.

6.3.6 Animation interactivity

Users did give some suggestions on how the speed and ease of using the animations could be increased. One way in which to improve the functionality of the animations, according to them, is to introduce a method in which specific months could be accessed immediately without having to wait through the entire animation to come to them. This interactivity is quite well studied and encouraged by Fairbairn et al (2001), among others. The interaction with the data, seeing how it changes according to one’s manipulation of different parameters, seems to be one of the ways in which cognition is maximized. Having a time slide bar was one of the early suggestions in pre-trials and users agreed that this may be one way to increase the usability of the animations so that they can answer the same types of questions that the static maps are good at answering.

6.3.7 Number of stills

Another suggestion that the participants had was to reduce the number of stills / visualizations. They suggested a bimonthly or seasonal aggregation of the data since there is not a great amount of variation at some points in the year. Unless very month-specific information is needed, this level of detail should be sufficient for answering most other questions relating to the distribution and magnitude of effort over the year. This would, of course, reduce the time of the animation, but it would also make the "paper load" of the static maps more manageable. So, it was both a technical and practical solution, given in Question 26.

It seems, from the subjective answers of the users, that they see the utility of using an animation but that they are still quite comfortable using static, paper maps. If we were to improve the functionality of the animations and include several capabilities, such as the ability to dynamically select an individual month, then the users would be more willing to incorporate this technology into their everyday usage.

6.4 Learnability

The learnability, or information retention, is somewhat hard to test, given that the second set of questions can easily be biased by the first, when testing whether animations are more useful than static maps. Nonetheless, this sort of bias is, in a way, a sign of information retention. The ability to retain information has also been directly tested in some of the questions, as this is one area where animations are perhaps not as convenient as statics.

For questions which were the same on both questionnaires, all of them would have been answered more quickly using the animation if the users had been biased by the static questionnaire. Out of the 18 chances for bias (nine questions per person), only five of them were more quickly answered with the static visualizations. This could mean that indeed, the users were biased. But as stated above, if they did recognize the question and remember the answer, at least this is some evidence of information retention, albeit not necessarily for spatial effort data.

There were seven questions on the animations that specifically required the users to remember something about the visualizations month-to-moth in order to answer them. There were several different kinds of questions though. There were ones that required remembering several occurrences of single ICES quads throughout the year, remembering spatial locations of occurrences of effort, remembering sizes of graduated circles (quantities), remembering general quantities of effort over the whole map, and remembering and aggregating monthly quantities. It appears though, that users will not remember things if they don't have to, or perhaps their memory recall is not fast enough because for all except one of these questions, the static maps were more useful in terms of time. Surprisingly, the only question which was performed quicker with the animations was one of the hardest. This was question number 13, which required the users to roughly remember all the months and then say which one had the greatest cumulative effort.

It should have been somewhat difficult to do so also since the temporal contiguity (Harrower 2007) was not so strong for the animation. That is to say, the stills were rather long and perhaps the users would not have been able to recall the previous still. It could also be argued though that the stills were indeed long enough (but not too long) to help Germane Cognitive Learning (Harrower 2007), which is the state when Working Memory has enough time to transfer its contents into Long Term Memory before it has to process new visual information.

Whatever the case may be, it still does not make up for the fact that the users did not find it necessary to use the animations because they were more able to answer the questions when their memory did not have to be exercised. In this respect, it could be concluded that the learnability of the data in the animations was not strong. However, the learnability of the animation technology, itself, was quite good, as it is a simple interface with few options.

6.5 Design flexibility

The ability to accommodate a multidisciplinary group is the main objective of this research project. Bringing fishery stakeholder together and engendering an effective dialogue and common understanding about the state of plaice fisheries in the North Sea was the impetus for making the visualization. The visualization is meant to be a tool for bringing together three different stakeholder groups: fishermen, scientists and policy makers. While at the current stage of research, only scientists have been included in the usability testing, it is envisioned that the other groups will also be integrated in.

In order to accommodate all these groups it was necessary to craft a visualization which amplifies cognition as much as possible. Designing for novices was encouraged (Dykes et al. 2005), so the visualization was created as simply as it could be while still incorporating all of the useful features envisioned at the beginning. We have made every attempt to use data, information and methods which draw upon the knowledge that stakeholders already possess, because it is the interaction between this knowledge and new information which results in the most effective learning (MacEachren 1995).

During the course of the usability evaluation, very few errors were made by the users and there were only several instances when they were not able to complete a question due to confusion, therefore it seems as if the animation and questionnaires were quite straightforward and to the point. It still remains to be seen if the same results are gotten when other stakeholder groups do the usability evaluation, but currently, the process seems to allow proper involvement of a multidisciplinary user group quite well. As mentioned earlier, even within groups there can be much diversity.

6.6 Conclusions

Based on the criteria put out in the “Measurements of Usability” of Ehn and Löwgren (1997), the attempt to bring together multiple stakeholder groups by building an animation incorporating all their data for them seems to have had mixed results. Since there were two different products which were part of this research project, it will be necessary to discuss them separately. They are: the usability testing process, namely the questionnaires, and the visualizations themselves.

6.6.1 Usability testing and questionnaires

We were able to evaluate the animations and compare them to the static visualizations according to user performance, subjective satisfaction, learnability and design flexibility. Users seemed to perform quite well with both static and animated maps. The time criteria favored using the static maps, although it was clear that this phenomenon was not always logical or conforming to our expectations / hypotheses. For completion percentage, which we defined as the percentage of questions answered under the target time, the animations were more useful. There seemed to be several questions which were particularly difficult to answer using the animations and thus, the total time of completion suffered. In terms of numbers of errors, there were very few on either animation questions or static map questions, though there were several more for static maps.

For subjective satisfaction, we received some very useful and interesting responses. The participants continuously stated that both formats of visualization, static and animated, had their own strengths and weaknesses in terms of their usability. While animations are good for tracking changes of one particular ICES quad over the course of the time span, static maps are still the preferred method for comparing effort quantities between non-consecutive months. Participants also noted that while static map might take slightly longer (a fact which we found in the objective questions to be absolutely false for total time) they may still be preferable if only one of the two formats were available. Had the animations included several features which were suggested by the participants, perhaps this would not have been the case. They said they would have liked to see more interactivity with the animation, particularly in the form of a scrollbar or some other feature to allow dynamic zooming to a particular month. It was revealed, however, that the cartographic features were well suited to the data and clearly understood.

It was slightly less clear though what the level of learnability of the animation was. Clearly, the data format and application used to deliver the animation were simple and didn't include any complicated functions. It was apparently difficult, however, to retain the base information displayed by the animation. That is, the data patterns were not easily mentally retained through the use of the animation. Participants more often had success using the static visualizations, where remembering data was not necessary and thus the time it took them to complete the questions was less.

The design flexibility seems to have been one of the strong points of this project. From the outset, it was understood that the application and visualization should be designed with the user in mind and that all possible distracting or unnecessary elements should be averted. It is true that this animation has not yet been tested on a wide sample of users, but so far, it seems to be performing at a level which is satisfactory to people of varied backgrounds and experience levels.

6.6.2 Visualizations

While the animation is usable, it is still not perfected. As discussed previously, even though the animation has improved on several aspects of information visualization that the static maps were not particularly strong in, the paper maps still hold the advantage in that they are simpler, cheaper and more well known to people. Animations will not be able to completely replace static maps until their functionality is increased for all kinds of analyses. Animations are particularly good at tracking the progression of one quad over a long period of time. They are also quite effective at enhancing the user's understanding of the spatial patterns of effort over the life of the data. Animations are poor, however, at answering simple queries for one given month, as it takes time to progress to that month and then the dynamic nature of the animation means that the data will not stay onscreen for very long. Lastly, and this was the most apparent conclusion from the usability testing, animations do not suffice when asked to compare two non-consecutive months of data, as the strain on the memory is quite significant.

6.7 Recommendations for future research

There are quite a number of possibilities for future research in this area. Some ideas relate to the number of participants in the usability study, the incorporation of stakeholders' suggestions into a new iteration of the animation, methodological maintenance related to the usability study and ultimately, the base data, itself.

Continuing with this project as a starting point, it would be nice to include more stakeholders from different groups (fishermen or policy makers) in the usability tests. Getting comments and suggestions from them would be invaluable, as their insights are based on completely different experiences and viewpoints which were not part of the evaluation done with fishery scientists. Through this strategy, we can perhaps get a better idea of how the users interact with the visualizations on a cognitive level. This knowledge will allow us to design future animation iterations while being better informed.

Ideally, it would also be useful to incorporate into the animation some of the suggestions that were given already. Key among them would be the presence of a dynamic selection tool, such as a time scrollbar for choosing the exact month's visualization and either viewing it as a still, as if it were static, or commencing the animation starting from that date. Certainly, in order to make information retention (learnability) more reasonably attained, the interactivity level of the animations must improve; something that is not at all possible with the animation capabilities of ArcGIS. The use of ArcGIS in this research, although originally practical, turned out to be more of a handicap than an asset. This was one of the reasons the users had difficulty with the current animation. We should explore the use of other software, such as Google Earth.

Another reason for difficulty was that some parts of the usability study, particularly the questionnaire were not clear to the participants. One glaring error, although maybe not avoidable, was the use of the same questions in both questionnaires. This caused some obvious bias in the answers for the animation questions. Trying at all costs, to avoid these types of influences should be made in the next version of this project. Another error which would be difficult to safeguard against is the misunderstanding which can sometimes occur between speakers of different languages. The primary stakeholders in this research are Dutch speakers, while the researcher is an English speaker, so there are bound to be times where language, especially on the questionnaires, is an impediment. Foreseeable problem areas should also be addressed though, such as giving extremely explicit directions at each stage of the questionnaire. At some points, we did not receive the exact type of answer we were looking for and this influenced the conclusions that could be drawn.

Lastly, the vast majority of the difficulties encountered in this research project were either directly or indirectly related to the quality of the data used. It was assumed early on that the quality of the data would not have a great deal of influence on the later stages of the project, but unfortunately, it did. Data for future use must be complete. Particularly, the completeness of the discards dataset should be improved, although it is acknowledged that since this is a dataset based on voluntary reporting, this may be easier said than done. Still though, in order for data to be displayed per month, per ICES quad, per mesh size, it must exist in this form. Until this data exists, bringing stakeholders together around a common tool such as a GIS animation will be a monumental task.

Bibliography

Aarts, G. M. and van Helmond, A. T. M. (2007). Discard sampling of plaice (Pleuronectes platessa) and cod (Gadus morhua) in the North Sea by the Dutch demersal fleet from 2004 to 2006. IJmuiden, IMARES**:** 42.

Allen, R. B. (1997). Mental Models and User Models. Handbook of Human-Computer Interaction. Helander, M. G., Landauer, T.K. and Prabhu, P.V. Amsterdam, Elsevier**:** 49-64.

Andrienko, G., Andrienko, N. and Gatalsky, P. (2005). Impact of Data and Task Characteristics on Design of Spatio-Temporal Data Visualization Tools. Exploring Geovisualization. Dykes, J., MacEachren, A. and Kraak, M. J.

Andrienko, N., Andrienko, G. and Gatalsky, P. (2003). "Exploratory spatio-temporal visualization: an analytical review." Journal of Visual Languages & Computing **14**(6): 503-541.

Arias, E., Eden, H. and Fisher, G. (1997). Enhancing Communication, Facilitating Shared Understanding, and Creating Better Artifacts by Integrating Physical and Computational Media for Design. Proceedings of the conference on Designing interactive systems: processes, practices, methods, and techniques, Amsterdam, The Netherlands, ACM Press, New York, NY.

Begg, G. A., Hare, J. A. and Sheehan, D. D. (1999). "The role of life history parameters as indicators of stock structure." Fisheries Research **43**(1-3): 141-163.

Bishop, I. D. and Lange, E. (2005). Visualization Classified. Visualization in landscape and environmental planning : technology and applications. Bishop, I. D. and Lange, E. London, Taylor & Francis**:** 23-34.

Booth, A. J. (2000). "Incorporating the spatial component of fisheries data into stock assessment models." ICES Journal of Marine Science **57**(4): 858-865.

Caddy, J. F. (2001). Spatial modelling in GIS fisheries applications Towards the use of Geographic Information Systems as a decision support tool for the management of Mediterranean fisheries. Bensch, A. and Taconet, M. Rome, Italy, Informes y Estudios COPEMED. **4**.

Casselman, B. (1999). "Visual Explanations: A Book Review." Notices of the AMS **46**(1): 43-46.

Chang, K. T. (2006). Introduction to Geographic Information Systems. NY, McGraw-Hill.

Daan, N. and van der Mheen, H. W. (2004). Outstanding Environmental Issues in relation to European Fisheries. RIVM, M.-e. N. IJmuiden, RIVO-Netherlands Institute for Fisheries Research**:** 21.

Devillers, R., Mostafavi, M. A., Fortin, M. J., Rose, G., Fotheringham, S., Meaden, G. J., Carette, V., Schoonbaert, S. and Windle, M. (2006). GeoCod: Geomatics for the Sustainable Management of Fish Stocks. GEOIDE annual conference, Banff Center, Alberta, Canada.

Drapeau, L. (2001). Aspects méthodologiques des SIG pour la compréhension des interactions entre pêcheries Towards the use of Geographic Information Systems as a decision support tool for the management of Mediterranean fisheries. Bensch, A. and Taconet, M. Rome, Italy, Informes y Estudios COPEMED. **4**.

Dykes, J., MacEachren, A. and Kraak, M. J. (2005). Exploring Geovisualization. Exploring Geovisualization. Dykes, J., MacEachren, A. and Kraak, M. J. Amsterdam, Elsevier**:** 3-22.

Eastwood, P. D. and Meaden, G. J. (2000). Spatial modelling of spawning habitat suitability for the sole (Solea solea L.) in the eastern English Channel and southern North Sea. Proceedings of ICES 2000 Annual Science Conference, Bruges, Belgium.

Edsall, R. M. and Sidney, R. L. (2005). Applications of a Cognitively Informed Framework for the Design of Interactive Spatio-temporal Representations. Exploring Geovisualization. Dykes, J., MacEachren, A. and Kraak, M. J. Amsterdam, Elsevier**:** 577-590.

Ehn, P. and Löwgren, J. (1997). Design for Quality-in-use: Human-Computer Interaction Meets Information Systems Development. Handbook of Human-Computer Interaction. Helander, M. G., Landauer, T. K. and Prabhu, P. V. Amsterdam, Elsevier.

Fairbairn, D., Andrienko, G., Andrienko, N., Buziek, G. and Dykes, J. (2001). "Representation and its relationship with cartographic visualization." Cartography and Geographic Information Science **28**(1): 13-28.

Fortunati, L., Garofalo, G. and Demontis, R. (2002). "TSDV: A GIS tool for inspecting trawl survey data." ICES Journal of Marine Science **59**(1): 168-178.

Fuhrmann, S., Ahonen-Rainnio, P., Edsall, R. M., Fabrikant, S. I., Koua, E. L., Tobon, C., Ware, C. and Wilson, S. (2005). Making Useful and Useable Geovisualization: Design and Evaluation Issues. Exploring Geovisualization. Dykes, J., MacEachren, A. and Kraak, M. J. Amsterdam, Elsevier**:** 553-566.

Fuhrmann, S. and Pike, W. (2005). User-centered Design of Collaborative Geovisualization Tools. Exploring Geovisualization. Dykes, J., MacEachren, A. and Kraak, M. J. Amsterdam, Elsevier**:** 591-610.

Furnas, G. W. (1986). Generalized fisheye views. Proceedings of ACM CHI'86 Conference, NY, ACM.

Garces, L. R., Silvestre, G. T., Stobutzki, I., Gayanilo, F. C., Valdez, F., Saupi, M., Boonvanich, T., Roongratri, M., Thouc, P., Purwanto, Haroon, I., Kurup, K. N., Srinath, M., Rodrigo, H. A. B., Santos, M. D., Torres, F. S. B., Tan, M. K. and Pauly, D. (2006). "A regional database management system-the fisheries resource information system and tools (FiRST): Its design, utility and future directions." Fisheries Research (Amsterdam) **78**(2-3): 119-129.

Gould, J. D., Boies, S. J. and Ukelson, J. (1997). How To Design Usable Systems. Handbook of Human-Computer Interaction. Helander, M. G., Landauer, T. K. and Prabhu, P. V. Amsterdam, Elsevier**:** 231-253.

Harrower, M. (2007). "The cognitive limits of animated maps." Cartographica **42**(4): 269-277.

Houde, S. and Hill, C. (1997). What do Prototypes Prototype? Handbook of Human-Computer Interaction. Helander, M. G., Landauer, T. K. and Prabhu, P. V. Amsterdam, Elsevier**:** 367-382.

ICES (2006). Report of the ICES Advisory Committee on Fishery Management, Advisory Committee on the Marine Environment and Advisory Committee on Ecosystems. Copenhagen, International Council for the Exploration of the Sea**:** 90.

ICES (2007a). Factors affecting the distribution of North Sea fish. ICES-FishMap. Copenhagen, International Council for the Exploration of the Seas**:** 10.

ICES (2007b). The North Sea Fish Community. ICES-Fishmap. Copenhagen, International Council for the Exploration of the Seas**:** 12.

ICES (2007c). Plaice. ICES-FishMap. Copenhagen, International Council for the Exploration of the Seas**:** 10.

ICES (2009). Glossary. ICES-FishMap, International Council for the Exploration of the Seas**:** 5.

ISO (1998). Ergonomic requirements for office work with visual display terminals - Part 11: Guidance on usability, ISO. **9241-11**.

Kraak, M. J. (2003). "Geovisualization illustrated." Isprs Journal of Photogrammetry and Remote Sensing **57**(5-6): 390-399.

Kwan, M. P. (2004). "GIS methods in time-geographic research: Geocomputation and geovisualization of human activity patterns." Geografiska Annaler Series B: Human Geography. 2004; 86(4): 267 280.

Landauer, T. K. (1997). Behavioral Research Methods in Human-Computer Interaction. Handbook of Human-Computer Interaction. Helander, M. G., Landauer, T.K. and Prabhu, P.V. Amsterdam, Elsevier**:** 203-228.

Larkin, J. W. and Simon, H. A. (1987). "Why a diagram is (sometimes) worth ten thousand words." Cognitive Science **11**: 65-99.

Lohse, G. L. (1997). Models of Graphical Perception. Handbook of Human-Computer Interaction. Helander, M. G., Landauer, T.K. and Prabhu, P.V. Amsterdam, Elsevier**:** 107-136.

MacEachren, A. (1995). How Maps Work: Representation, Visualization qnd Design. NY, Guilford.

MacEachren, A. (2005). Moving Geovisualization toward Support for Group Work. Exploring Geovisualization. Dykes, J., MacEachren, A. and Kraak, M. J. Amsterdam, Elsevier**:** 445-462.

MacEachren, A., Brewer, I. and Steiner, E. (2001). Geovisualization to mediate collaborative work: Tools to Support Different-place Knowledge Construction and Decision-Making. 20th International Cartographic Conference, Beijing, China, ICA.

MacEachren, A., Gahegan, M., Pike, W., Brewer, I., Cai, G., Lengerich, E. and Hardisty, F. (2004). "Geovisualization for knowledge construction and decision-support." Computer Graphics & Applications **24**: 13-17.

MacEachren, A. M. (2001). "An evolving cognitive-semiotic approach to geographic visualization and knowledge construction." Information Design Journal **10**(1): 26-36.

MacEachren, A. M., Gahegan, M. and Pike, W. (2004). "Visualization for constructing and sharing geo-scientific concepts." Proceedings of the National Academy of Sciences of the United States of America **101 Suppl. 1**: 5279-5286.

Marrs, S. J., Tuck, I. D., Atkinson, R. J. A., Stevenson, T. D. I. and Hall, C. (2002). "Position data loggers and logbooks as tools in fisheries research: results of a pilot study and some recommendations." Fisheries Research **58**(1): 109-117.

Norman, D. A. (1993). Things That Make Us Smart. Reading, MA, Addison-Wesley.

Olson, G. M. and Olson, J. S. (1997). Research on Computer Supported Cooperative Work. Handbook of Human-Computer Interaction. Helander, M. G., Landauer, T. K. and Prabhu, P. V. Amsterdam, Elsevier**:** 1433-1454.

Plaisant, C. (2005). Information Visualization and the Challenge of Universal Usability. Exploring Geovisualization. Dykes, J., MacEachren, A. and Kraak, M. J. Amsterdam, Elsevier.

Politis, P. (2003). Application of GIS in Marine Fisheries Trawl Surveys, University of Rhode Island**:** 5.

Rijnsdorp, A. D., Daan, N., Dekker, W., Poos, J. J. and van Densen, W. L. T. (2006). "Sustainable use of flatfish resources: Addressing the credibility crisis in mixed fisheries management." Journal of Sea Research **57**(2007): 114-125.

Roberts, J. C. (2005). Exploratory Visualization with Multiple Linked Views. Exploring Geovisualization. Dykes, J., MacEachren, A. and Kraak, M. J. Amsterdam, Elsevier**:** 159-180.

Staples, L. (1993). Representation in virtual space: visual convention in the graphical user interface. Conference on Human Factors in Computing Systems - Proceedings.

Stoner, A. W., Manderson, J. P. and Pessutti, J. P. (2001). "Spatially explicit analysis of estuarine habitat for juvenile winter flounder: Combining generalized additive models and geographic information systems." Marine Ecology Progress Series **213**: 253-271.

Valavanis, V. D. (2002). Geographic Information Systems In Oceanography And Fisheries. London, Taylor and Francis.
Valavanis, V. D., Georgakarakos, S., Kapantagakis, A., Palialexis, A. and Katara, I. (2004). "A GIS environmental modelling approach to essential fish habitat designation." Ecological Modelling **178**: 417-427.
van Densen, W. (2007). Communicating spatio-temporal patterns in the fishery for plaice and sole, IMARES**:** 12.
van Densen, W. L. T. (2008). Thesis Proposal meeting. Gold, C. Wageningen.
van Keeken, O. A., van Hoppe, M., Grift, R. E. and Rijnsdorp, A. D. (2007). "Changes in the spatial distribution of North Sea plaice (Pleuronectes platessa) and implications for fisheries management." Journal of Sea Research **57**(2-3 SPEC. ISS.): 187-197.
Virzi, R. A. (1997). Usability Inspection Methods. Handbook of Human-Computer Interaction. Helander, M. G., Landauer, T.K. and Prabhu, P.V. Amsterdam, Elsevier**:** 705-716.
Visvalingam, M. (1994). Visualisation in GIS, Cartography and ViSC. Visualization in Geographical Information Systems. Hearnshaw, H. M. and Unwin, D. J. Chichester, UK, John Wiley & Sons.
Waterworth, J. A. and Chignell, M. H. (1997). Multimedia Interaction. Handbook of Human-Computer Interaction. Helander, M. G., Landauer, T. K. and Prabhu, P. V. Amsterdam, Elsevier**:** 915-946.
Wennhage, H., Pihl, L. and Stal, J. (2007). "Distribution and quality of plaice (Pleuronectes platessa) nursery grounds on the Swedish west coast." Journal of Sea Research **57**(2-3 SPEC. ISS.): 218-229.
Wright, D. and Bartlett, D. (2000). Marine and Coastal Geographical Information Systems. London, Taylor and Francis.
Zeller, D. and Pauly, D. (2001). "Visualisation of standardized life-history patterns." Fish and Fisheries **2**(4): 344-355.

Appendices

Appendix 1: Data availability limits

Part One

200	ICES quads in North Sea
*****	multiplied by
24	months in time period
=	equals
4800	data records possible

Part Two

We expect to have 4800 records for each of the following variables:

Part Three

Considering 4800 possible records, are the variables complete or at least manageable?

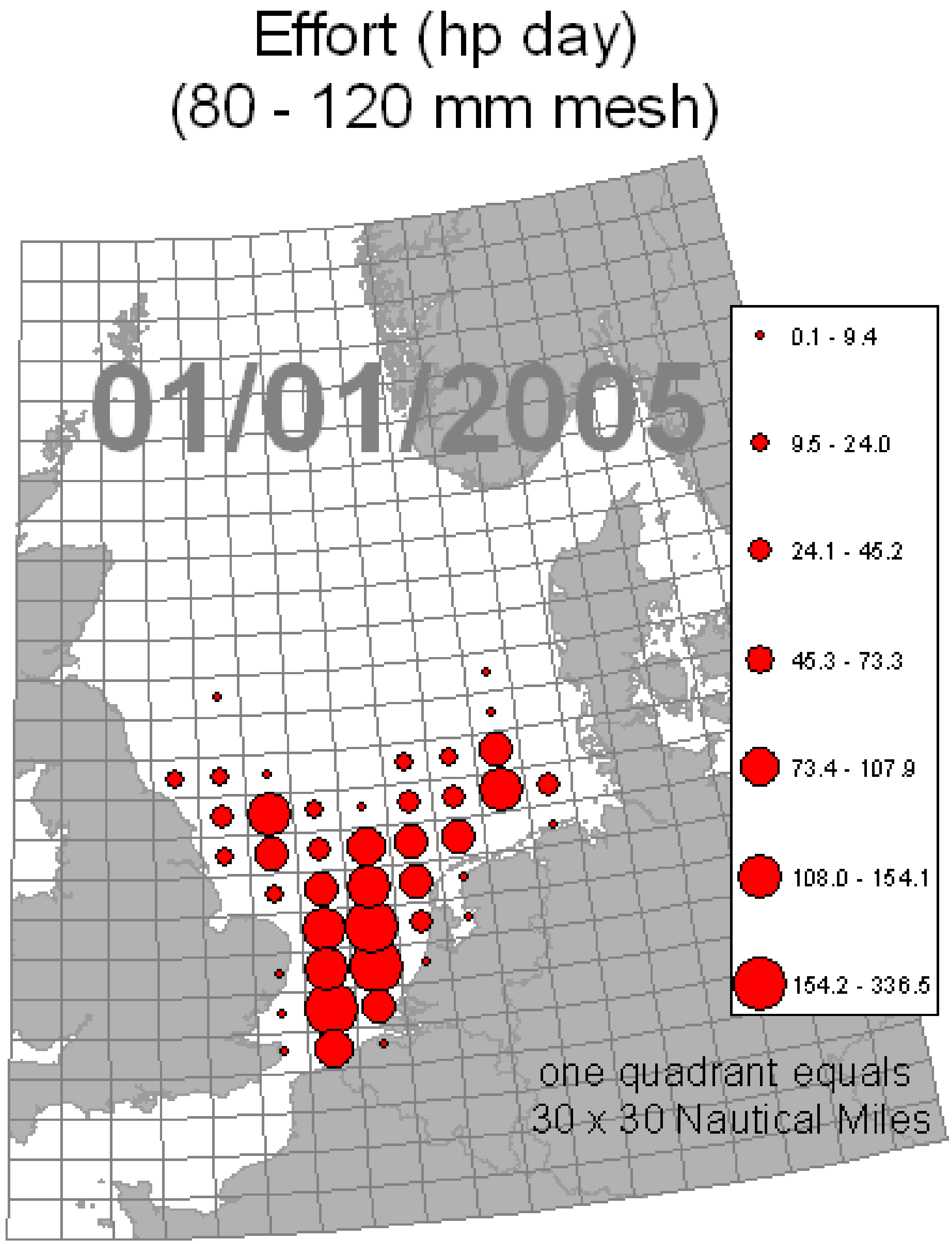
CPUE (per class) →	yes
discard percentage →	no
effort →	yes
mesh size →	no

Part Four

Conclusions:

1. Nearly complete variables come from the official market survey conducted by Floor Quirijns, while the variables which are not at all complete came from a commercial volunteer survey compiled by Edwin van Helmond.
2. The lack of discard data from the commercial survey is exacerbated by the fact that the datasets are further fractioned according to three mesh sizes when visualized.
3. Of course, visualizing CPUE as a total, or divided into classes is possible, since this dataset is quite complete.
4. We can also easily visualize the effort dataset, as it is also nearly complete.
5. Visualizing discards (not compared with CPUE) is also a possibility, either as a whole or per mesh size.
6. It would also be possible to visualize mesh sizes to get a spatial and temporal understanding of that variable regardless of catches.

Appendix 2: Visualization Layout



Appendix 3: Personal Questionnaire

1. What is your name?

2. What is your job title?

3. Do you consider yourself a:

    a. Fisheries scientist?
    b. Policy maker?
    c. Fisherman?

4. How long have you worked in the fisheries science industry?

5. What level would you rate yourself in terms of experience with:

    - Static maps? 1-5; 5 is most experienced

    - Dynamic / animated maps?

Appendix 4: Questionnaire for static visualizations

Answers are in **Bold**

1. What variable / phenomenon is presented in this animation?
    - **Effort**
    - Catch
    - Discards
    - Number of vessels

2. What unit do the symbols represent? **(hpday)**

3. What geographic location is displayed? **(North Sea or similar)**

4. What does the grid represent?
    - Exclusive Economic Zones (EEZ)
    - **ICES quadrants**
    - Conservation areas
    - Fishing holes

5. What size is each grid cell?
    - 3 square Nautical Miles
    - 90 x 90 NM
    - 270 sq. NM
    - **30 x 30 NM**

6. How long is the timeline of the visualized data?
    - 1 year
    - 12 months
    - 2 decades
    - **2 years**

7. Are the geographic distributions of effort roughly similar month-to-month? **(no)**

8. During which months and years is the quad north of Norwich, England free from effort? **(10, 11, 12 2005; 8,9,11 2006)**

9. During which month and year does an occurrence of effort occur furthest south? **(01/01/06)**

10. During which months and years does an occurrence of effort occur furthest east? **(6, 7 2005; 5, 6, 7 2006)**

11. Does the quad west of Zeeland increase or decrease in effort from 01/06/2005 – 01/06/2006? **(Decrease)**

12. Does this quad ever experience the minimum effort in this animation? **(no)**

13. For both years, which month would it appear that the cumulative effort is at its minimum?

    a. 2005
        i. April
        ii. June
        iii. December

    b. 2006
        i. February
        ii. April
        iii. June

14. What time of year are high efforts most spatially dispersed?
    a. Summer
    b. Winter
    c. Autumn

15. In which month do efforts seem to "jump" North? **(May)**

DIRECTIONS: For the following two questions, please indicate your level of agreement with the statement (1 is low, 5 is high).

16. The quantities and/or patterns of effort displayed in these statics are similar to those which I previously drew on the MapTable. Please explain what was different or the same.

17. It useful to be able to view all the months at the same time.

---

18. At which time of year does coastal Denmark seem to experience the most fishing effort?
    a. Late winter
    b. Late spring
    c. Early fall
    d. Early summer

19. What can you say about the immediate West coast of Denmark in 01/11/2006?
    a. Activity but very thinly spread
    b. Boats all over
    c. Devoid of effort
    d. Very high effort

20. During which month and year do two spatial outliers occur in the far North?
    a. 01/05/2005
    b. 01/05/2006
    c. 01/06/2005
    d. 01/10/2006

21. When does there appear to be effort on land in Flevoland?
    a. 01/03/2005
    b. 01/06/2005
    c. 01/01/2006
    d. 01/10/2006

DIRECTIONS: For the following six questions, please indicate your level of agreement with the statement (1 is low, 5 is high).

22. You are able to make meaningful conclusions from the static visualizations. If so, what are they?
    -
    -
    -

23. In the following scenarios, you can imagine the static maps being more useful than the animation.
    - Checking the intensity of effort for one ICES quad in one specific month
    - Comparing effort distribution for consecutive months
    - Comparing effort distribution for the same months in different years
    - Tracking the changes in effort of a single ICES quad throughout the year

24. If an animation was not available, a paper map would suffice.

25. The following characteristics of the static maps should be improved. Please specify after indicating your level of agreement.
    - The placement of elements on screen
    - The size of onscreen elements
    - The color of onscreen elements
    - The cartographic elements (graduated symbols, grid, etc.)

26. There were too many static maps for me to be able to effectively respond to the questions.
    - I would have preferred 15
    - I would have preferred 10
    - I would have preferred 5

27. I prefer the static maps over the animation.

Appendix 5: Questionnaire for animated visualizations

Answers are in **Bold**

1. What variable / phenomenon is presented in this animation?
    - **Effort**
    - Catch
    - Discards
    - Number of vessels

2. What unit do the symbols represent? **(hpday)**

3. What geographic location is displayed? **(North Sea or similar)**

4. What does the grid represent?
    - Exclusive Economic Zones (EEZ)
    - **ICES quadrants**
    - Conservation areas
    - Fishing holes

5. What size is each grid cell?
    - 3 square Nautical Miles
    - 90 x 90 NM
    - 270 sq. NM
    - **30 x 30 NM**

6. How long is the timeline of the visualized data?
    - 1 year
    - 12 months
    - 2 decades
    - **2 years**

7. Are the geographic distributions of effort roughly similar month-to-month? **(no)**

8. During which months and years is the quad north of the Groningen Wad free from effort? **(1, 2, 4, 11 2005; 2, 3, 5 2006)**

9. Disregarding two spatial outliers in the far North, which month and year sees effort at its highest latitude? **(01/07/06)**

10. Which month and year does effort extend furthest West? **(01/01/05 and 01/02/05)**

11. Does the quad west of Zeeland increase or decrease from 01/03/2005 – 01/03/2006? **(decrease)**

12. Does this quad ever experience the maximum effort in this animation? **(yes 01/04/06)**

13. For both years, which month would it appear that the cumulative effort is greatest?

- 2005
    i. September
    ii. October
    iii. December

- 2006
    i. January
    ii. October
    iii. November

14. What time of year are high efforts most spatially concentrated?
    - **Early spring**
    - Late summer
    - Early winter

15. In which month do efforts seem to "jump" North? **(May)**

DIRECTIONS: For the following two questions, please indicate your level of agreement with the statement (1 is low, 5 is high).

16. The quantities and/or patterns of effort displayed in this animation are similar to those which I previously drew on the MapTable. Please explain what was different or the same.

17. The movement of the data in the animation aids in my understanding of the spatial patterns of effort.

---

18. At which time of year does coastal Denmark seem to experience the most fishing effort?
    - Late winter
    - Late spring
    - Early fall
    - **Early summer**

19. What can you say about the immediate West coast of Denmark in 01/11/2006?
    - Activity but very thinly spread
    - Boats all over
    - **Devoid of effort**
    - Very high effort

20. During which month and year do two spatial outliers occur in the far North, off the western coast of Norway?
    - **01/05/2005**
    - 01/05/2006
    - 01/06/2005
    - 01/10/2006

21. What is remarkable about the Thames River estuary in 01/06/2005?
    - The river dries up
    - Plaice fisheries are replaced by sole fisheries
    - Southern England is flooded
    - **There is effort which appears on dry land**

DIRECTIONS: For the following six questions, please indicate your level of agreement with the statement (1 is low, 5 is high).

22. You are able to make meaningful conclusions from the animated visualizations. If so, what are they?
    -
    -
    -

23. In the following scenarios, you can imagine the animation being more useful than the static maps.
    a. Checking the intensity of effort for one ICES quad in one specific month
    b. Comparing effort distribution for consecutive months
    c. Comparing effort distribution for the same months in different years
    d. Tracking the changes in effort of a single ICES quad throughout the year

24. If static paper maps were not available, an animation would suffice.

25. The following characteristics of the animation should be improved. Please specify after indicating your level of agreement.
    a. The placement of elements on screen
    b. The size of onscreen elements
    c. The color of onscreen elements
    d. The cartographic elements (graduated symbols, grid, etc.)
    e. The length of time stills.
    f. The progression from one time still to another.
    g. The interactivity of the animation.

26. Would you like to be able to progress and rewind incrementally by way of a slide bar which would denote the months at each increment?

27. I prefer the animation over the static maps.

Appendix 6: Results of Questionnaires

STATIC	Person A			B					
	answers	time	time +/-	answers	time	time +/-	notes	**target time**	**question #**
		4	-11		14	-1		**15**	**1**
		24	9		9	-6		**15**	**2**
	D	11	1		16	6		**10**	**3**
		9	-11		16	-4		**20**	**4**
		26	1		10	-15		**25**	**5**
		14	-31		20	-25		**45**	**6**
	F	26	-9		74	39		**35**	**7**
	F	94	19	F	95	20	B: partially right	**75**	**8**
		57	12		40	-5		**45**	**9**
	F	42	-3		70	25		**45**	**10**
	V	43	-2	V	237	192		**45**	**11**
	C	51	16		33	-2		**35**	**12**
	D	87	-3		270	180		**90**	**13**
	D	47	2	D	29	-16		**45**	**14**
		38	-7		40	-5		**45**	**15**
	1			V					**16**
	yes			yes			no numbers		**17**
		44	-16	D	100	40		**60**	**18**
		41	16		52	27		**25**	**19**
		49	-11		31	-29		**60**	**20**
		45	20	F	93	68		**25**	**21**
	2						B: quantities and geography		**22**
	4,4,4,1			1,2,3,4			B: inverse of what I expected		**23**
	5			yes			paper takes longer		**24**
	4,4,4,4			ok			see comments		**25**
	4,4,1						seasonal or bi-monthly ok		**26**
	V			V			impossible to answer		**27**
		752			1249				**Total time**
	4			2					**# wrong**

Key: **1-5=mark** **V=void** **D=debatable** **C=confusion** **F=wrong**

Appendix 6 continued

		ANIMATED	Person A →			B →			
target time	**question #**		answers	time	time +/-	answers	time	time +/-	notes
15	**1**			4	-11		8	-7	
15	**2**			15	0		8	-7	
10	**3**			21	11		3	-7	
20	**4**			8	-12		4	-16	
25	**5**			7	-18		5	-20	
45	**6**			9	-36		6	-39	B: played several times
35	**7**			10	-25		100	65	
75	**8**		F	142	67		216	141	A: partially right. B: replay animation
45	**9**			81	36		56	11	
45	**10**		F	75	30		204	159	A: partially right
45	**11**		V	63	18	V	67	22	my mistake
35	**12**			57	22		91	56	
90	**13**		D	79	-11	D	141	51	
45	**14**			47	2		265	220	
45	**15**			29	-16		53	8	
	16					V			
	17		4			yes			particularly interesting
60	**18**		D	69	9		107	47	
25	**19**			25	0		70	45	
60	**20**		F	58	-2		20	-40	A: mix-up?
25	**21**			89	64		23	-2	
	22		5						A: good for concentrating on one quad. B: good for movement - faster
	23		3,5,2,5			no,yes,no,yes			
	24		yes			no			A: if works properly. B: static for comparing, animated for general patterns through time
	25		e,f,g			d,e,g			B: grid not regular, faster changes, scrolling better
	26		C			yes			
	27		Both			both			both useful
Total time				888			1447		
# wrong			3			0			

Appendix 7: User comparison

questions	Static time		Animation time		Time difference		Conclusion
	User A	B	A	B	A	B	which better?
1	4	14	4	8	0	-6	anim
2	24	9	15	8	-9	-1	anim
3	11	16	21	3	10	-13	
4	9	16	8	4	-1	-12	anim
5	26	10	7	5	-19	-5	anim
6	14	20	9	6	-5	-14	anim
7	26	74	10	100	-16	26	
8	94	95	142	216	48	121	static
9	57	40	81	56	24	16	static
10	42	70	75	204	33	134	static
11	43	237	63	67	20	-170	
12	51	33	57	91	6	58	static
13	87	270	79	141	-8	-129	anim
14	47	29	47	265	0	236	static
15	38	40	29	53	-9	13	
18	44	100	69	107	25	7	static
19	41	52	25	70	-16	18	
20	49	31	58	20	9	-11	
21	45	93	89	23	44	-70	
							6 for static
					negative = animated faster		6 for animation

Appendix 8: MapTable exercise

MapTable Introduction

Using the MapTable was suggested during preliminary stages of this project. Unfortunately, it did not integrate well into the methodology which was ultimately chosen. Nonetheless, it seemed like an interesting activity for understanding how different stakeholder groups understood the problem of discards of plaice in the North Sea fishing industry, so we chose it as a sort of “add-on” exercise for the usability meeting.

MapTable Exercise Setup

After the Personal Questions, but before the main usability testing, there will be an activity employing the MapTable. Stakeholders are meant to undertake this activity individually, just as for the rest of the testing.

They will be asked to sketch on the MapTable ten separate maps. They are divided into predators (fishermen) and prey (plaice). Three different variables will be available for the predator maps. They are effort (kg/hpday), catch (kg) and discard (kg). The stakeholders will sketch one map per variable. For the prey maps, two variables exist: landable sized fish and undersized fish.

For each variable, they will be asked to sketch two different maps; one for winter and one for summer. Quantities of the variable shall be represented as High, Medium or Low. The drawing (representation) of the variables will be left completely up to the stakeholders and their choices of representation techniques will be of utmost interest.

As each individual map is completed to the best of their knowledge / hypothesis, it will be saved for later analysis. The activity will be entirely video-taped in order to draw conclusions on the stakeholders’ actions. It is expected that each map may take up to five minutes to complete depending on the level of precision used.

Results of testing

Initially, we envisioned having the MapTable activity at the outset of the meeting. With the first participant, we did this, although it was deemed unnecessary to continue with the second, since he is not a North Sea plaice specialist and would not have the knowledge necessary to complete this activity. Since the first participant is only one of a handful of people in the Netherlands who know about this data, any other non-experts would also not be able to complete the exercise with any sort of certainty. Additionally, the MapTable was not functioning terribly well, including saving the data which was inputted. As a result, there is no useful quantifiable data available for objective analysis.

On the questionnaire there was one question included in the subjective set of questions. Question 16 focused on the MapTable exercise and how its results compared with the visualized effort in the static and animated maps. The first participant was the only one to take part in the MapTable exercise, therefore we cannot make good comparisons. Nonetheless, he noted that the visualized effort, seemed to extend further North than what he expected reality to show (what he drew on the MapTable).

Discussion

The first interesting conclusion that can be drawn from the subjective questions is that even in our small sample size, we saw evidence of the exact problem we are trying to ameliorate. During the MapTable exercise, the first participant inputted some data which he believed to be representative of reality. While doing the questionnaire (Question 16), he mentioned that the effort which was depicted in the visualizations seemed to expand too far North, in his mind. From this, we can hypothesize that he was not accustomed to viewing the exact data which had been used for the creation of the visualizations and was, at the very least, unfamiliar with it, and perhaps even skeptical of its validity. By designing a visualization which is pleasing to all stakeholder groups, we hope to avoid circumstances where the users are surprised by something in the visualization.

Conclusion

Unfortunately, there were very few useful conclusions which were drawn from the MapTable exercise either. To avoid this next time, it should be exactly clear to the users what they are expected to do with the MapTable exercise, but of course, these directions for use and background information cannot influence or bias the users in any way. This is a fine line which must be worked out in the future.