DOI: 10.1111/cgf.12907
Eurographics Conference on Visualization (EuroVis) 2016
K.-L. Ma, G. Santucci, and J. van Wijk
(Guest Editors)

Volume 35 (2016), Number 3

Dynamic Change Arcs to Explore Model Forecasts

C. St. Jean[1], C. Ware[1], and R. Gamble[2]

[1]University of New Hampshire, Durham, New Hampshire, United States
[2]NOAA Fisheries, Woods Hole, Massachusetts, United States

Abstract

In many planning applications, a computational model is used to make predictions about the effects of management or engineering decisions. To understand the implications of alternative scenarios, a user typically adjusts one or more of the input parameters, runs the model, and examines the outcomes using simple charts. For example, a time series showing changes in productivity or revenue might be generated. While this approach can be effective in showing the projected effects of changes to the model's input parameters, it fails to show the mechanisms that cause those changes. In order to promote understanding of model mechanics using a simple graphical device, we propose dynamic change arcs. Dynamic change arcs graphically reveal the internal model structure as cause and effect linkages. They are signed to show both positive and negative effects. We implemented this concept using a species interaction model developed for fisheries management based on a system of Lotka-Volterra equations. The model has 10 economically important fish species and incorporates both predation and competition between species. The model predicts that changing the catch of one species can sometimes result in changes in biomass of another species through multi-step causal chains. The dynamic change arcs make it possible to interpret the resulting complex causal chains and interaction effects. We carried out an experiment to evaluate three alternative forms of arcs for portraying causal connections in the model. The results show that all linkage representations enabled participants to reason better about complex chains of causality than not showing linkages. However, none of them were significantly better than the others.

Categories and Subject Descriptors (according to ACM CCS): H.5.2 [Information Interfaces and Presentation]: User Interfaces—Interactions Styles

1. Introduction

In more and more design applications, a model is used to make predictions and its results are displayed graphically, usually by means of one or more time series plots. The classic example of this is the business model spreadsheet. Such models use various parameters representing costs of production, product marketing, distribution, and so on to generate a profit forecast projected out several years. These models enable business executives to explore *what if* scenarios. For example, what if the cost of raw materials rises by a certain amount? What if a lower rate of interest can be obtained for a business loan? Changing a single number in the spreadsheet can produce a new forecast. VisiCalc, the first "killer app" for personal computers, supported exactly this kind of activity and is often credited with the rise of the personal computer in business [Ram80].

Spreadsheet applications like VisiCalc and its successor Microsoft Excel are powerful because they show the *effect* of changing parameter values instantaneously, but spreadsheets fail to show the *reasons* for those changes. The only way of discovering the chain of causal linkages that led to a particular outcome is to delve into the spreadsheet code. Our work aims to provide a partial solution to this problem in the form of *dynamic change arcs*. These are graphical devices designed to show internal model linkages and enable the user to understand both the consequences of a change in a model parameter and the causal linkages leading to those consequences. We implemented our design ideas in an interactive visualization of an ecosystem-based model built for fisheries management and we carried out a study where participants used this visualization to produce explanations of the effects of changes in fishing practice.

2. Prior research

There have been a number of different approaches to the representation of causality for data visualization. It is useful to classify these according to how causal linkages are displayed and how the effects are shown in response to changes of the model parameters. (A summary of the following review is given in Table 1.)

Work by Igarishi et al. [IMCZ98] aimed to reveal the causal network structure that is implicit in many spreadsheet models. When the user moused-over a spreadsheet cell, lines emerged to show incoming and outgoing relationships with other cells. This provided a kind of interactive data flow graph. The size of the effect was shown by a change in the numbers in destination cells, or by bars resembling those found in bar graphs.

The Influence Explorer and the Attribute Explorer were experi-

Paper	**Representation of causal linkages**	**Representation of causal effects**	**Evaluation**
Kadaba et al. 2007 [KIL07]	Lines with animated 'bullets'; size is effect strength	Target node changes size	Both simple and complex relations
Igarashi et al. 1998 [IMCZ98]	Animated lines linking spreadsheet cells; magnitude not shown	Change in numbers in cells	None
Ware et al. 1999 [WNB99]	Animated waves or balls	Target nodes change size instantaneously	Temporal contingencies; simple and causal effects
Zapata-Rivera et al. 1999 [ZRNG99]	Animated temporal order; lines connecting cells on a spreadsheet	Size or color change	None
Tweedie et al. 1995 [TSDS95]	None	Distributions of outcomes change	None
Neufeld et al. 2006 [NSK06]	Animated links: moving waves or streaks of light	Size or color change	None
Neufeld and Kristtorn 2005 [NK05]	Node link diagram: fixed-width links	Distribution change	Modes of representation
Eberlein and Peterson 1992 [EP92]	Node link diagram: fixed-width links	Time series plots	None (commercial software)

Table 1: Summary of approaches to the representation of causal models.

mental visualization tools intended to help with complex interactive design decisions [TSDS95, TSWB94]. They linked a visualization with a Monte Carlo simulation to show effects of design decisions. The user was able to adjust design parameters using sliders and see instantly how the different simulations performed. The results were expressed with distribution plots showing the likelihood of different outcomes. Other work by Neufeld and Kristtorn [NK05] showed changes in distributions according an underlying causal statistical model based on theoretical, not Monte-Carlo, distributions.

Several studies have examined ways of expressing causal linkages and effects so they are perceptually immediate. The idea is that people will literally see causal effects as opposed to having to reason about them. These studies, inspired by the work of Michotte and Thines [MT63], leverage perceptual effects where the motion of one object is seen to cause some change in another. They showed that viewers strongly perceive causality when viewing a shape that begins to move after being contacted by another moving shape, as when a billiard ball strikes another. Their series of experiments on the temporal contingencies of the percept found that if there was more than a 200 msec delay between the contact and the second object moving, the perception of causality was lost. This has been incorporated into systems that use animated node-link diagrams to convey causal chains [WNB99, NSK06]. One common causal metaphor is an animated ball, which is emitted from one node and travels to another target node, 'causing' it to vibrate, increase in size, or change color. This is used to indicate the causal effect of one variable, represented by the first node, on another variable, represented by the second node. Kadaba et al. [KIL07] compared static and animated causal visualizations. In their static design, positive influences were indicated with a plus sign (+) glyph and negative influences were indicated a minus sign (−) glyph attached to the link between two entities in the network. In the animated version, the size of an animated round glyph which they called a "bullet" represented the magnitude of the influence on the recipient node. As a bullet hit a recipient node, it grew or shrank to show a positive or negative effect. They found that subjects could interpret animated and static representations equally accurately, but formed responses slightly quicker with animations. Ware has shown that multi-touch screens can be used to convey more complex causal effects such as causal enhancements, effect reductions, and causal blocking effects [War13].

Although the Michotte-inspired animated methods can convey causal effects immediately and can be interpreted with little or no training, there are problems using the method in more complex causal networks [War13]. Consider two nodes with a single causal link, where clicking on one node causes a ball to be emitted. The second node expands after being struck by this ball to signal a positive causal effect, but the second node has to be reset to its original size to see this effect again. A complex causal network can be represented by a directed graph, with both positive and negative causal effects, but multiple causal pathways of different lengths may terminate on a particular node, so arrival times will not be synchronized. Resets could occur at confusing intervals. Additionally, the Michotte-based causal methods are ill-suited to showing causal effects resulting in a changing trend, as in a business model, because they only show effects by transient changes in nodes.

Vensim, a commercial tool for visualizing and analyzing simulation results, supports an interactive activity the authors called "causal tracing" [EP92]. It allows users to view multiple time series plots corresponding to causal influences on a particular effect and thereby infer the causal chain. However, the time series plots are not integrated with the causal influence graph and they do not change in real time when model parameters are adjusted.

2.1. Key design ideas

Our design goal was to make it possible to see the effects of a management decision and to understand the causal chains that resulted in those effects. We chose to abandon the ideal of using Michottian causality perception; instead we have perceptually instantaneous feedback in the form of time series plots showing forecasts that are recomputed in real time in response to adjustments of model parameters using sliders. Time series plots are connected by arc diagram links designed to graphically indicate the amount of *change* relative to the status quo. We call these links 'dynamic change arcs.' Before giving the design details, it is useful to understand the application domain which motivated this research.

2.2. The fisheries model

Our interactive visualization was driven by a food web model developed for fisheries management which simulates predator-prey interactions and competition between several commercially important species. The model is a system of Lotka-Volterra equations [Lot26, Vol26]. Lotka-Volterra equations are a pair of differential equations where a change in one species is a function of the biomass of another. For example, the rate of change in a fox population might depend on the number of rabbits and the rate of change of the rabbit population might depend inversely on the number of foxes. The equations are of the form $\delta A/\delta t = A(p - q \cdot B)$, meaning the change in species A over time is given by the mass of species B multiplied by a scaling value q, which represents how much one species eats or competes with another.

The fisheries model underlying our visualization is MS-PROD, which is a multispecies production model developed by NOAA scientists Gamble and Link [GL09]. This model is a system of equations where the change in biomass of one species is a function of two kinds of interaction with other species in the model. One kind of interaction is a predator prey-relationship (given by α in the model); the other interaction is competition between species (given by β). Competition accounts for cases when two species may compete with one another for a resource such as habitat space or food sources. Another term in the model is harvesting, by humans means such as fishing. The goal of fisheries management is to obtain a situation where harvesting is sustainable over a long period. The MS-PROD model is the following formula:

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

where N is species biomass, β_{ij} represents competition between species i and species j, and α_{ij}, represents predation of species i on species j. H_i is the harvest effort on species i. The K terms represent carrying capacity, the maximum population of the species in the absence of competition, predation, or harvest.

10 species of fish interact with each other in the version of MS-PROD we used. The model authors provided us with a sample parameter file that listed these 10 species chosen from the Northeast United States Continental Shelf Large Marine Ecosystem (NEUSLME), listed here by functional group, which is a set of species that share habitat, ecosystem function, and other characteristics:

- **Elasmobranchs**: skates, spiny dogfish
- **Flatfish**: windowpane, winter flounder, yellowtail flounder
- **Groundfish**: cod, haddock, redfish
- **Small pelagics**: herring, mackerel

It should be noted that this parameter file has only been tuned roughly to the NEUSLME to provide a set of interesting interactions to test this project's ability to visualize the cause and effect linkages between species. Thus, it should not be considered to describe the real world system—only whether the Dynamic Change Arcs effectively display outputs from time series that are linked by causal relationships.

The MS-PROD model runs simulations for 30 years with an annual time step to predict individual biomasses.

3. Detailed design

There are three critical elements in our design: *time series plots* to show forecast trends, *arcs* to show model linkages, and *interactive sliders* to allow the user to change parameter values. These are illustrated together in Figure 1. A critical part of our design is a graphical device known as an "arc diagram". This name was coined by Wattenberg, though the method was invented earlier; Knuth used arc diagrams to illustrate interaction of characters in a novel [Wat02, Knu93]. While the arc diagram may fail to properly depict the structure of a network, Heer et al. point out its one-dimensionality allows for other features to be easily displayed near the nodes [HBO10]. This property is useful in our design since it fits well with a stack of time series plots. The time series plots stand for nodes in the food web network, with each plot representing a single species. The time series plots are organized into four functional groups, each distinguished using a different color. The four interactive sliders control the fishing effort for each functional group. By adjusting a slider, the amount of fishing effort for a particular functional group can be changed and the results visualized.

3.1. Parameter adjustment sliders

The sliders on the left-hand side of the interface panel in Figure 1 enable users to adjust the amount of fish caught in different functional groups. In reality, fishing for a particular group of fish would result in by-catch of other species, but the MS-PROD model does not account for by-catch. Underneath each slider, a colored rectangle indicates differences from the baseline effort settings. Blue indicates the effort value was increased since the baseline was set, while red indicates the effort value was decreased. In the example shown, the effort for the elasmobranch group was doubled. The time series plots show the resulting forecast changes in the biomass of the entire set of species according to the model. When a slider is adjusted, the model is recomputed and the time series are updated in real time.

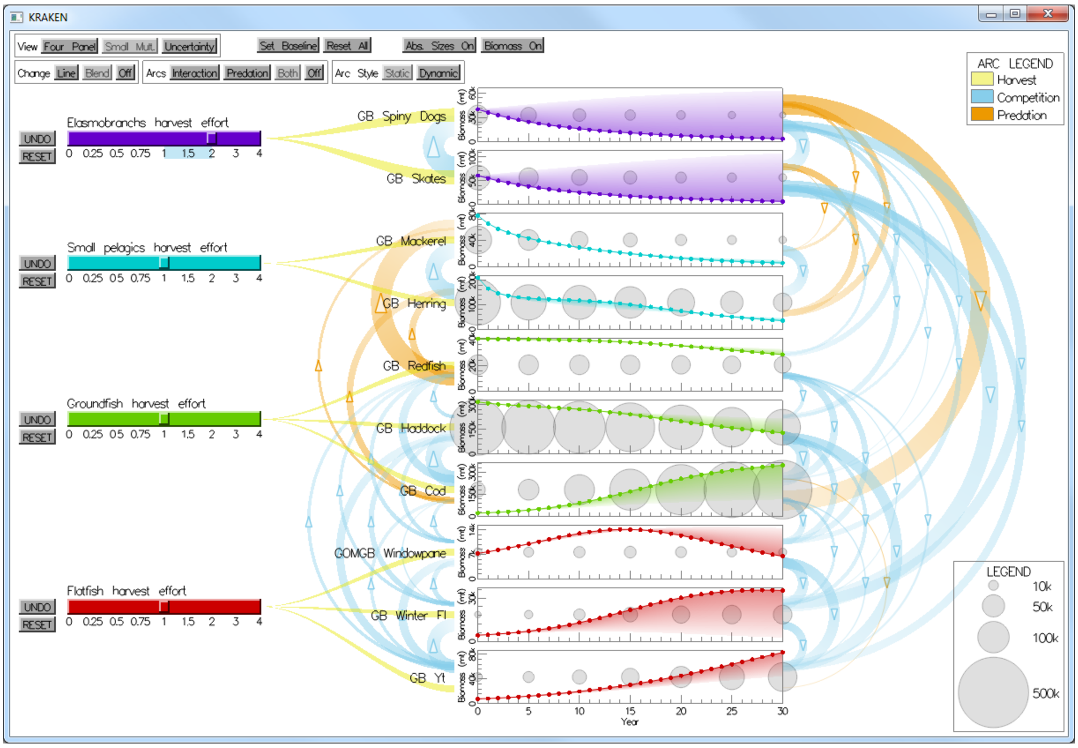


Figure 1: The overall design is illustrated.

3.2. Time series plots

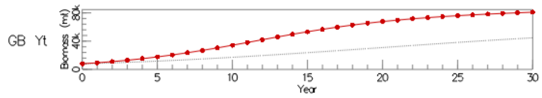
The time series plots show 30 year projections based on the model. Since the biomasses of the species varies considerably, each plot has a different y-axis scale. Absolute biomass indicators (shown in the background of the time series plots in Figure 1) were introduced to make comparison across species possible and to prevent the assumption that the plots have similar scales. These indicators show the absolute biomass of the population as the area of a circle and are drawn at five-year intervals across the 30 year time span.

3.3. Visualization of change

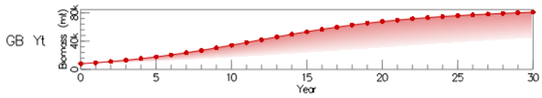
To understand and compare decisions, modelers and other stakeholders need to be able to see *changes* in biomass resulting from *changes* in the fishing effort. We designed two alternatives to represent time series change from a baseline. The first, seen in Figure 2a, shows the baseline forecast as a dotted gray line and the current forecast as a heavier colored line with dots at yearly intervals. The second, shown in Figure 2b, is a shaded area originating from the curve of the current forecast. The shading diminishes in opacity as it approaches the curve of the baseline forecast. Stakeholders in early evaluations preferred the shaded version, so it became part of the standard presentation.

3.4. Interspecies relationships

The MS-PROD model is a system of Lotka-Volterra equations, so *explaining*, as opposed to simply showing, the results of a particular parameter change is a matter of making these parameter values explicit. This is the role of the arcs in Figure 1. These relationships can be stated in terms of cause and effect. For example, increasing the catch of elasmobranchs (spiny dogfish and skates) causes a decrease in the population of these species. A secondary effect is that winter flounder and yellowtail flounder populations increase since they face less competition from the elasmobranchs.



(a) The status quo time series is given by the gray line. The colored, dotted line indicates the new forecast.



(b) The status quo time series is given by the bottom of the shaded portion.

Figure 2: Alternative representations of 'before' and 'after' time series.

Two colors were used for the arcs to differentiate the types of model interaction terms: orange for predation and blue for competition. Both predation and competition relationships are directed and three cues indicate this in the non-animated versions. First, we used fading opacity based on Holten and van Wijk's recommendation of dark-to-light shading along each arc [HvW09]. Second, triangular marks are drawn in the middle of the arcs to point from the source species to the recipient species. Third, our arcs follow a clockwise direction; arcs on the right-hand side are all directed downward, while arcs on the left-hand side go upward. This serves both to reinforce directionality and to provide a principled way of decluttering the diagram. Without this left-right separation, it would be difficult to show reciprocal relationships between species. The directionality can also be indicated with animation, where the color of the arc alternates with gray and the stripes of colors travel from the source species (e.g., predator) to the recipient species (e.g., prey).

We created three alternative versions of the arcs to show the underlying model parameters: static arcs, dynamic arcs, and dynamic arcs with animation.

Static. With the static style of the arcs, all arcs are shown always, as shown in Figure 1. The width of an arc corresponds to the magnitude of the relationship as defined by the predation or competition term in the model. The downside is that viewing all arcs at once can be overwhelming as the display becomes somewhat cluttered.

In addition, reasoning about cause and effect relationships using the static arcs can be difficult. The effect of a change in the amount of fishing of species A on species B that it eats (or competes with) is weighted by the amount of change in A, as well as the biomass of the two species. Thus, although the static arc view provides the information needed to understand causal chains, the reasoning process is complex.

Dynamic. Dynamic arcs, illustrated in Figure 3, were motivated by the need to simplify reasoning about causal chains occurring in the model. Another benefit is they reduce the visual clutter created by the static arcs. The width of dynamic arcs depends on the size of the causal influence. If there is no change in effect due to a parameter change, the arcs do not appear. Dynamic arcs are a simplification of the effects of an iterative process; in a forecast, the model is recomputed at yearly intervals and the biomass values of a species and its influence on other species has 30 different values in a 30 year forecast. The arcs can represent only a single value.

In our final design, the width of each directed arc is proportional to a weight w_{ij}: the effect of the jth species on the ith species. In the case of predation,

$$w_{ij} = \alpha_{ij} \cdot 100000 \cdot (N'_{j,30} - N_{j,30})/(N_{i,0} + 100000) \quad (2)$$

where α_{ij} represents the predation of species j on species i, $N_{i,0}$ represents the initial biomass of the prey species, $N_{j,30}$ is the biomass at year 30 for the predator species according to the current forecast, and $N'_{j,30}$ is the biomass at year 30 for the predator species according to the baseline forecast. If the predator species biomass at year 30 did not change between the forecasts, then $(N'_{j,30} - N_{j,30})$ equals zero, resulting in a w of zero, so the arc will not be drawn. In other words, the link width is given by the size of the predation coefficient, weighted by the overall change in biomass of the predator species over 30 years and inversely weighted by the biomass of the prey species at the start of the forecast.

In the MS-PROD model, the rate of change in a species is proportional to its own biomass. In this regard, a somewhat subtle point must be made; in a *visual sense* the effects on a prey species are inversely weighed by the biomass of the prey, simply because of the different scales used for the different time series plots. In Figure 3, for example, the change in biomass of haddock appears small relative to the change in biomass of mackerel, but the actual absolute change in haddock is more than twice as large since there are far more haddock. Because of the different scales, what we perceive in the time series plots is a *relative change*, not an absolute change. Back in Equation 1, N_i is a multiplier for all terms on the right-hand side, but we can eliminate it by dividing both sides by N_i to give the relative change $dN_i/N_i dt$. The bottom term $N_{i,0}$ in Equation 2 is there solely to prevent arcs getting too wide in the case of the largest prey species. It has a small effect for the other species.

Width values calculated using Equation 2 can be either positive or negative. Both competition and predation relationships inhibit the growth of the recipient species since the source species either consumes the recipient species itself or its resources. If there is an increase in a predator species' biomass, the effect on a prey species will be negative and their population will decline. The reverse is true: if a predator species decreases, then the prey species will benefit and its population will increase. 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 signs ($-$) 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 [KIL09]. Plus signs were drawn in black; minus signs were drawn in white outlined in black. Several sign glyphs are drawn along each arc to allow the user to determine the arc's signage.

The dynamic links between the harvest sliders and the harvested species behave similarly. Their width is a function of the change in harvest from the baseline.

Figure 3 shows the dynamic arc representation for the same scenario given in Figure 1. There has been an increase in fishing effort on the functional group of elasmobranch from 1.0 to 2.0. Spiny dogfish and skates decrease as a result of the increase in fishing effort, as is indicated by the shaded area between the baseline forecast and the current forecast. There are also secondary effects: spiny dogfish strongly predate cod, so the cod population increases, as indicated by the ($+$) signs on the orange arc from spiny dogfish to cod. Elasmobranchs compete with winter flounder and yellowtail flounder, causing a population increase for the flounders. However, windowpane population declined as a result of the increase in cod, since windowpane and cod compete. While these results may not be a true representation of the NEUSLME, the dynamic arcs allow the modeler to immediately see if a cause and effect relationship is modeled as desired with the set of parameters provided.

Dynamic with animation. The final style for the arcs is dynamic with animation. All of the rules concerning non-animated dynamic arcs apply. Additionally, the plus signs ($+$) or minus signs ($-$) travel from the source species (e.g., the predator) to the recipient species (e.g. the prey). Also, the color of the arc—orange or blue—alternates with the color gray. These alternating stripes of colors also travel along the arc to give a stronger indication of the directionality.

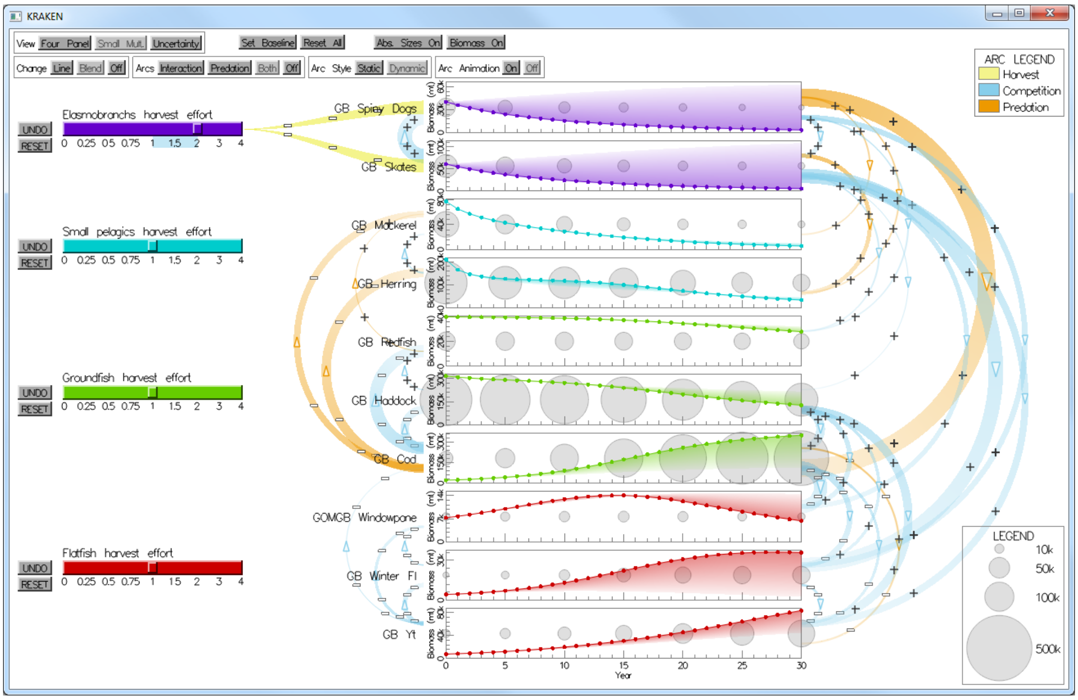


Figure 3: The design illustrated with dynamic arcs to show causal relationships between the species. Arc thickness reflects the magnitude of the causal effect.

4. Evaluation

We were interested in how different arc visualization alternatives enhance a user's understanding of the complex relationships between the fish species and the effects of those relationships as a result of changes in fish catch. Our hypothesis was that there would be benefits to showing causal arcs and these benefits would be greater for the dynamic change arcs. To test this, we conducted a user study to measure the performance of different arc depiction alternatives.

4.1. Method

In our study, participants manipulated a dynamic slider to change the fish harvest for a functional group, then reported on the resulting changes in biomass by answering a series of standardized questions. First, they were asked to explain how the harvest change affected the population forecast of a specific species. Next, they were asked to explain the *causes* of that change, with questions that ranged from straightforward to complex. An example of a simple question would be to explain a change in a species that had declined because the fishing effort had increased for that species. A complex question would require the participant to explain a causal chain of predation and/or competition. For example, under the given model parameters, spiny dogfish (a member of the elasmobranch group) eats cod, so increasing the catch of elasmobranchs results in a decrease of spiny dogfish, which in turn leads to an increase of cod. We hypothesized that only the complex questions would benefit from the presence of arcs.

4.2. Conditions

There were four experimental conditions.

(A) **No arcs**: Only the time series are displayed.
(B) **Static arcs**: Arcs are drawn between the time series to show predation or competition. Each arc's width is based on the model coefficient defining the relationship between a pair of species. The arcs are always shown.
(C) **Dynamic arcs without animation**: The arcs change in width according to causal linkage.
(D) **Dynamic arcs with animation**: The arcs change in width dynamically and also are animated to help indicate the direction of the relationship.

4.3. Procedure

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

A between-participants design was used: each participant conducted the experiment task for only one of the four conditions. The experiment began with a brief training session which was tailored according to the experimental condition—i.e., arcs were explained only for conditions B, C, and D; the meaning of dynamic arcs was explained only for conditions C and D. Feedback about the quality of the participant's answers was given only during training.

After training, the participant followed on screen instructions to manipulate one of the sliders controlling the fish catch. The participant then answered questions about the resulting effects and the reasons for the effects. The experiment lasted approximately fifteen minutes.

4.4. 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 entered their answers using the laptop keyboard and mouse.

4.5. Participants

There were 92 participants who took part in the study, all of whom were recruited by a poster affixed to the back of the privacy screen. The responses of three participants were eliminated because of errors in even the most basic questions, indicating a lack of understanding of the task. Participants were randomly assigned to the four conditions, such that there were at least 20 in each condition, and were rewarded with a set of pens.

	No arcs	Static arcs	Dyn.	Dyn. & animated
Low quant.	11	10	9	10
High quant.	12	14	11	12

Table 2: The numbers of students recruited for each of the four conditions, separated by quantitative level.

Participants reported their college at the university (e.g., engineering, business, liberal arts), which we used to group them in terms of quantitative skills. Students who reported being from the College of Liberal Arts were placed in the “low quantitative” category and students from any other college were placed in “high quantitative.” Our thought was that students in fields such business, science, or engineering were more likely to have experience reading charts. We tried to obtain equal numbers in each category but were not entirely successful. The numbers we recruited are shown in Table 2.

4.5.1. 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)?” For example, “What was the effect on haddock?” Participants answered this “What?” 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 were all reset to one and a new instruction was given for the next set of questions until all questions were answered.

As mentioned earlier, the “Why?” questions varied in difficulty. They were designed so that the questions would fit in one of two difficulty categories:

- **First-order**: These questions were simpler because they involved a fish species whose biomass changed directly as a result of increased or decreased fishing effort.
- **Higher-order**: These questions were harder because they involved a fish species whose biomass changed as an indirect result of fishing effort change. The explanation required following a second-order or higher causal effect.

There were three instructions for adjusting the harvest effort and seven pairs of “What?” and “Why?” questions. All participants were given the same instructions and asked the same questions in the same order, regardless of condition.

For example, the participant may be instructed, “Double the harvest effort on elasmobranchs.” The participant would then be asked, “(a) What was the effect on cod? (b) Why?” Answering correctly requires looking at a second order effect: “(a) Cod increased a lot. (b) Spiny dogfish is a type of elasmobranch, so its biomass went down because it was being fished more. Spiny dogfish prey on cod, so the cod biomass increased since there were less predators.” There were even more difficult questions such as, “(a) What is the effect on haddock? (b) Why?” The correct explanation would look something like, “(a) Haddock decreased a lot. (b) Spiny dogfish, which are elasmobranchs, prey on cod, so the cod biomass increases as more spiny dogfish are fished. Cod competes with haddock, so as the cod biomass increases, the effect of the competition is stronger and the haddock biomass declines.” Both of these examples fall into the higher-order difficulty category.

4.6. Results

The answers to the questions of the evaluation were graded on a scale of zero (i.e., completely wrong) to three (i.e., completely correct). “What?” questions, which were answered using a drop-down, were graded automatically, while “Why?” questions were graded by two graders. The average of the two scores was taken and these averages were used in our analyses. The correlation coefficient (Pearson’s r) between the scores assigned by the two graders was 0.7.

The results are summarized in Figure 4. Separate ANOVAs were run with Tukey HSD tests for each of these three types of questions (what, first-order, and higher-order). There were no significant differences between the four conditions for the seven “What?” questions and the three first-order “Why?” questions. However, there was a significant effect for the four conditions with the higher-order “Why?” questions, shown by ($F[3,81] = 13.2$; $p < 0.001$). A Tukey

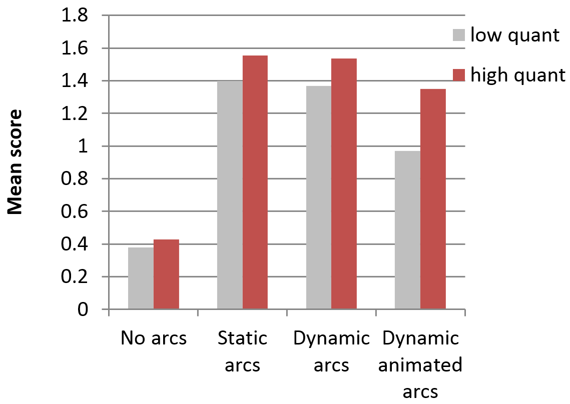


Figure 4: Mean grades for the higher-order explanations given by study participants under the four conditions. Means of both high and low quantitative participants are shown.

HSD test on the four conditions showed that all of the arc conditions were better than the no arcs conditions. However, there were no significant differences between the three conditions with arcs.

The effect for high versus low quantitative ($F[1,81] = 1.82$; $p = 0.18$) failed to reach significance.

4.7. Discussion

Our results suggest that using either static or dynamic arcs is better than using no arcs at all for asking higher-order "Why?" questions about the underlying relationships between the species, while "What?" questions could be answered even without any causal relationship depiction. This was true for both high and low quantitative participants. Dynamic arcs performed slightly better without animation than with animation; perhaps participants found the animation distracting or confusing.

5. Expert feedback

Our visualization was assessed by two model developers and one fisherman who has been a member of the New England Fishery Management Council. The response of these three expert users was positive. The modelers liked the ability to change fishing effort parameters and see an instantaneous result with the altered forecasts. They appreciated that using the visualization was quicker and more informative than the alternative of rerunning the model and graphing its output in Microsoft Excel. All three users preferred dynamic arcs over static arcs for depicting the causal relationships between the fish species. Two of the experts admitted that the underlying relationships and their counter-intuitive effects can confuse even them without the aid of a visualization, despite being quite familiar with the 10 species involved. Dynamic arcs made these complex interactions much easier for them to interpret; one of them noted that dynamic arcs "help to follow the flow [of the relationships] more easily." The three experts agreed that the visualization could play an important role in promoting the use of more complex models like MS-PROD.

6. Conclusion

The broad goal of this research has been to investigate ways of visually representing causal chains in a complex model to allow users to reason about why various effects occur when changes are made to critical model parameters.

The key components of our solution are as follows:

- Sliders to dynamically change model input parameters.
- Real time recalculation of model forecasts.
- Nodes containing time series showing the forecasts and making clear *differences* from some baseline forecast.
- Links showing *changes in causal effects* of one model component on another as a result of a change in model parameters.

Our evaluation of the depictions of the interspecies relationships in a fisheries model showed that having weighted causal links is superior to no links for answering higher-order questions about forecast changes. Domain experts preferred the dynamic change arcs over the static arcs, but the more formal evaluation with undergraduate students failed to show an improvement for this mode of representation.

The prior work can be grouped into two categories: one where transitory changes occur (the works shown above the double line in Table 1), the other where they remain visible to be studied (the works shown below the double line in Table 1). Our visualization differs from both categories in that it shows both the before and after views by depicting deviation from the status quo, also while selectively displaying causal arcs that help explain that deviation.

It is pertinent to ask, how general purpose is our solution? Can it be used in other fields such as business, economics, medicine, or engineering? There are two major issues relating to this. First, will it *scale* to more complex models? Our design combining stacked time series plots with a dynamic arc diagram works well for 10 interacting components of a system, and this number could perhaps be doubled and still be clear, but larger models would require a different approach. An alternative for larger causal networks might be to use a spring layout node-link diagram where each node contains a much smaller time series plot. So long as interactions between components only involved small subgraphs of the network, the dynamic change arc approach would be effective in decluttering the diagram. Interactive methods where topologically nearby nodes and links are enhanced (e.g., one and two links away from a selected node) have been demonstrated to make node-link diagrams with several hundred nodes usable for reasoning [WB05]. An interactive hierarchical network view using a method such as the intelligent zoom technique could also help with larger networks [BHDH95, SZG*96]. Also, fisheye methods might be used to expand the nodes and their embedded time series plots for sub-components of a model relating to a selected node [CCF97, GKN05]. Second, how are the arcs *visually weighted to express the model*? In the case of MS-PROD, our dynamic change arcs were a simplification of a series of interactions occurring over 30 discrete time steps. This is a relatively complex case and many models are simpler. To use dynamic change arcs, only a function that describes a node to node causal effect (either positive or negative) is needed. We believe many models may fit into this class, such as the disease factor models discussed by Kadaba et al. [KIL07].

Another limitation of our design is that dynamic change arcs are based on the endpoint of a 30 year time series and the magnitudes of the changes over this period. This can work well where the forecasts have simple monotonic trends, but if there are oscillations up and down this simplification will necessarily fail to capture the complexity. Nevertheless, a time series approach can show far more detail than the simple Michotte-derived animations used to show immediate causality [MT63, WNB99].

A more concrete goal of this research was to help the fishery management community make informed decisions by using the MS-PROD model through our interface. Achieving this long-term goal would require a public unveiling of the MS-PROD model by its authors, which unfortunately is unlikely to occur soon. On the other hand, our visualization is already being used by the model authors, including in several meetings and seminars they have held. Our informal interviews with expert users has led us to believe we have succeeded in creating an effective visualization and that the dynamic arc version is the most informative. More widespread release will depend on validation of the model itself. In the future, evaluation of individual components of the visualization and exploration of alternative techniques could improve the product further.

Funding

This work was supported by NOAA Grant #NA05NOS4001153 and NSF Grant #1016607.

Acknowledgements

We wish to thank David Goethel and Michael Fogarty for their feedback. Joel Wyatt and Rachel Quimby helped run the studies.

References

[BHDH95] BARTRAM L., HO A., DILL J., HENIGMAN F.: The continuous zoom: A constrained fisheye technique for viewing and navigating large information spaces. In *Proceedings of the 8th Annual ACM Symposium on User Interface and Software Technology* (New York, NY, USA, 1995), UIST '95, ACM, pp. 207–215. URL: http://doi.acm.org/10.1145/215585.215977, doi:10.1145/215585.215977. 8

[CCF97] CARPENDALE M., COWPERTHWAITE D., FRACCHIA F.: Extending distortion viewing from 2d to 3d. *Computer Graphics and Applications, IEEE 17*, 4 (Jul 1997), 42–51. doi:10.1109/38.595268. 8

[EP92] EBERLEIN R. L., PETERSON D. W.: Understanding models with Vensim™. *European Journal of Operational Research 59*, 1 (May 1992), 216–219. 2

[GKN05] GANSNER E. R., KOREN Y., NORTH S. C.: Topological fisheye views for visualizing large graphs. *IEEE Trans. Vis. Comput. Graph. 11*, 4 (2005), 457–468. 8

[GL09] GAMBLE R. J., LINK J. S.: Analyzing the tradeoffs among ecological and fishing effects on an example fish community: A multispecies (fisheries) production model. *Ecological Modelling 220*, 19 (2009), 2570–2582. 3

[HBO10] HEER J., BOSTOCK M., OGIEVETSKY V.: A tour through the visualization zoo. *Commun. ACM 53*, 6 (June 2010), 59–67. 3

[HvW09] HOLTEN D., VAN WIJK J. J.: A user study on visualizing directed edges in graphs. In *Proceedings of the SIGCHI Conference on Human Factors in Computing Systems* (New York, NY, USA, 2009), CHI '09, ACM, pp. 2299–2308. 4

[IMCZ98] IGARASHI T., MACKINLAY J., CHANG B.-W., ZELLWEGER P.: Fluid visualization of spreadsheet structures. In *Visual Languages, 1998. Proceedings. 1998 IEEE Symposium on* (Sep 1998), pp. 118–125. doi:10.1109/VL.1998.706154. 1, 2

[KIL07] KADABA N., IRANI P., LEBOE J.: Visualizing causal semantics using animations. *IEEE Transactions on Visualization and Computer Graphics 13*, 6 (2007), 1254–1261. doi:10.1109/TVCG.2007.70618. 2, 8

[KIL09] KADABA N. R., IRANI P. P., LEBOE J.: Analyzing animated representations of complex causal semantics. In *Proceedings of the 6th Symposium on Applied Perception in Graphics and Visualization* (New York, NY, USA, 2009), APGV '09, ACM, pp. 77–84. URL: http://doi.acm.org/10.1145/1620993.1621009, doi:10.1145/1620993.1621009. 5

[Knu93] KNUTH D. E.: *The Stanford GraphBase: a platform for combinatorial computing*. ACM, New York, NY, USA, 1993. 3

[Lot26] LOTKA A. J.: The frequency distribution of scientific productivity. *Journal of the Washington Academy of Sciences 16*, 12 (1926), 317–323. 3

[MT63] MICHOTTE A., THINÉS G.: La causalité perceptive. *Journal de Psychologie Normale et Pathologique* (1963), 9–36. 2, 9

[NK05] NEUFELD E., KRISTTORN S.: Picturing causality - the serendipitous semiotics of causal graphs. In *Smart Graphics* (2005), Butz A., Fisher B. D., Krüger A., Olivier P., (Eds.), vol. 3638 of *Lecture Notes in Computer Science*, Springer, pp. 252–262. 2

[NSK06] NEUFELD E., SOLHEIM J., KRISTTORN S.: Experiments in the perception of causality. In *Smart Graphics* (2006), Butz A., Fisher B., Krüger A., Olivier P., (Eds.), vol. 4073 of *Lecture Notes in Computer Science*, Springer, pp. 36–49. 2

[Ram80] RAMSDELL R. E.: The power of visicalc. *BYTE* (Nov. 1980), 190–192. 1

[SZG*96] SCHAFFER D., ZUO Z., GREENBERG S., BARTRAM L., DILL J., DUBS S., ROSEMAN M.: Navigating hierarchically clustered networks through fisheye and full-zoom methods. *ACM Trans. Comput.-Hum. Interact. 3*, 2 (June 1996), 162–188. URL: http://doi.acm.org/10.1145/230562.230577, doi:10.1145/230562.230577. 8

[TSDS95] TWEEDIE L., SPENCE R., DAWKES H., SU H.: The influence explorer. In *CHI 95 Conference Companion* (1995), Miller J., Katz I. R., Mack R. L., Marks L., (Eds.), ACM, pp. 129–130. 2

[TSWB94] TWEEDIE L., SPENCE B., WILLIAMS D., BHOGAL R.: The attribute explorer. In *Conference Companion on Human Factors in Computing Systems* (New York, NY, USA, 1994), CHI '94, ACM, pp. 435–436. URL: http://doi.acm.org/10.1145/259963.260433, doi:10.1145/259963.260433. 2

[Vol26] VOLTERRA V.: Fluctuations in the abundance of a species considered mathematically. *Nature 118* (1926), 558–560. 3

[War13] WARE C.: Perceiving complex causation through interaction. In *Proceedings of the Symposium on Computational Aesthetics* (New York, NY, USA, 2013), CAE '13, ACM, pp. 29–35. URL: http://doi.acm.org/10.1145/2487276.2487279, doi:10.1145/2487276.2487279. 2

[Wat02] WATTENBERG M.: Arc diagrams: Visualizing structure in strings. In *Proceedings of the IEEE Symposium on Information Visualization* (Washington, DC, USA, 2002), INFOVIS '02, IEEE Computer Society, pp. 110–116. 3

[WB05] WARE C., BOBROW R.: Supporting visual queries on medium-sized node-link diagrams. *Information Visualization 4* (2005), 49–58. 8

[WNB99] WARE C., NEUFELD E., BARTRAM L.: Visualizing causal relations. *IEEE Symposium on Information Visualization: Late Breaking Hot Topics* (1999), 39–42. 2, 9

[ZRNG99] ZAPATA-RIVERA J.-D., NEUFELD E., GREER J. E.: Visualization of bayesian belief networks. In *IEEE Visualization 1999 Late Breaking Hot Topics Proceedings* (1999), Press, pp. 85–88. 2