Using Interactive Visualization to Enhance Understanding of a Fisheries Model

BY

Carmen Rose St. Jean

B.S., University of New Hampshire, 2010

THESIS

Submitted to the University of New Hampshire
in Partial Fulfillment of
the Requirements for the Degree of

Master of Science

in

Computer Science

May, 2014

This thesis has been examined and approved.

Thesis director, Colin Ware
Professor of Computer Science

R. Daniel Bergeron
Professor of Computer Science

Matt Plumlee
Affiliate Assistant Professor in Computer Science

Date

TABLE OF CONTENTS

LIST OF FIGURES

CHAPTER 1

Introduction

Fishery managers have only one "lever" to pull when it comes to fishery management: the ability to set harvest quotas. Fishermen work within these quotas by exerting various levels of fishing effort. Production models have been designed to help managers and fisherman understand the ecosystem they work within and the implications of their decisions. In this context, a production model is a mathematical model, based on data, that simulates interactions between species and predicts species biomass as a function of fishing effort, climate change, and other variables. MS-PROD is a multispecies production model developed by NOAA scientists Gamble and Link (Gamble and Link, 2009). A visualization may enhance such a model by making its inner workings more explicit and may be useful for decision making. The goal of this research has been to explore design alternatives and evaluate the effectiveness of different modes of portrayal and interaction to make a visualization of the MS-PROD model that will be a valuable tool to the modelers and stakeholders alike.

1.1 Models

Both the short- and long-term effects of human exploitation on an ocean ecosystem, such as the species inhabiting the Gulf of Maine, are not easily understood since experiments which would allow ecosystem managers to investigate the impact of different levels of exploitation over many years are either impractical or impossible to conduct on a large scale. Fortunately, ecosystem models can be used instead to help gain a better understanding of an ecosystem.

1.1.1 Lotka-Volterra Equations

Ecosystem models are abstract representations of an ecological system, and can range from an individual species in its environment to an entire community of species. A classic example is the Lotka-Volterra model, which is a pair of differential equations for describing the non-linear interactions between a predator species and a prey species (Lotka, Volterra, 1926, 1926):

$$\frac{dN_1}{dt} = N_1 \left(\alpha - \beta N_2\right) \tag{1.1}$$

$$\frac{dN_2}{dt} = -N_2 \left(\gamma - \delta N_1\right) \tag{1.2}$$

where N_1 is the number of prey animals, N_2 is the number of predator animals, t is time, α is the prey animals' growth rate, β is the rate at which the predators destroys the prey animals, γ is the death rate of the predators, and δ is the rate at which the predators increase from consuming the prey animals. The model can be generalized to discuss an arbitrary number of species rather than just a single pair.

The Lotka-Volterra model can be modified to take competition instead of predation into account, as in the Rosenzweig-MacArthur model (Rosenzweig and Macarthur, 1963) and the Leslie-Gower model (Leslie and Gower, 1960). These adaptations to the model also consider carrying capacity, which is the maximum number of a species that can be sustained indefinitely in a particular environment:

$$\frac{dN_1}{dt} = r_1 N_1 \left(1 - \left(\frac{N_1 + \alpha_{12} N_2}{K_1}\right)\right) \tag{1.3}$$

$$\frac{dN_2}{dt} = r_2 N_2 \left(1 - \left(\frac{N_2 + \alpha_{21} N_1}{K_2}\right)\right) \tag{1.4}$$

where r_i is the growth rate for species i, K_i is the carrying capacity for species i, and α_{ij} is the effect species j has on species i. As with Lotka-Volterra, this model concerns only two species, but it can be generalized to include more than two.

Both Lotka-Volterra and Leslie-Gower do not incorporate a factor which is critical when discussing fisheries management: the effect of harvest. The Schaefer model adds a term to account for the effect of harvest on an individual species (Schaefer, 1957):

$$\frac{dN}{dt} = rN\left(1 - \left(\frac{N}{K}\right)\right) - qEN \tag{1.5}$$

where N is the number (or biomass) of the species, r is the growth rate, K is the carrying capacity, q is the catchability coefficient, and E is the fishing effort.

Simple models, when available and correct, are generally preferred; since fewer components are needed to describe their real-world counterparts, they are more easily understood and implemented. All three of these models are subjectively simple in that they only consider a few ecological factors each. However, in reality, ecosystems are complex systems which require management that recognizes them as such (Christensen et al., 1996). Thus, a more holistic approach called ecosystem-based fishery management (EBFM) has been advocated (Panel, 1999). However, this approach has not often been implemented due to a lack of models which consider all necessary ecological factors. Gamble and Link developed a multispecies production model (MS-PROD) to fill this gap (Gamble and Link, 2009).

1.1.2 The MS-PROD Model

The MS-PROD model forecasts biomass for species separated into functional groups, which are biological groupings of species that perform similar functions within their ecosystem. The model is built upon the Schaefer production model and also includes Lotka-Volterra terms for predation, Leslie-Gower terms for competition, and carrying capacities for functional groups (K_G) as well as for the entire system (K_σ):

$$\frac{dN_i}{dt} = \underbrace{r_i}_{\substack{\text{Growth}\\\text{rate}}} N_i \left(\underbrace{1 - \frac{N_i}{K_G}}_{\substack{\text{Prevents}\\\text{infinite}\\\text{growth}}} \underbrace{- \frac{\sum_{j=1}^{g} \beta_{ij} N_j}{K_G} - \frac{\sum_{j=1}^{G} \beta_{ij} N_j}{K_\sigma - K_G}}_{\text{Competition}} \right) \underbrace{- N_i \sum_{j=1}^{P} \alpha_{ij} N_j}_{\text{Predation}} - \underbrace{H_i N_i}_{\text{Harvest}} \tag{1.6}$$

where N_i is the number (or biomass) of species i, t is a unit of time, r_i is growth rate for species i, β_{ij} is the competition of species j on i, α_{ij} is the predation of species j on i, H_i is the harvest rate on species i, g is the number of species within species i's group, G is the number of groups, and P is the number of predators.

This model is distinguished from other multispecies production models by describing stocks with explicit ecological and harvest factors. Each species to be included in the simulation must be specified in the parameter file by listing growth rate, functional group membership, initial biomass, carrying capacity, and catchability. Additionally, matrices representing inter-species relationships must be provided to describe the relationship between every pair of species. Such matrices are required for *predation*, where one species consumes the other, and *competition*, where one species affects the other in any manner besides predation. (In ecology, the word "interaction" is often used instead of "competition," but we have chosen to use "competition" so that the term "interaction" will not be overloaded since "interactive" has a separate meaning in a visualization context.)

The MS-PROD authors provided us with a parameter file which listed ten key species chosen from the Northeast United States Continental Shelf Large Marine Ecosystem (NEUS LME), listed here by functional group:

- **Elasmobranchs:** Skates, Spiny Dogfish
- **Flatfish:** Windowpane, Winter Flounder, Yellowfin Tuna
- **Groundfish:** Cod, Haddock, Redfish
- **Small Pelagics:** Herring, Mackerel

The MS-PROD model runs simulations for 30 years with an annual time step to predict individual biomasses. While this outputted information is potentially valuable to fishery managers, it was lacking an interactive graphical user interface.

1.2 Visualization Methods

When designing any complex interactive visualization, a large number of design choices must be made. This section reviews existing research of visualization methods which have relevance to the problem of creating an effective interface to a fisheries ecosystem model—in particular, methods for representing time series, networks, causality, and uncertainty.

1.2.1 Time Series

Fisheries management is focused on the sustainability of choices concerning fish stocks. A main purpose of ecosystem management is to ensure that future generations can enjoy the same natural resources (Christensen et al., 1996). As such, the MS-PROD model provides biomass forecasts for 30 years. Therefore, visualization techniques for temporal data must be explored.

Line Charts

The line chart was first invented by William Playfair in 1786 to communicate time series data, seen in Figure 1-1 (Playfair, 1786). Today, it remains a common method for visualizing time-oriented data in many fields, including science, economics, planning, and engineering to name a few. Line charts typically encode time on the horizontal axis, progressing from left to right, and some time-varying value on the vertical axis. Points in the chart are connected by line segments such that the slope of the line indicates the rate of change between time steps.

Multiple time series can be part of a single line chart; each series needs only to be distinguished by a color and/or line style. However, as the number of time series on a single line chart increases, it becomes more difficult to identify an individual series. Javed et al. evaluated the four different plotting techniques for multiple time series illustrated in Figure 1-2 (Javed et al., 2010). The first of the techniques is the "simple line chart," which was Playfair's original line chart with all series plotted together. A slight variation on that is "small multiples," where each series had its own line chart though all charts share the same

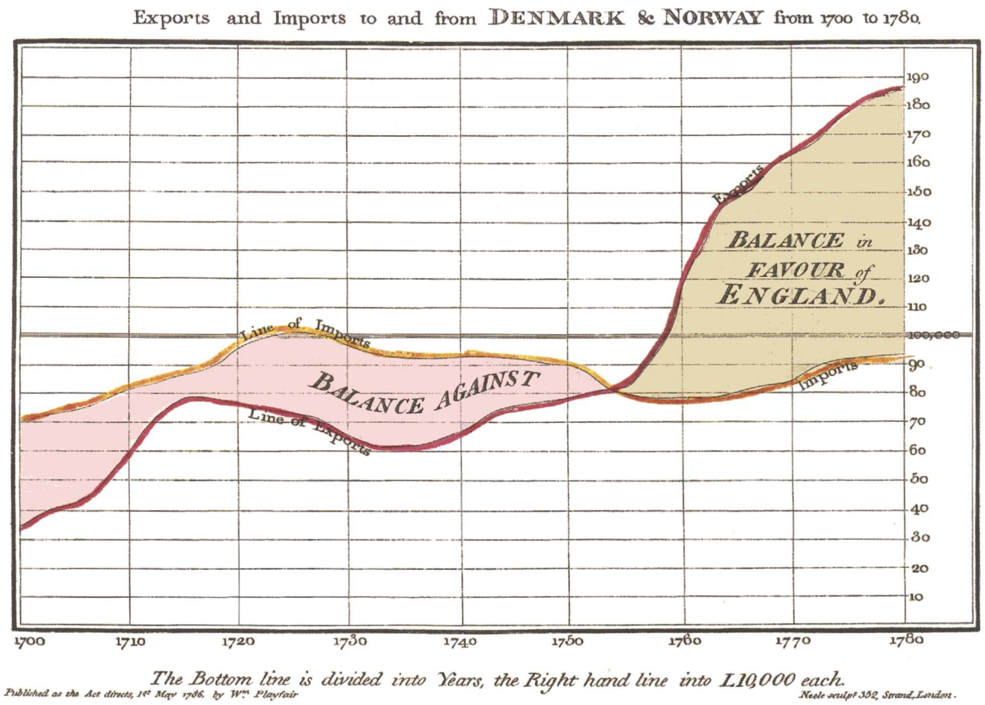
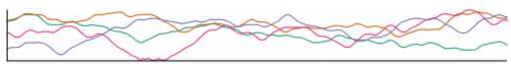


Figure 1-1: Playfair's original time series chart (Playfair, 1786).

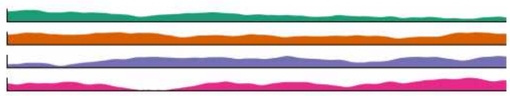
axis scales. Horizon graphs, originally developed by Saito et al., wrap around a baseline in two color tones to save space (Saito et al., 2005). Lastly, braided graphs feature all series on one chart with the coloring under the curves alternating as series intersect each other. The user evaluation by Javed et al. revealed that a simple line graph with all time series on one plot or a single graph for each time series is better suited to a variety of tasks than a horizon graph or a braided graph. They also found that users complete tasks more correctly when there is more display space allocated to the graphs. They did not recommend using a higher number of simultaneous time series—their study used eight at the most—because it leads to a decline in correctness of task completion.

Some line charts are more effective at conveying the nature of the data than others because of the way different drawing techniques affect interpretability. Cleveland et al. found the shape of a line chart—defined as the height of the chart divided by the width of the chart—to be a critical factor (Cleveland et al., 1988). The shape of the chart directly

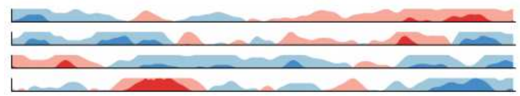


(a) A simple line graph.

(b) A braided graph.

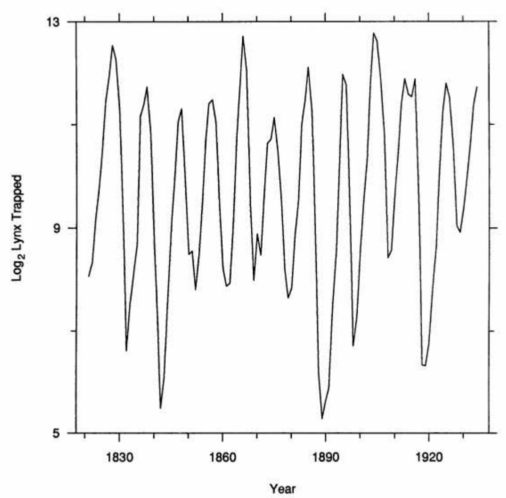


(c) Small multiples.

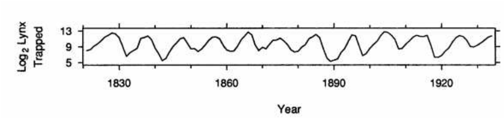


(d) Horizon graphs.

Figure 1-2: Four possible methods for visualizing multiple time series. Reprinted from "Graphical perception of multiple time series" by Javed et al. with permission. ©2010 IEEE

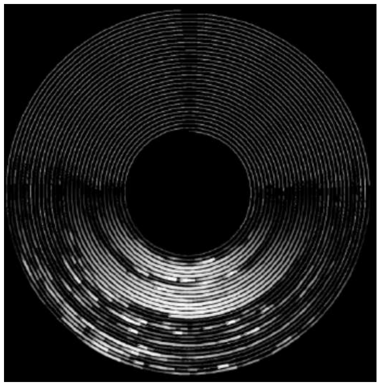


(a) Shape parameter of 1.0.

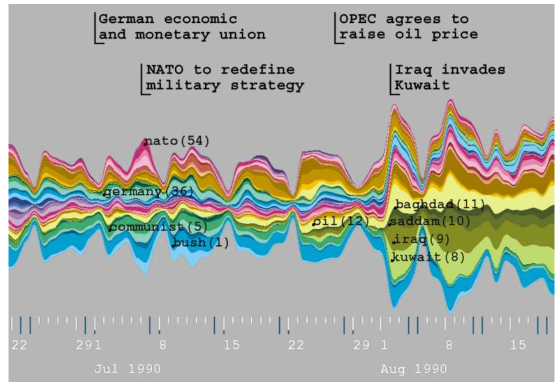


(b) Shape parameter of 0.074.

Figure 1-3: Two effects of chart shape on Canadian lynx data (Cleveland et al., 1988).



(a) Spiral time series. Reprinted from "Visualizing time-series on spirals" by Weber et al. with permission. ©2001 IEEE



(b) ThemeRiver time series. Reprinted from "ThemeRiver: visualizing theme changes over time" by Havre et al. with permission. ©2000 IEEE

Figure 1-4: Two alternative time series visualizations.

impacts the slopes of line segments, which viewers interpret in order to understand the dependence of the y variable on the x variable. Figure 1-3a, a time series of Canadian lynx trapping data, features a shape of 1.0 and seems to imply rapid increases and decreases in the population. On the other hand, Figure 1-3b has a shape of 0.074 and shows more clearly that the population rises somewhat steadily and declines somewhat rapidly, which Figure 1-3a failed to show. Their user evaluation found that judgment of two slopes is influenced by the orientation mid-angle, defined as the average of the minimum slope orientation and the maximum slope orientation. They proposed line chart shape should be selected such that orientations are as close to $\pm 45°$ as is possible, like in Figure 1-3b.

Alternatives

There are many alternatives to and variations of Playfair's original time series. One example is Weber et al.'s spiral time series, seen in Figure 1-4a (Weber et al., 2001). The spiral time series was designed for cyclic data. Cycles are emphasized in a properly-parameterized spiral visualization, however it may be difficult to describe periodic behavior in unknown datasets or determine if that behavior even exists. Another example is the ThemeRiver by

Havre et al, seen in Figure 1-4b (Havre et al., 2000). Each "current" in the ThemeRiver represents an entity or subject and must be of a distinctive color. Positioning along the y-axis is meaningless, instead the abundance of the entity or subject over time is indicated by the width of the band of color. The overall width of the stream is the sum of the widths of all individual bands.

1.2.2 Networks

The input parameters to the MS-PROD model includes predation and competition matrices. The model may be better understood if these relationships can be incorporated into the visualization. Relationships are often visualized through a node-link diagram, which typically represents entities as nodes and links as relationships between the nodes they connect. There are many types of node-link diagrams used for illustrating networks; those which are relevant to our research are discussed in the following sections.

Force-Directed Layouts

One option for showing fish species interactions would be to use a force-directed layout as Gaichas and Francis did, seen in Figure 1-5 (Gaichas and Francis, 2008). Here, the nodes represent an individual species in the Gulf of Alaska, while the links represent a predator-prey interaction. In a force-directed layout, nodes repel each other, while related nodes become pulled toward each other by links (Heer et al., 2010). The result is an aesthetically pleasing layout where there are relatively few link crossings and links are of approximately similar length. The color of the node can be used to indicate group membership, while the size can represent the magnitude of some property of the node. Likewise, the drawing style of the link can be varied to encode different types of relationships. Depending on the size of the network, a force-directed layout can be dense, like in Figure 1-5, making it difficult to discern individual nodes or links. Interactive methods can alleviate this by allowing the user to zoom or to click a node and see only the subset of the network directly connected to that node.

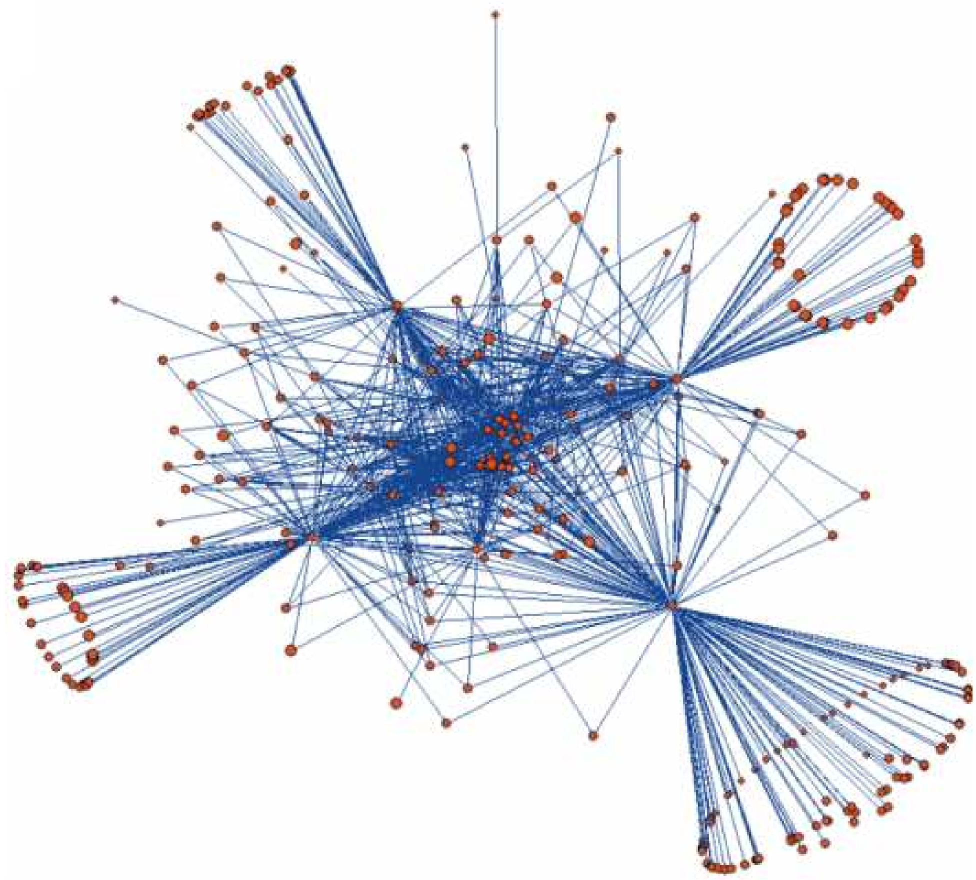


Figure 1-5: A force-directed visualization of a food web of Gulf of Alaska data (Gaichas and Francis, 2008).

Arc Diagrams

An alternative for force-directed layout is an arc diagram. The name arc diagram was coined by Wattenberg (Wattenberg, 2002), though they were invented earlier. Knuth used arc diagrams to illustrate interaction of characters in Victor Hugo's novel *Les Misérables*, seen in Figure 1-6 (Knuth, 1993). Each character is represented with a circular node, where size indicates the number of appearances in the novel. The nodes are arranged linearly, colored and ordered according to clusters of characters that appear together frequently. Semi-transparent arcs are drawn between the characters which appear in the same chapter, with the thickness of the arc representing the number of such appearances. While the arc diagram may fail to properly depict the structure of a network, Heer et al. point out it is

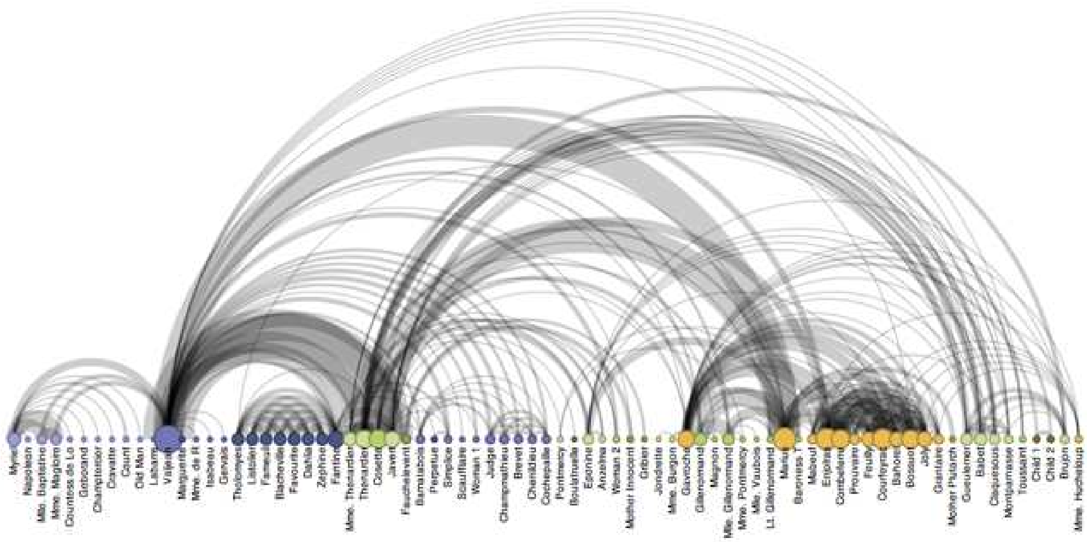


Figure 1-6: Knuth's arc diagram of *Les Misérables* characters (Knuth, 1993).

advantageous because the one-dimensionality allows for other features to be easily displayed near the nodes (Heer et al., 2010), such as text labels.

Directed Edges

Relationships in a network may be directional, such as the predator-prey relationship. In a visualization of such a network, the direction of the edges must be encoded so these relationships can be understood. Holten and van Wijk studied the effectiveness of different techniques for indicating directionality of edges in a graph, seen in Figure 1-7 (Holten and van Wijk, 2009). The traditional arrowhead was found to perform poorly, while tapered edges performed best. As for an intensity-based direction cue, a dark-to-light representation was found to be clearer than light-to-dark.

Matrix Representations

Node-link diagrams can have occlusion problems when they are highly-connected, so a matrix-based representation of a network is a possible alternative (Heer et al., 2010). In many cases, networks are stored as an adjacency matrix, so all that needs to be done is visualize that matrix as a grid, where the cell at the ith row and the jth column represents

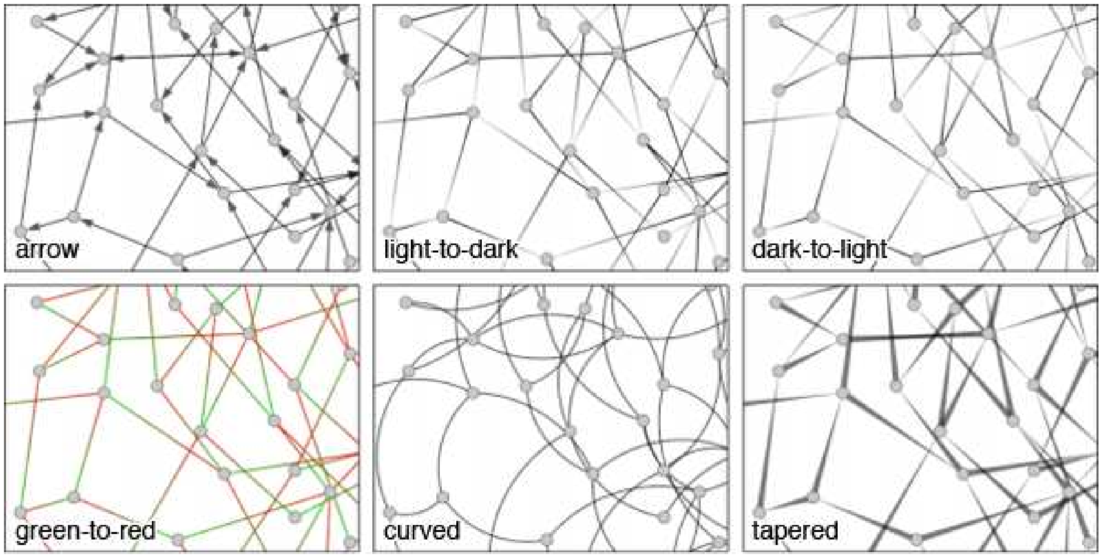


Figure 1-7: Different types of directed edges (Holten and van Wijk, 2009).

the relationship from entity i to entity j. Figure 1-8 shows Knuth's visualization of *Les Misérables* characters in matrix-form (Knuth, 1993). The color of the cell indicates the presence or type of a relationship, with some neutral color indicating the lack of a relationship. Ghoniem et al. showed that a matrix-based view is suitable for large or dense networks for tasks that involve finding or counting links or nodes (Ghoniem et al., 2004). With proper ordering of the rows and columns, the structure of the network can be effectively displayed, however path-finding tasks may be difficult.

1.2.3 Causality

Gamble and Link found that both inter-species relationships and harvesting by humans can contribute to changes in species biomass according to their MS-PROD model (Gamble and Link, 2009). For example, an increase in biomass for one species could possibly hinder growth for another species. This is a type of cause and effect relationship, therefore it is necessary to consider the various techniques for representing causality in a network, especially in an interactive context.

Michotte and Thinés suggested that viewers infer causality after viewing an object being set into motion after being struck by another object (Michotte and Thinés, 1963). This

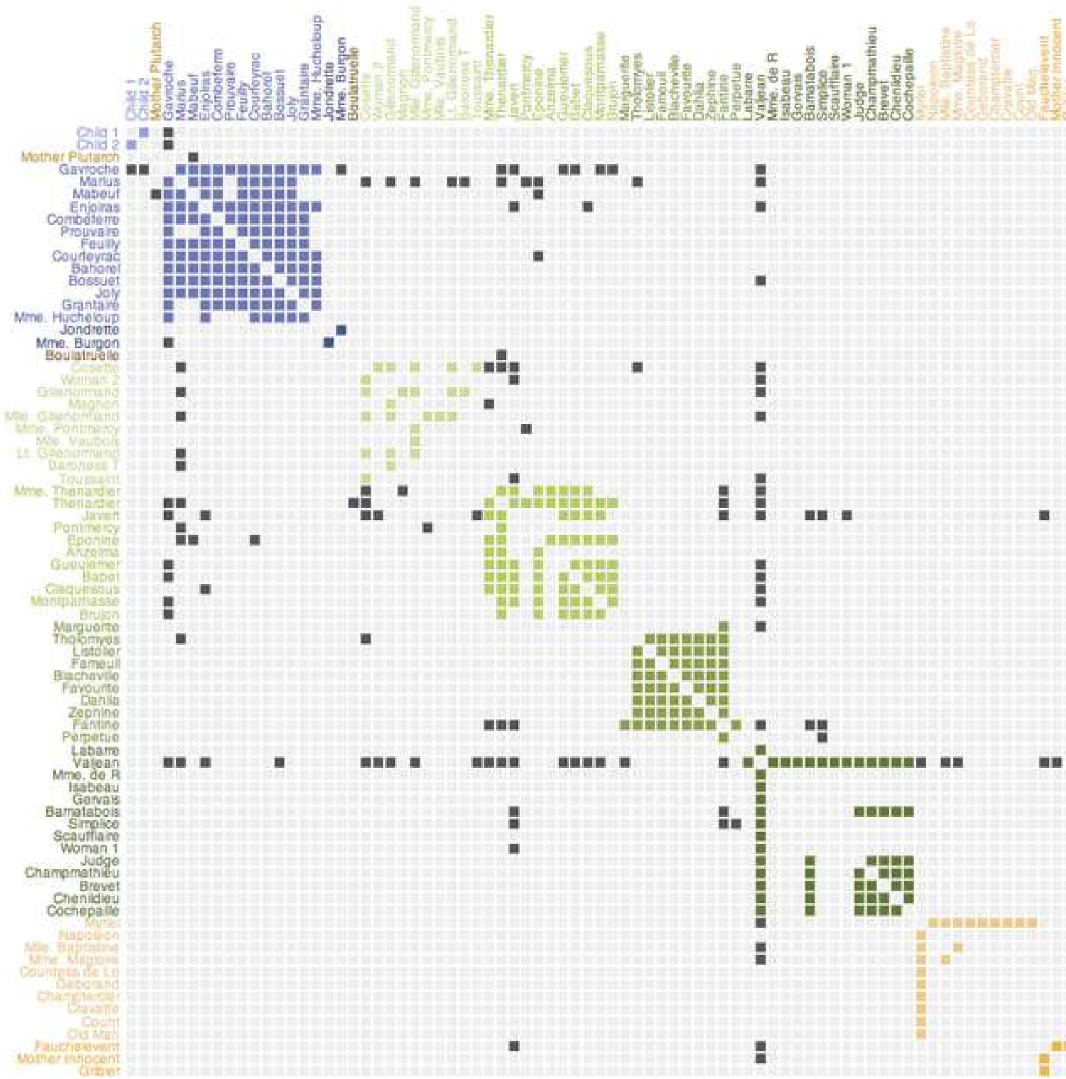


Figure 1-8: A matrix-based visualization of an adjacency matrix (Knuth, 1993).

served as a basis for Ware et al.'s visual causality vector (VCV), which communicates causal relationships between two nodes in a network visualization (Ware et al., 1999). They studied how several animated metaphors—shown in Figure 1-9—and different timing rules for a VCV affect the perception of causality. Their evaluation showed that temporal synchrony between the animation of the metaphor and the changes in the recipient node is more critical than the type of metaphor for showing causal relationships. Ware later revisited this work in the context of multi-touch screens to convey causal effect enhancements, causal effect reductions, and causal blocking effects using colored pulses (Ware, 2013). The user

(a) Pinball. (b) Prod. (c) Wave.

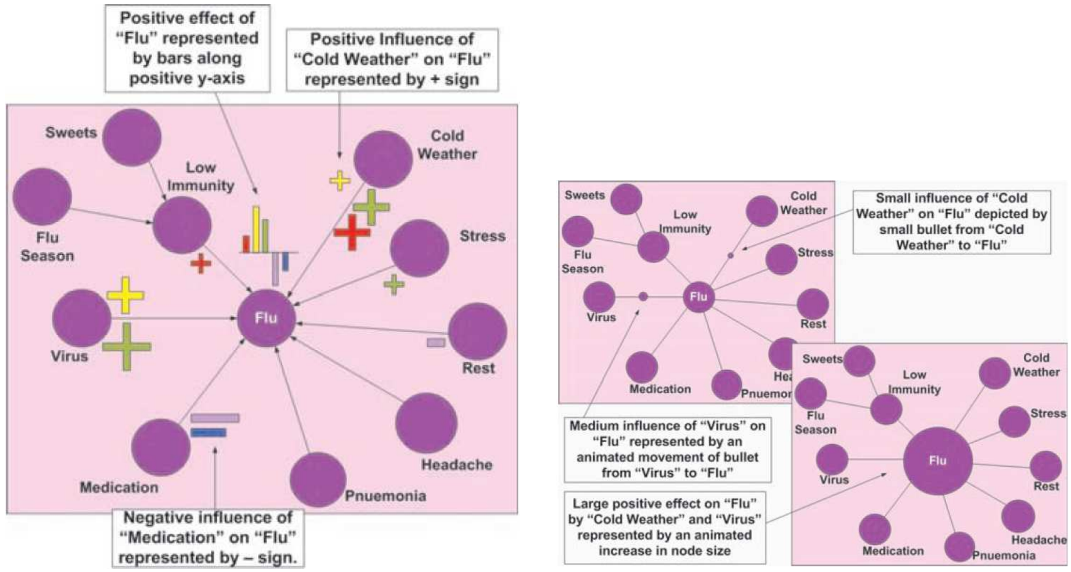
Figure 1-9: Three metaphors for conveying causality. Reprinted from “Visualizing causal relations” by Ware et al. with permission. ©1999 IEEE

evaluation conducted by Ware showed that causal blocking effects and positive enhancements were well understood with this design, while negative causal effects were less reliably judged. Still, these methods were recommended for showing simple causal relationships.

Kadaba et al. expanded upon Ware et al.'s VCV work to compare between static and animated causal visualizations (Kadaba et al., 2007). In their static design, positive influences were indicated with a plus sign (+) glyph and negative influences were indicated with a minus sign (−) glyph attached to the link between two entities in the network, as in Figure 1-10a. The size of the glyph represented the magnitude of the influence on the recipient node and glyphs of the same color described a multiplication effect on the recipient. Their animated design featured bullets traveling along the links toward the recipient node to indicate causal influences, shown in Figure 1-10b. As a bullet hit the recipient node, the size of the recipient node changed. They found that subjects can interpret animated and static representations equally accurately, but subjects formed responses slightly quicker with animated representations.

1.2.4 Uncertainty

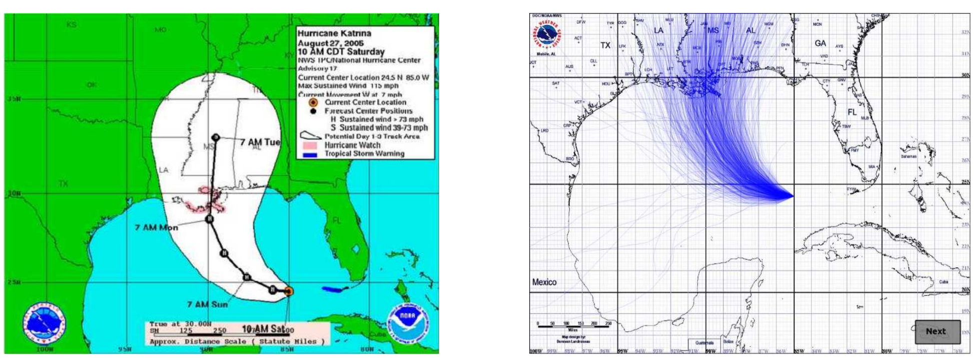
A model is a simplification of reality, therefore there are many reasons why it may be inaccurate. There may be inaccurately estimated parameters used in the model, such as the predation matrix coeffecients or the starting values for the numbers of fish species. Furthemore, relevant factors may not have been taken into account when designing the model, such as climate. Inaccuracies may even result from approximation techniques, such



(a) Static.

(b) Animated.

Figure 1-10: Two alternatives for visualizing causal influences. Reprinted from "Visualizing causal semantics using animations" by Kadaba et al. with permission. ©2007 IEEE



(a) Error cone.

(b) Cox et al.'s method.

Figure 1-11: Two visualizations of uncertainty for a hurricane advisory (Cox et al., 2013).

as the use of a Runge-Kutta method. Therefore, it is common for a model's output to be regarded with some uncertainty. To clarify, *error* describes inaccuracies when the correct answer is known, while *uncertainty* describes inaccuracies when the answer is unknown (Hunter and Goodchild, 1993). Due to uncertainty, model output is best understood as a range of expected values which is likely to contain the true value. According to Potter et al., scientific data should be considered to be incomplete without representations of uncertainty (Potter et al., 2010), so we have investigated some different portrayals of uncertainty.

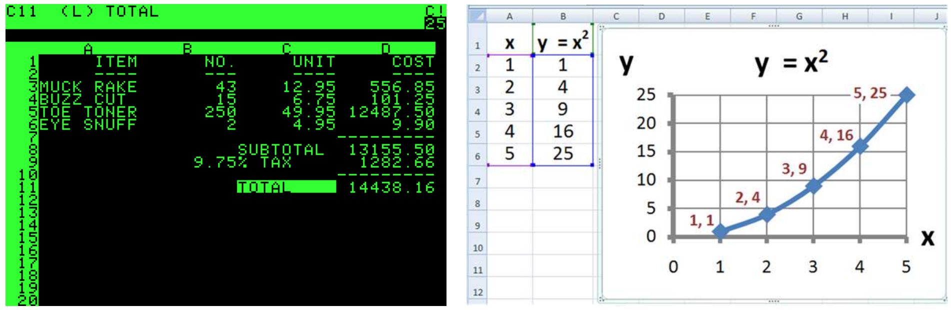
Cox et al. explored methods for depicting uncertainty for hurricane advisories (Cox et al., 2013). The traditional error cone, as in Figure 1-11a, is used by NOAA's National Hurricane Center, but it can be poorly understood by the average citizen. They designed their own alternative representation, as in Figure 1-11b, where many possible hurricane tracks are drawn to describe the range of possible outcomes. Their user evaluation found that neither method was significantly better than the other for all cases, but a qualitative analysis showed that nearly all users prefer their new method over the error cone.

1.3 Understanding Models

Users studying models can benefit from the aid of a visualization, because patterns and trends may be difficult—if not impossible—to discern from only a table of numerical values. The learning process can be even further enhanced through interaction with the model. If interactivity is supported by the visualization, then users can adjust parameter values, perceive a change (or perhaps no change) in the results, and begin to understand the degree of influence different parameters possess.

1.3.1 Spreadsheet Programs

VisiCalc was a very early example of software assisting the understanding of models through visualizations (Grad, 2007). As a business student, Bricklin wished there was a faster way to change the input or fix mistakes when working out financial models by hand (Bricklin and Frankston, 1999). To address this, he worked with Frankston to develop VisiCalc, seen



(a) A screenshot of VisiCalc (GNU General Public License).

(b) A chart made using Microsoft Excel (public domain).

Figure 1-12: Two examples of spreadsheet applications.

in Figure 1-12a. As the world's first electronic spreadsheet, VisiCalc consisted of rows and columns containing either text, numerical values, or formulas. Result cells were instantly updated according to changed inputs or adjusted formulas, allowing a user to work with models in a more efficient and dynamic manner. VisiCalc was superseded by Lotus 1-2-3, which was in turn supplanted by Microsoft Excel. Microsoft Excel remains popular and features graphing tools which can generate charts, such as in Figure 1-12b.

1.3.2 The Influence Explorer

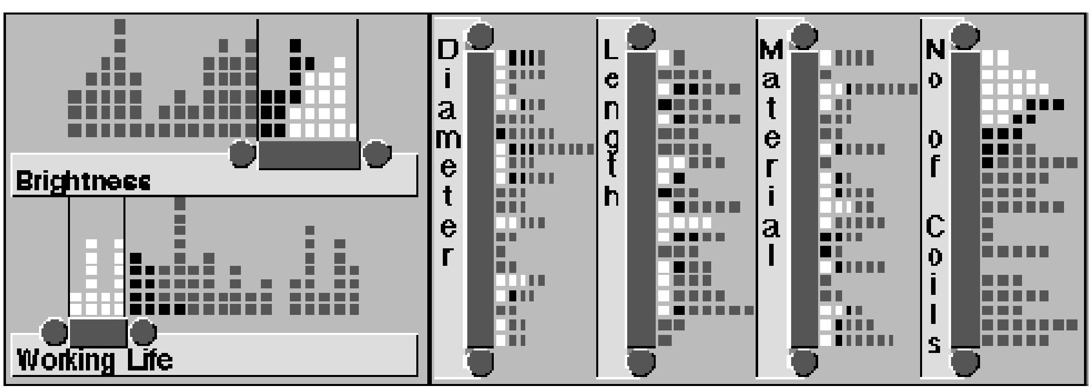


Figure 1-13: A screenshot of the Influence Explorer (Tweedie et al., 1995).

The Influence Explorer by Tweedie et al. is a good example of a more complex interactive visualization (Tweedie et al., 1995). They developed an interface for understanding the relationships between different attributes in a design process. Design parameter values of the Influence Explorer are initially randomly selected in a Monte Carlo simulation to represent many different possible designs. Each cell in the histogram is a simulation run. For each attribute, there is a histogram including each of the items. The attribute ranges are controlled by sliders. When the user adjusts the slider of a given attribute, all items that are within that range are highlighted on all of the histograms. Figure 1-13 is a screenshot of the Influence Explorer being used to test the performance of different light bulb designs; white indicates the design passed, black it failed one specification, and grey it failed two specifications. In a user evaluation, industrial designers found the ability to interactively explore the effects of different parameter ranges to be valuable.

1.3.3 Vensim

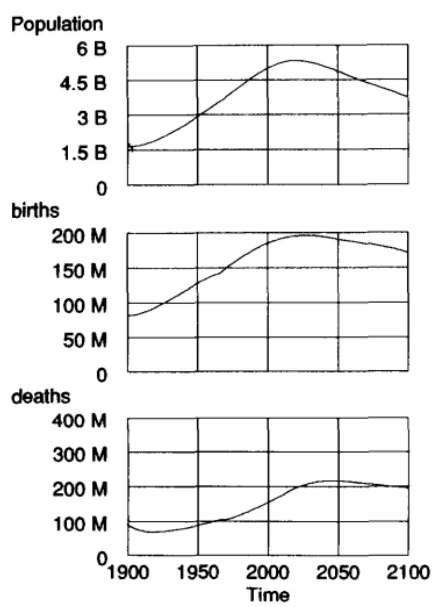


Figure 1-14: A screenshot of the strip graphs in Vensim (Eberlein and Peterson, 1992).

Eberlein and Peterson recognized that both unskilled and skilled model users have the same need: to quickly obtain a thorough understanding of a model and its implications (Eberlein and Peterson, 1992). This motivated their development of Vensim, which is a commercial tool for visualizing and analyzing simulation results. Vensim allows users to run a model under different conditions with a simple mouse click, enabling the user to learn the effects of different actions with ease. Various features enhance this learning process—e.g., "causal tracing" strip graphs, shown in Figure 1-14. Rather than simply seeing a chart of the projected population, a user exploring the causal tracing feature can begin to understand the various components that contributed to the population simulations—births and deaths—by seeing each component on its own chart. A major emphasis of Eberlein and Peterson is that the same visualization tool can and should be used for both the development and teaching phases of a model, especially since those phases may not be discrete.

CHAPTER 2

Requirements and Design

The motivation of this research is to provide an interactive interface to the NOAA MS-PROD model so as to help fishermen, fisheries managers, and other stakeholders understand the implications of decisions, such as changing catch quotas for particular kinds of fishing activity—e.g., bottom trawling versus mid-water trawling. We planned our interactive interface with the intent that users would gain insight into:

- **Implications of the model:** E.g., how do two different sets of fishing effort values affect the biomass predictions of the ten species?
- **The model itself:** E.g., why does the abundance of one species increase when another species is caught?

To accomplish this goal, we determined our interactive interface of the MS-PROD model would need to visualize:

- The predicted biomass over time for all modeled species
- Changes in predicted biomass as a result of changes in fishing effort for a subset of species
- Causal relationships which help to explain why the model predicts the effects of changes in fishing effort for a subset of species
- Uncertainty of the model

Since setting fishing quotas is the only action fishery managers can take, we determined that the only interaction available to users would be to adjust the fishing effort level; all other model parameters are not modifiable once the model and its visualization are running. This interaction is done by means of a set of sliders with the goal of allowing the user to immediately see the impact of management decisions on fisheries. The user adjusts sliders which represent harvest effort and watches the biomass plots change instantaneously as the model is re-run according to the new effort values. Like Eberlein and Peterson, we aimed to turn the "time consuming and tedious" task of working with a model into a "fast and fun" interactive experience (Eberlein and Peterson, 1992). Different views and features of the model visualization are described and discussed in the sections below.

It is important to note that the MS-PROD model allows for different fishing effort values for each species for each year in the thirty-year period that the model produces forecasts for. This is to reflect that fishing quotas can change from year to year. Since there are ten species and thirty years, this means there are potentially 300 separate fishing effort values that could be set by a user. However, it would be cluttered and confusing to have an effort slider for each of those 300 values. Therefore, two simplifications were made. First, fishing effort should be controlled by functional group rather than individual; the model authors requested this because it is more realistic than fishing by individual species. Second, fishing effort for a functional group is constant across all thirty years; while this is unrealistic, it is easier to understand and perhaps, in a sustainable, "perfect world," fishing quotas would not need adjusting over the years. Thus, each slider controls the fishing effort for all thirty years for a particular functional group.

2.1 Visualization of Predicted Biomass

First and foremost, we had to decide how to display the thirty-year biomass forecast data output by the model. Time series line charts—of both the simple line chart and small multiples varieties—were chosen because casual viewers can understand them without further instructions, as opposed to, say a horizon graph. Another advantage to line charts is that they tend to have a reasonable amount of whitespace where additional information can be displayed, such as uncertainty or alternate forecasts. A key design issue concerned the representation of change between two forecasts; line charts provide the ability to display this change. A number of alternatives for change were implemented and are described in the sections that follow.

2.1.1 Alternative Screen Layouts

And early design decision concerned the issue of how many individual charts should be used to display the data and how to arrange those charts. Two alternative screen layouts were developed for displaying these time series on line charts: a “four panel” view in Figure 2-1 and a “small multiple” view in Figure 2-2, described in further detail below. With both views, the time series for the species are organized by functional group. A functional group is a biological grouping of species which perform similar functions within their ecosystem—e.g., mackerel and herring are both members of the “small pelagic” group since they live in the water column.

There were several reasons for the decision to arrange by functional group rather than placing all time series on a single line chart. First, harvest effort is controlled by functional group using the sliders, so it must be somehow indicated which species are part of which functional group. With multiple time series, this can be encoded through positioning by arranging the slider of a functional group to be near the time series of that group’s species. Second, it would have been impractical to simply place all time series on a single line chart because would have been difficult to select enough distinct colors to represent each of the ten species. Furthermore, the biomass of some species is significantly larger than others. By scaling the y-axis according to the largest biomass value of all species in the entire 30-year time span, the lines representing some species would have been crowded at the bottom of the chart and seemed to be flat even when they were not.

Four Panel View

It was observed that species of similar functional groups tend to have biomass values in similar numeric ranges, so a decision was made to have one line chart per functional group

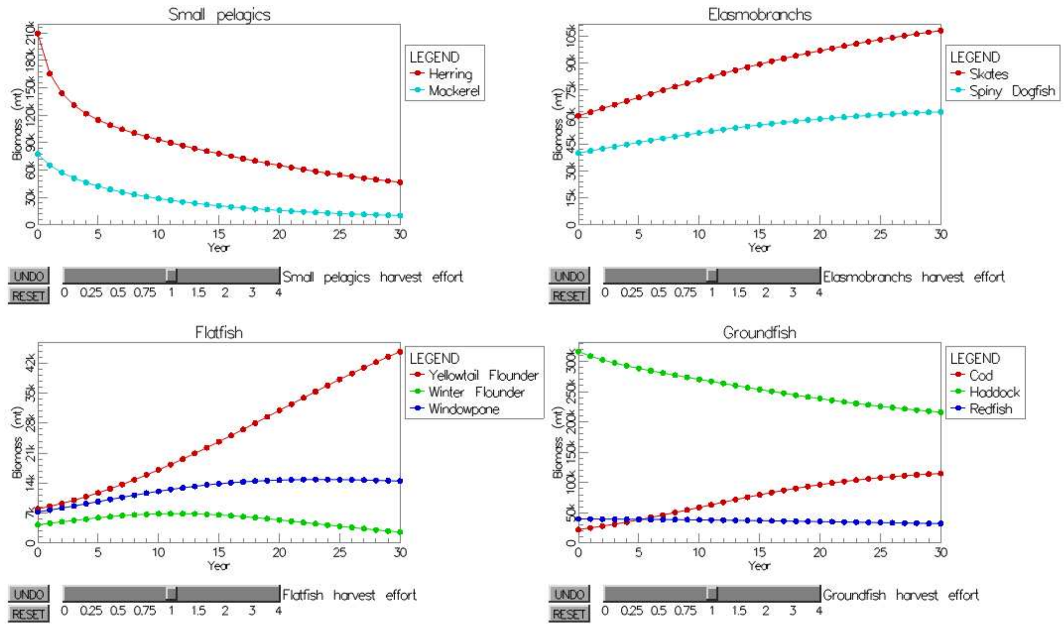


Figure 2-1: The "four panel" view of our MS-PROD visualization.

for displaying biomasses. This view is called the "four panel" view and is shown in Figure 2-1.

The major advantage of this "four panel" view is that comparison of species within a group is easy. For MS-PROD, there are only two or three species per functional group, so the line chart for each group tends to not suffer from occlusion problems. Direct and indirect effects of changes in harvest effort are easily differentiated with the "four panel" approach—e.g., if the user adjusts the effort slider only for elasmobranchs, yet sees the biomasses change on the groundfish chart, then the user can begin to understand there is some kind of relationship between elasmobranchs and groundfish.

Small Multiples View

The alternative to the "four panel" view is to view each species on its own plot, which we call the "small multiple" view. The main purpose of this view was to support the addition of arc graph connections between species, as are discussed later in Section 2.3.1. Shown in Figure 2-2, each plot is sorted and colored according to functional group membership.

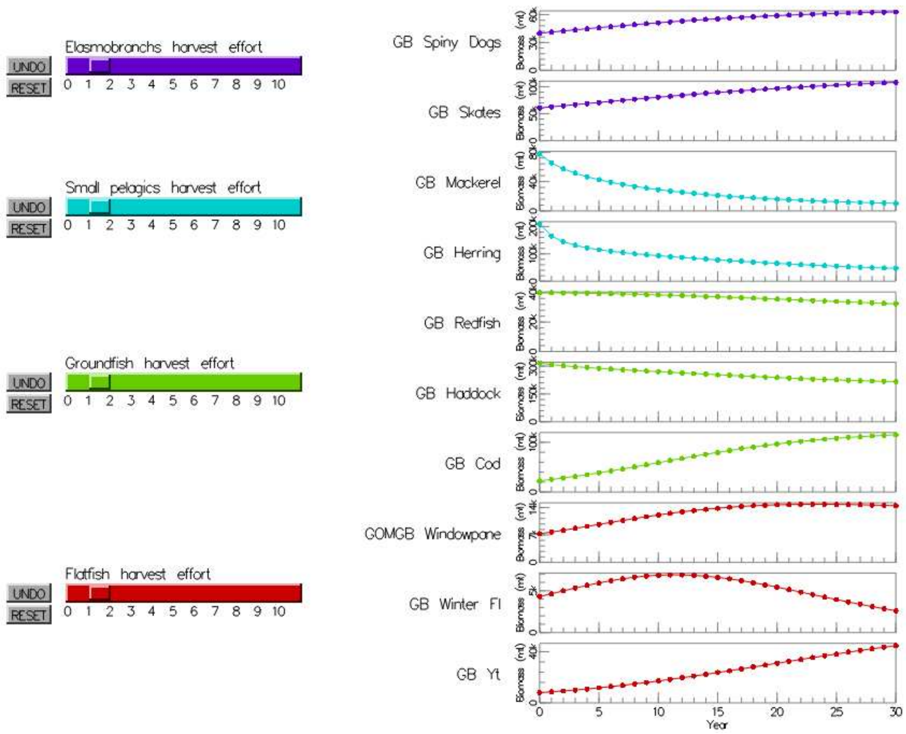


Figure 2-2: The "small multiples" view of our MS-PROD visualization.

The harvest effort sliders are also colored by functional group and positioned near the plots of the corresponding group. This allows for the ability to differentiate between direct and indirect effects of changes in harvest effort.

With each species on its own plot, it is much easier to interpret the biomass predictions of an individual species, since the y-dimension of a plot needs to be scaled to the data of one species only. It is also easier to perceive increases or decreases in biomass because no species suffer from the "flattening" that can occur when a series is displayed on the same plot as a series that has significantly higher values. On the other hand, this makes biomass comparison between species somewhat difficult because all the y-axis scales are different and the user must either refer to the y-axis labels or hover over a specific point on a chart, which causes a label to appear, in order to determine the absolute value of the biomass at a point in time.

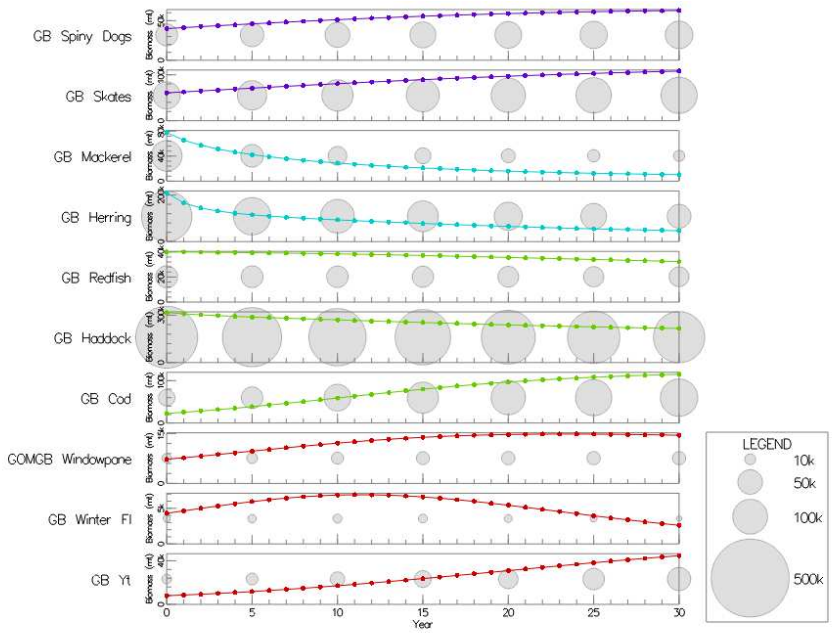


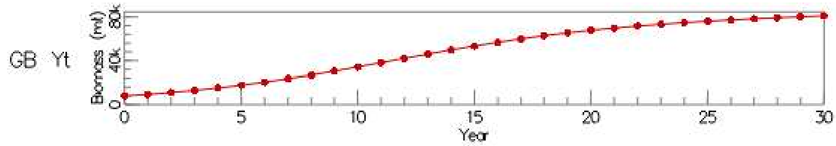
Figure 2-3: Absolute biomass indicators overlaying the "small multiples" view.

Absolute Biomass Indicators

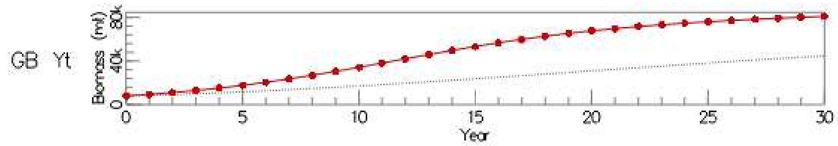
Absolute biomass indicators, seen in Figure 2-3, were introduced to show how biomass changes over time by showing the absolute biomass of the population as the area of a circle. This makes comparison across species possible. To avoid occlusion, these indicators are drawn every five years within the thirty-year time span.

2.2 Visualization of Change

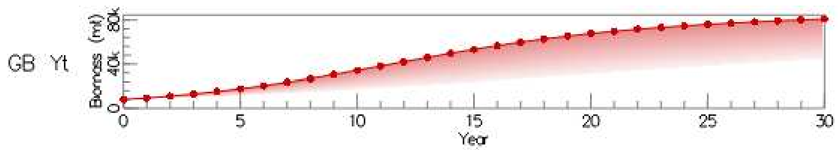
In order for modelers and other stakeholders to understand and compare decisions, the ability to perceive changes in biomass resulting from changes in the fishing effort is required. Therefore, we have introduced a feature that allows the user to compare the forecast effects of a change in fishing effort compared to the forecasts of the *status quo* [baseline]. There are three alternatives for displaying forecast differences with the baseline. The first is to simply have the biomass plots change instantaneously as the harvest effort sliders are adjusted, as in Figure 2-4a. In this case, the user must remember the previous curve in order to judge a



(a) Change shown by interaction.



(b) The status quo is shown as the dotted gray line.



(c) The area between the status quo graph and the new forecast is shaded.

Figure 2-4: The three options for depicting change between current biomass predictions and baseline predictions.

change. The second is a dotted gray line which shows the forecast of the baseline in addition to the current forecast, shown in Figure 2-4b. The third is a shaded area originating from the curve of the current forecast that diminishes in opacity as it approaches the curve of the baseline forecast, as in Figure 2-4c.

Figure 2-5 shows all of the time series in the "small multiples" view with the effort sliders and the blended change option enabled. Again, the blended area in a line chart is colored from the current biomass line to the line as it was when the baseline was set; a colored area above the line indicates the biomass declined, while a colored area beneath represents the biomass increased—e.g., the "skate" population declined dramatically with the new effort values, the "winter flounder" population increased due to the changes, and "redfish" seemed to be unaffected.

Underneath each slider, a colored rectangle—if present—indicates differences from the baseline effort settings. Blue indicates the effort value has been increased since the baseline was set—e.g., again in Figure 2-5, the effort for "elasmobranchs" was originally set to 1.0

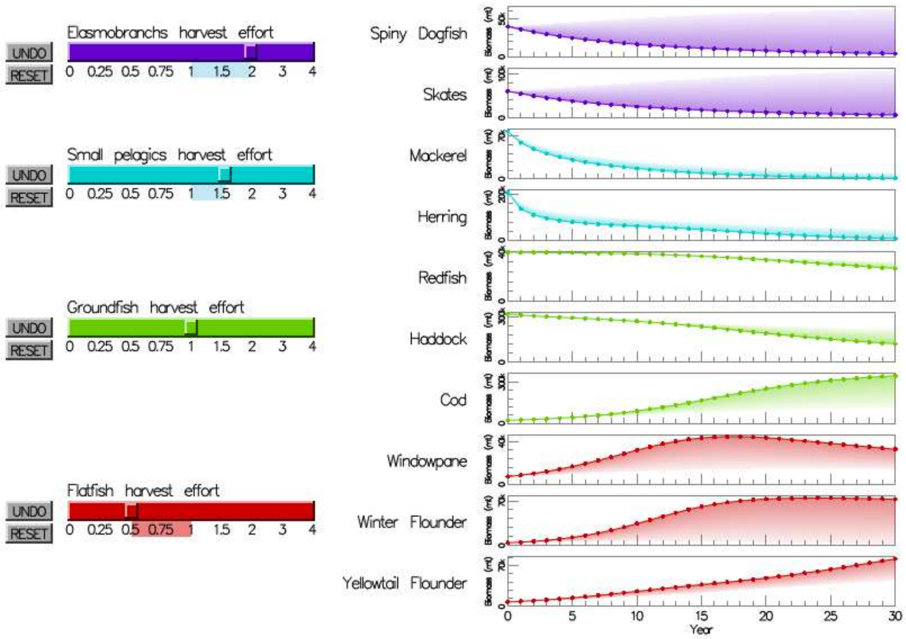


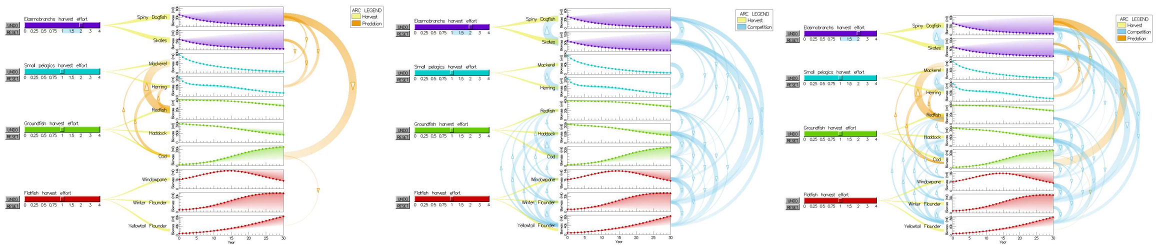
Figure 2-5: Showing change between different effort values.

and now it is approximately 2.0. Red represents the effort value has been decreased since the baseline was set—e.g., the effort for "flatfish" was originally set to 1.0 and now it is approximately 0.5.

This feature is available in both the "four panel" and "small multiples" views. The baseline effort setting can be defined at any time with a simple button at the top of the screen. This resets the baseline as the current slider settings. Buttons are also available to undo or reset changes to the effort values.

2.3 Visualization of Causal Relationships

Understanding of the model requires understanding of the underlying relationships between species—namely, *predation* and *competition*—and *harvest*. As defined earlier, predation is when one species consumes another and competition accounts for any way species might impact another in a way that is not predation. Harvest is the process of humans catching fish in the wild. Our interface must explain to users which species impact each other, how harvesting impacts the species, and the magnitude of those relationships. We have chosen



(a) Predation arcs. (b) Competition arcs. (c) Both types of arcs.

Figure 2-6: Static arcs drawn between species charts to represent relationships.

to illustrate these relationships with a node-link diagram, where the nodes are the time series and effort sliders.

2.3.1 Inter-Species Relationships

The inter-species relationships are illustrated with arc diagram network visualizations between time series, as in Figure 2-6, in the small multiple view. In the input parameter file, predation and competition coefficient matrices are defined to represent the relationship between each pair of species. Each non-zero coefficient is represented with an arc connecting the time series plots of the corresponding species. This is only available in the small multiples view because multiple species are represented in each line chart of the four panel view.

The arc diagrams are easily combined with the small multiples; in fact, small multiples were chosen because they would work with an arc diagram. A separate node-link diagram, such as one with a force-directed layout, may have been confusing because it would require the user to mentally associate the randomly distributed nodes in the network with line charts. An arc diagram does not suffer from this problem since, in our case, the time series themselves are the nodes. Furthermore, the arc diagram enhances the entire visualization without occluding the time series. Arc diagrams are also well suited to smaller datasets with clusters of nodes, which applies to our dataset since the fish are segmented into functional groups.

Directionality

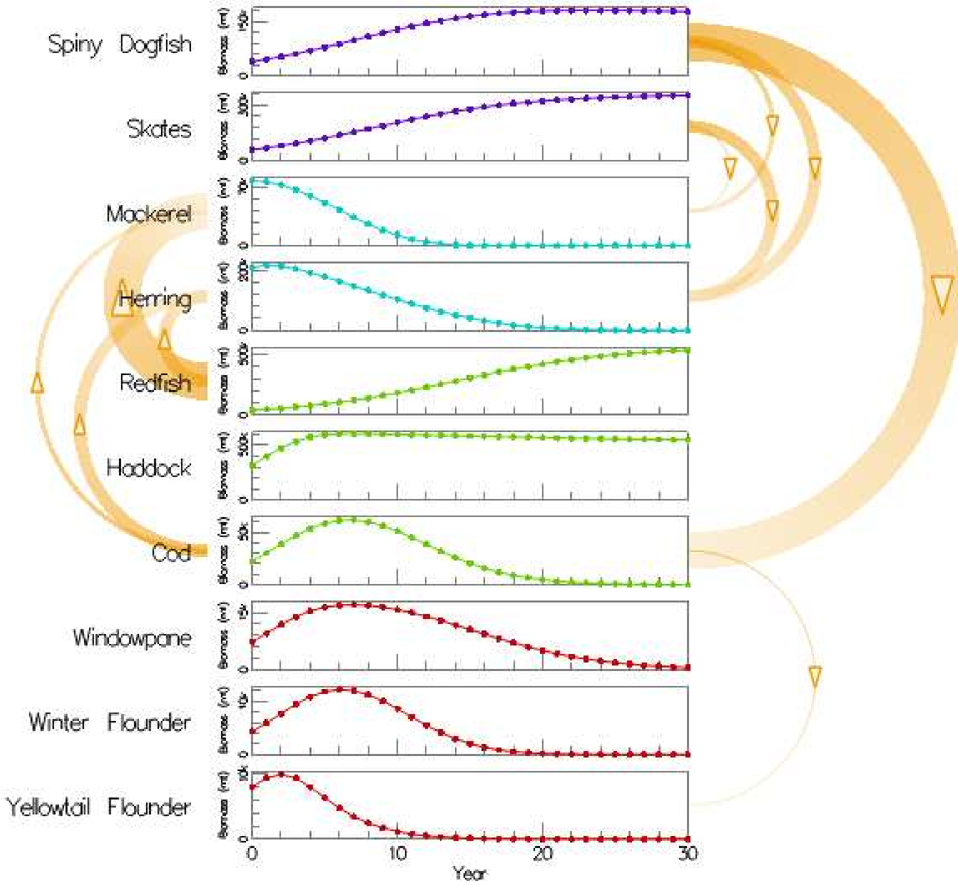


Figure 2-7: Static predation arcs with directionality indicated by a triangular arrow, fading color intensity, and clockwise direction.

The predation and competition relationships are directed, where one species of fish is the "source" and the other species is the "recipient." Therefore, our arcs have been drawn with fading opacity to indicate the direction, since Holten and van Wijk recommended a dark-to-light representation for an intensity-based cue (Holten and van Wijk, 2009). Additionally, our arcs follow a clockwise direction; arcs on the right-hand side are all directed downward, whereas arcs on the left-hand side are directed upward. The arcs can also be animated with a pulsing effect that originates from the source species to the recipient species. These directional cues are necessary because there may be reciprocal "Fish A affects Fish B" and

"Fish B affects Fish A" relationships, especially for the competition type of relationship, so arcs must be drawn on both sides of the time series, as seen in Figure 2-6. Additionally, triangular marks have been drawn in the middle of the arcs to point from the source species to the recipient species. These three directionality cues can be seen in Figure 2-7.

Arc Type

The interface provides a few options for viewing the inter-species relationships. Firstly, users have the ability to view either predation as in Figure 2-6a, competition as in Figure 2-6b, or both as in Figure 2-6c. Secondly, the arcs can be viewed statically, dynamically without animation, or dynamically with animation; these three arc types are described below. Regardless of the selected arc type, all predation arcs are colored orange and all competition arcs are colored sky-blue. The user can mouse over a particular arc, which causes that arc to highlight—while the other arcs fade—and displays a label which spells out the relationship in words and shows the input parameter matrix original coefficient—e.g., "Skates compete with Winter Flounder (0.6)."

Static With the static style of the arcs, all arcs are drawn at all times, as shown in all three subfigures of Figure 2-6. The width of an arc corresponds to the magnitude of the relationship, as defined in the predation coefficient matrix or the competition coefficient matrix in the original parameter file. A benefit of this type of arc is that it is possible to see all interactions between the species at all times. However, the downside is that viewing all arcs at once can be overwhelming because the display becomes somewhat cluttered.

Dynamic Dynamic arcs, shown in Figure 2-8, were motivated by need to (1) clarify the sized change in a causal influence path and (2) eliminate or at least reduce the visual clutter created by the static arcs. We considered the fact that many users might use our interface as a means of comparing two biomass forecasts that resulted from changes in fishing effort values. Such users might wonder what specifically caused the differences between the two forecasts, especially if there are indirect, unintuitive effects. Therefore, we designed dynamic arcs, which are drawn only selectively to help explain the differences in between the current forecast and the baseline forecast.

The dynamic arcs are simplification of a time series of effects, where the width of each arc is proportional to a weight w:

$$w = \frac{100000}{r_0 + 100000} \cdot (s_{30} - s'_{30}) \tag{2.1}$$

where r_0 represents the initial biomass of the recipient species, s_{30} is the biomass at year 30 for the source species according to the current forecast, and s'_{30} is the biomass at year 30 for the source species according to the baseline forecast. If the source species biomass at year 30 did not change between the species, then $(s_{30} - s'_{30})$ equals zero, resulting in a w of zero, so the arc will not be drawn. In other words, the source species must have experienced change between the two forecasts in order to possibly explain a change that occurred in the recipient species; if there was no such change, then no arc is drawn. The width of the arc increases as the difference between the two forecasts for the source at year 30 increases. The arc width is inversely proportional to the size of the recipient species because relatively small species will have a low impact on large species.

To summarize, the width of an arc in dynamic mode is (a) proportional to the difference for the final values of the baseline and current forecasts for the source species and (b) inversely proportional to the initial biomass of the recipient species. Therefore, no arcs are drawn when the current forecast and the baseline forecast are the same. As the sliders are adjusted either positively or negatively, more arcs may appear to assist in explaining indirect effects. An arc grows in width as the source species experiences more dramatic change between the forecasts.

Additionally, the weight w is negative if $s_{30} < s'_{30}$. This is significant in explaining *how* the change in the source species affects the recipient species—i.e., was the change that the source species experienced "good" or "bad" from the perspective of the recipient species? Both relationships, competition and predation, inhibit the growth of the recipient species because the source species either consumes the recipient species itself or its resources.

Therefore, if the source species biomass decreases, then the recipient species biomass may be able to grow more. Of course, other species may be dampening the growth of the recipient species, so the recipient species may not necessarily experience a growth, but the potential exists. Conversely, if the source species biomass increases, then the recipient species may suffer more and experience a decrease in biomass. We chose to use plus signs (+) for the cases where w is negative—i.e., when the source species declines between forecasts which is "good" for the recipient species—and minus sign (−) for the cases where w is positive—i.e., when the source species increases between forecasts which is "bad" for the recipient species—as Kadaba et al. used in their static causal visualizations (Kadaba et al., 2007). Plus signs were drawn in black and minus signs were drawn in white with a black outline to allow for some redundant coding. Several signs are drawn along each arc to allow the user to easily determine the signage of a dynamic arc.

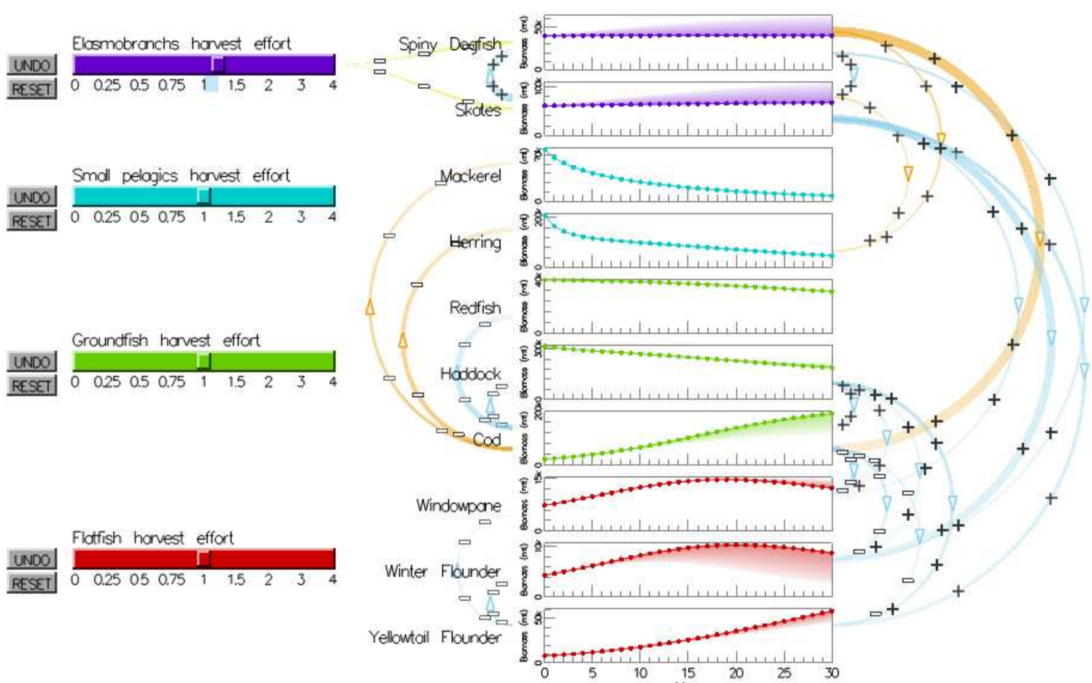


Figure 2-8: Dynamic arcs drawn as a result of slightly increasing the fishing effort on elasmobranchs.

Again, Figure 2-8 shows dynamic arcs which resulted from slightly increasing the fishing effort on elasmobranchs from 1.0 to about 1.25. The spiny dogfish and skate biomasses both decreased as a result of the increase in fishing, as is indicated by the shaded area between the baseline forecast and the current forecast. This is good from the perspective of all of the fish that either spiny dogfish or elasmobranchs predate on or compete with, therefore all of the arcs drawn from these two species show plus signs. For example, spiny dogfish predate on cod, so the arc between them has plus signs. The cod biomass increased, which seems to corroborate with the plus signs drawn on the arc. There are also indirect effects that the arcs help to explain. Cod competes with windowpane, so minus signs are drawn on the arc from cod to windowpane, which helps to explain the slight decrease in the windowpane biomass.

We have realized it is possible for dynamic arcs to be misinterpreted. Users may incorrectly assume that absence of an arc indicates that the relationship is no longer present—e.g., if the predation arc from spiny dogfish to cod is not present, then the spiny dogfish are not eating the cod. However, we believed that users would understand the true meaning of the arcs after a brief explanation—e.g., the arc between spiny dogfish and cod is no longer being drawn because there are no changes in the cod population that might be explained by the spiny dogfish. When informally showing the model to users during and after development, we found that users properly interpreted the arcs after a demonstration or interaction with the model. Another possible criticism of the dynamic arcs is that sometimes many arcs are drawn and it can become cluttered, such as when several sliders are pulled to the extremes. However, generally less arcs are drawn in dynamic mode than static mode even in extreme cases. Also, fishery managers may potentially be more interested in comparing the long-term effects of slight adjustments to fishing quota than dramatic adjustments that either eliminate fishing completely or wipe out entire stocks through overfishing.

Dynamic with Animation The third and final type for viewing the inter-species arcs is dynamic arcs with animation, seen in Figure 2-9. In this mode, the rules for the appearance of arcs and their width is the same as in the non-animated dynamic version. However, the

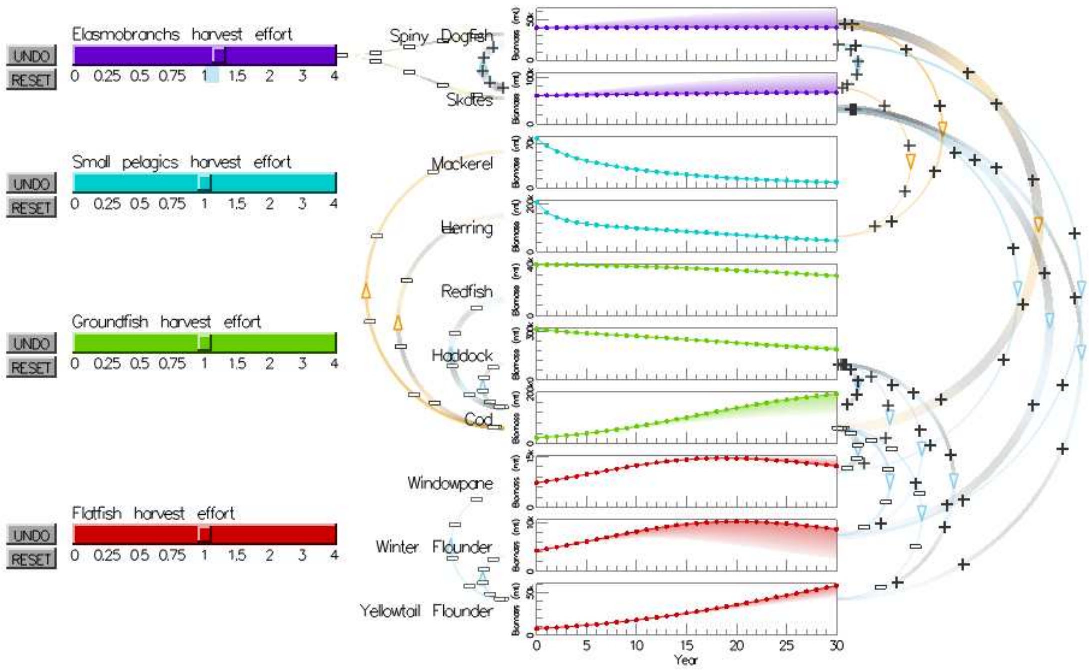


Figure 2-9: The same scenario as in Figure 2-8, though with animated dynamic arc.

plus (+) or minus (−) signs travel along the arc from the source species to the recipient species. Additionally, the color of the arc alternates between gray and either blue for competition or orange for predation. The alteration in color also moves from source to recipient. Both of these cues help to highlight the directionality of the arc.

2.3.2 Harvest Influences

The other type of relationship that must be elucidated by our visualization is the harvest relationship. While it is clear what the harvest effort value is for a particular slider, it could perhaps be clearer which species were directly affected by the harvest and how. Therefore, a spline curve is drawn between a harvest type and a fish species, which corresponds to a fishing effort slider and a biomass small multiple, respectively, in our interface. These links are drawn using Hermite spline curves which are colored yellow. There are three styles for determining the width of these spline curves, which correspond to the styles for inter-species arcs: static, dynamic, and dynamic with animation. The controls for setting the inter-species arc style also set the harvest spline curve style.

Static

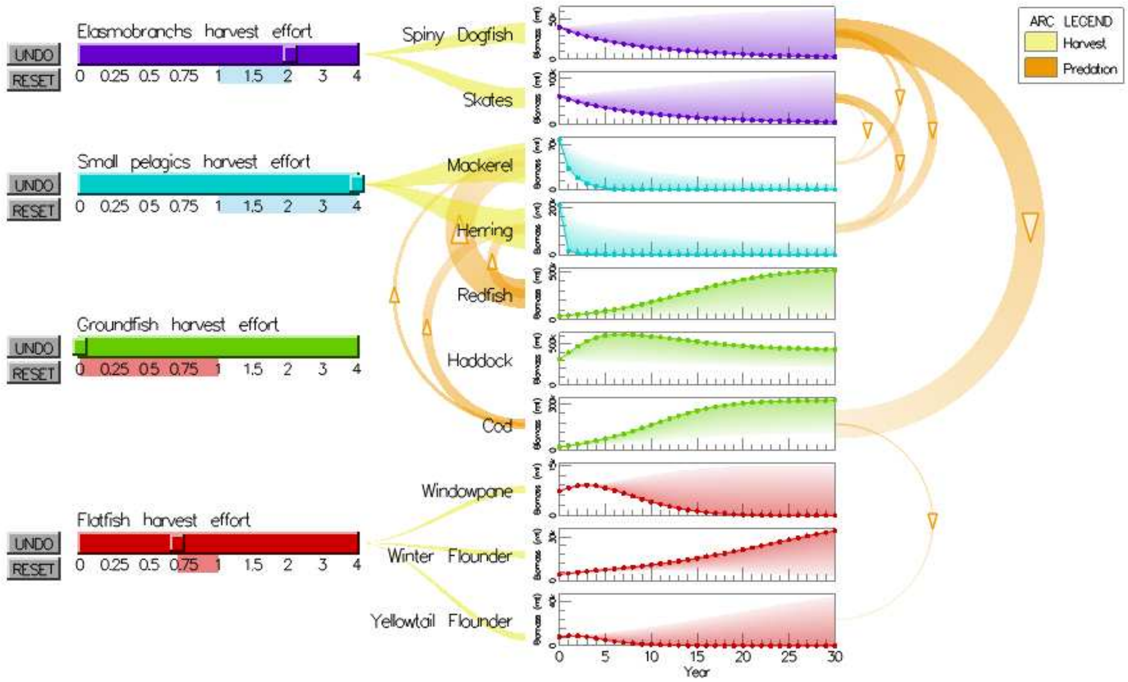


Figure 2-10: Static harvest spline curves in yellow between effort sliders and the small multiples.

In static mode, the harvest spline curve's width is directly proportional to the current value of the harvest slider, as in Figure 2-10. Therefore, when harvest effort for a specific functional group is set to zero (as it is for "Groundfish"), then no harvest spline curves are drawn emerging from that harvest effort slider. Likewise, the spline curves are their thickest when the matching effort slider is set to the maximum value of four (as it is for "Small pelagics").

Dynamic

Dynamic harvest splines can be seen in Figure 2-8. The width of a harvest spline in dynamic mode is proportional to the difference between the current value of the effort slider and the baseline value of the effort slider. In other words, the width depends on a weight w:

$$w = E - E' \tag{2.2}$$

where E is the current effort and E' is the baseline effort value for the functional group. This is to help highlight and explain the differences between the baseline and current forecasts. Therefore, when the harvest effort for a functional group has not been changed from the baseline, w is zero and no harvest splines are drawn from that functional group's slider.

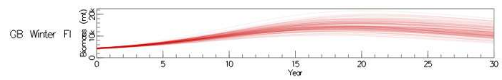
When $E' > E$ (i.e., when the harvest effort slider has been decreased), w is negative. Therefore, as with dynamic arcs for the inter-species relationships, the harvest splines have signage, which we interpreted from the perspective of the recipient—i.e., the fish being harvested. Thus, when w is negative, black plus signs (+) are drawn along the spline. Conversely, when w is positive, white minus signs (−) are drawn along the spline. From the perspective of the fish, it is "bad" to be fished more and "good" to be fished less. The meaning and design of these signs are the same as for the inter-species arcs.

Dynamic with Animated

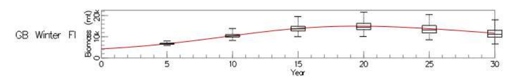
Dynamic harvest spline curves, as in Figure 2-9, are drawn following the rules for non-animated spline curves, though with animation added. The animation is similar to the animation for inter-species arcs: the signs travel from the harvest effort slider to the species biomass small multiple chart and the color is alternated with gray to create a pulsing effect. The intent was to give a clearer indication of the direction of the causal relationship.

2.4 Visualization of Uncertainty

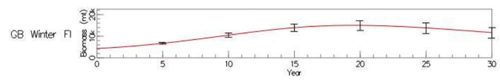
Since models are simplifications of reality, their output is best understood as a range of expected values. It is possible that a representation of uncertainty may aid decision making. To add uncertainty visualization to the MS-PROD model, our interface can perform Monte Carlo simulations by randomly varying the non-zero input parameter values ±10% using a normal distribution and computing 100 separate runs of the model.



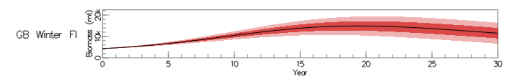
(a) Multi-line.



(b) Box plots.



(c) Error bars.



(d) Error bands.

Figure 2-11: The four methods for visualizing uncertainty of MS-PROD model simulations.

The resulting uncertainty can be displayed in four styles. First, a multi-line option shows the uncertainty by drawing one semi-transparent line for each run of the Monte Carlo simulation, as in Figure 2-11a. This was inspired by Cox et al.'s new method for displaying hurricane tracks (Cox et al., 2013). The second option displays a traditional box plot every five years, as in Figure 2-11b. Third, in Figure 2-11c, a summary line is drawn, with bars every five years. Lastly, Figure 2-11d shows a summary line in solid black line with bands.

In all options but the multi-line option, the user has the ability to display different statistical data in the selected style. Users can select between mean and median for the "summary" lines. Box plots, error bars, and error bands can represent either quartiles with minimum and maximum data values or two standard deviations. This allows the user to easily explore the distribution of data under different representation with a few clicks of the mouse. The multi-line option may appeal to less scientific users, while the other more traditional representations of statistical data may appeal to advanced users, such as modelers or fishery managers, but all users have all options at their disposal.

CHAPTER 3

Evaluation

The intention behind this work has been to develop an interactive visualization that effectively portrays the MS-PROD model and its implications. More specifically, we are interested in how different visualization alternatives enhance a user's understanding of the complex relationships between the fish species and their effects. In other words, is there a benefit to using dynamic, animated arcs—the most complicated representation of the relationships—over another method or even displaying no arcs at all? To investigate this question, we designed and conducted a user study to the performance of different arc depiction alternatives.

3.1 Method

The study was conducted at a screened-off table in a student union building at the University of New Hampshire. A paid undergraduate research assistant conducted the study.

The research assistant explained the MS-PROD model and our visualization, and showed a training example before leaving the participant to the experiment. Feedback was received only during the traiing phase. A single experiment lasted approximately ten minutes.

Responses from the study were graded by two paid undergraduate research assistants.

3.2 Participants

There were 80 participants who took part in the study, all of which were recruited by a poster affixed to the backside of the privacy screen. Participants were voluntary and were compensated with a pack of pens or a notebook. They were required to read and sign an IRB consent form before participating in the study.

3.3 Apparatus

We conducted the experiment using a standard Dell laptop with an extra Dell monitor. The window with the model visualization was maximized on the extra screen, while the window with the experiment questions was maximized on the laptop screen. Participants used the mouse to interact with the model visualization and recorded their answers using the laptop keyboard.

3.4 Experimental Conditions

Each participant conducted the experiment task for only one of the four conditions:

(A) No arcs

(B) Static arcs

(C) Dynamic arcs without animation

(D) Dynamic arcs with animation

Explanations and training phases were tailored according to the experimental condition—e.g., arcs were explained only for conditions B, C, and D; dynamic arcs were explained only for conditions C and D.

3.5 Task

Initially, all fishing effort sliders were set to the value of one. Participants were instructed to increase or decrease the fishing effort of a specific functional group, e.g.:

> *Using the sliders, double the harvest effort on elasmobranchs.*

Next, the participants were asked to answer one or more questions of the form, "*What was the effect on (fish species)?*" The questions were designed so that sometimes the fish species was a member of the function group for which the effort was just adjusted, while other times the fish species was not a member of that functional group, e.g.:

What is the effect on haddock?

Users answered this question with one of five options from a drop-down menu:

- *Increased a lot*
- *Increased a little*
- *Stayed about the same*
- *Decreased a little*
- *Decreased a lot*

Finally, the user was asked, "*Why? [Try to explain in no more than three sentences.]*" A large text box was provided for the participant to type a response. If this question was the last question in its set, then the sliders are all reset to one and a new instruction is given for the next set of questions.

In total, there were three instructions and eight questions. All participants were given the same instructions and asked the same questions in the same order, regardless of condition.

3.6 Results

ABC

BIBLIOGRAPHY

Dan Bricklin and Bob Frankston. VisiCalc: Information from its creators, Dan Bricklin and Bob Frankston. http://bricklin.com/visicalc.htm, 1999. [Online; accessed 2013-10-07].

Norman L. Christensen, Ann M. Bartuska, James H. Brown, Stephen Carpenter, Carla D'Antonio, Rober Francis, Jerry F. Franklin, James A. MacMahon, Reed F. Noss, David J. Parsons, Charles H. Peterson, Monica G. Turner, and Robert G. Woodmansee. The Report of the Ecological Society of America Committee on the Scientific Basis for Ecosystem Management. *Ecological Applications*, 6(3):665–691, August 1996. ISSN 10510761. doi: 10.2307/2269460. URL http://dx.doi.org/10.2307/2269460.

William S. Cleveland, Marylyn E. McGill, and Robert McGill. The shape parameter of a two-variable graph. *Journal of the American Statistical Association*, 83(402):pp. 289–300, 1988. ISSN 01621459. URL http://www.jstor.org/stable/2288843.

Jonathan Cox, Donald House, and Michael Lindell. Visualizing uncertainty in predicted hurricane tracks. *International Journal for Uncertainty Quantification*, 3(2):143–156, 2013. ISSN 2152-5080.

Robert L. Eberlein and David W. Peterson. Understanding models with Vensim™. *European Journal of Operational Research*, 59(1):216–219, May 1992. URL http://ideas.repec.org/a/eee/ejores/v59y1992i1p216-219.html.

Sarah K. Gaichas and Robert C. Francis. Network models for ecosystem-based fishery analysis: a review of concepts and application to the Gulf of Alaska marine food web. *Canadian Journal of Fisheries and Aquatic Sciences*, 65(9):1965–1982, 2008. URL http://www.ingentaconnect.com/content/nrc/cjfas/2008/00000065/00000009/art00014.

Robert J. Gamble and Jason S. Link. Analyzing the tradeoffs among ecological and fishing effects on an example fish community: A multispecies (fisheries) production model. *Ecological Modelling*, 220(19):2570 – 2582, 2009. ISSN 0304-3800. doi: 10.1016/j.ecolmodel.2009.06.022. URL `http://www.sciencedirect.com/science/article/pii/S0304380009003998`.

Mohammad Ghoniem, Jean-Daniel Fekete, and Philippe Castagliola. A comparison of the readability of graphs using node-link and matrix-based representations. In *Proceedings of the IEEE Symposium on Information Visualization*, INFOVIS '04, pages 17–24, Washington, DC, USA, 2004. IEEE Computer Society. ISBN 0-7803-8779-3. doi: 10.1109/INFOVIS.2004.1. URL `http://dx.doi.org/10.1109/INFOVIS.2004.1`.

B. Grad. The creation and the demise of VisiCalc. *Annals of the History of Computing, IEEE*, 29(3):20–31, 2007. ISSN 1058-6180. doi: 10.1109/MAHC.2007.4338439.

S. Havre, B. Hetzler, and L. Nowell. ThemeRiver: visualizing theme changes over time. In *IEEE Symposium on Information Visualization*, pages 115–123, 2000. doi: 10.1109/INFVIS.2000.885098.

Jeffrey Heer, Michael Bostock, and Vadim Ogievetsky. A tour through the visualization zoo. *Commun. ACM*, 53(6):59–67, June 2010. ISSN 0001-0782. doi: 10.1145/1743546.1743567. URL `http://doi.acm.org/10.1145/1743546.1743567`.

Danny Holten and Jarke J. van Wijk. A user study on visualizing directed edges in graphs. In *Proceedings of the SIGCHI Conference on Human Factors in Computing Systems*, CHI '09, pages 2299–2308, New York, NY, USA, 2009. ACM. ISBN 978-1-60558-246-7. doi: 10.1145/1518701.1519054. URL `http://doi.acm.org/10.1145/1518701.1519054`.

G. J. Hunter and M. F. Goodchild. Managing uncertainty in spatial databases: Putting theory into practice. *URISA*, pages 1–14, 1993.

Waqas Javed, Bryan McDonnel, and Niklas Elmqvist. Graphical perception of multiple time series. *IEEE Transactions on Visualization and Computer Graphics*, 16(6):927–934, November 2010. ISSN 1077-2626. doi: 10.1109/TVCG.2010.162. URL http://dx.doi.org/10.1109/TVCG.2010.162.

Nivedita Kadaba, Pourang Irani, and Jason Leboe. Visualizing causal semantics using animations. *IEEE Transactions on Visualization and Computer Graphics*, 13(6):1254–1261, 2007. doi: 10.1109/TVCG.2007.70618.

Donald E. Knuth. *The Stanford GraphBase: a platform for combinatorial computing*. ACM, New York, NY, USA, 1993. ISBN 0-201-54275-7.

P. H. Leslie and J. C. Gower. The properties of a stochastic model for the predator-prey type of interaction between two species. *Biometrika*, 47:219–301, 1960.

Alfred J. Lotka. The frequency distribution of scientific productivity. *Journal of the Washington Academy of Sciences*, 16(12):317–323, 1926.

Albert Michotte and Georges Thinés. La causalité perceptive. *Journal de Psychologie Normale et Pathologique*, pages 9–36, 1963.

United States. National Marine Fisheries Service. Ecosystem Principles Advisory Panel. *Ecosystem-based fishery management: a report to Congress*. U.S. Dept. of Commerce, National Oceanic and Atmospheric Administration, National Marine Fisheries Service, 1999. URL http://books.google.com/books?id=bLoxmLSz6IYC.

William Playfair. *The Commerical and Political Atlas: Representing, by means of stained copper-plate charts, the progress of the commerce, revenues, expenditure and debts of England during the whole of the Eighteenth Century*. 1786.

Kristin Potter, Joe Kniss, Richard Riesenfeld, and Chris R. Johnson. Visualizing summary statistics and uncertainty. *Computer Graphics Forum (Proceedings of Eurovis 2010)*, 29 (3):823–831, 2010.

M. L. Rosenzweig and R. H. Macarthur. Graphical Representation and Stability Conditions of Predator-Prey Interactions. *The American Naturalist*, 97(895):209+, January 1963. doi: 10.1086/282272. URL http://dx.doi.org/10.1086/282272.

Takafumi Saito, Hiroko Nakamura Miyamura, Mitsuyoshi Yamamoto, Hiroki Saito, Yuka Hoshiya, and Takumi Kaseda. Two-tone pseudo coloring: Compact visualization for one-dimensional data. *IEEE Symposium on Information Visualization*, 0:23, 2005. ISSN 1522-404x. doi: http://doi.ieeecomputersociety.org/10.1109/INFOVIS.2005.35.

Milner B. Schaefer. A study of the dynamics of the fishery for yellowfin tuna in the eastern tropical pacific ocean. *Inter-American Tropical Tuna Commission Bulletin*, 2(6):243–285, 1957.

Lisa Tweedie, Robert Spence, Huw Dawkes, and Hua Su. The influence explorer. In Jim Miller, Irvin R. Katz, Robert L. Mack, and Linn Marks, editors, *CHI 95 Conference Companion*, pages 129–130. ACM, 1995. ISBN 0-89791-755-3. URL http://dblp.uni-trier.de/db/conf/chi/chi1995c.html#TweedieSDS95.

Vito Volterra. Fluctuations in the abundance of a species considered mathematically. *Nature*, 118:558–560, 1926. doi: 10.1038/118558a0.

Colin Ware. Perceiving complex causation through interaction. In *Proceedings of the Symposium on Computational Aesthetics*, CAE '13, pages 29–35, New York, NY, USA, 2013. ACM. ISBN 978-1-4503-2203-4. doi: 10.1145/2487276.2487279. URL http://doi.acm.org/10.1145/2487276.2487279.

Colin Ware, Eric Neufeld, and Lyn Bartram. Visualizing causal relations. *IEEE Symposium on Information Visualization: Late Breaking Hot Topics*, pages 39–42, 1999.

Martin Wattenberg. Arc diagrams: Visualizing structure in strings. In *Proceedings of the IEEE Symposium on Information Visualization*, INFOVIS '02, pages 110–116, Washington, DC, USA, 2002. IEEE Computer Society. ISBN 0-7695-1751-X. URL http://dl.acm.org/citation.cfm?id=857191.857733.

M. Weber, M. Alexa, and W. Muller. Visualizing time-series on spirals. In *IEEE Symposium on Information Visualization*, pages 7–13, 2001. doi: 10.1109/INFVIS.2001.963273.