Modeling as an Aid to Robust Tactical Decision-Making

Gary L. Klein, Mark Pfaff, Jill L. Drury, Jennifer Mathieu, John James, and Paula Mahoney

16 June 2009

Core Idea

http://www.surfcity-hb.org/images/users/fire/cedar_fire3a.jpg

Let emergency responders visualize more futures and save more lives through Robust Decision Making (RDM)

Photos from Huntingdon Beach (CA) Fire-rescue; right: http://www.surfcity-hb.org/images/users/fire/fire_rescue.jpg

Why “robust”? Uncertainty!

Robust

“Optimal” solutions can be so sensitive to small changes in the environment that they can often be “brittle” choices

Robust solutions may not be the best choice under all conditions but are successful under the broadest swath of plausible futures

Robust Decision-Making (RDM) Analysis

- **Simulation model generates plausible futures for each course of action (COA) and calculates range of costs → "option awareness"**
- **Decisions involve choosing the most "robust" COA based on comparing the cost distributions**
- **A COA with a low, tight cost range indicates it is relatively successful even when worst case conditions occur**

RDM bridges the “Situation Space” and “Decision Space” gap

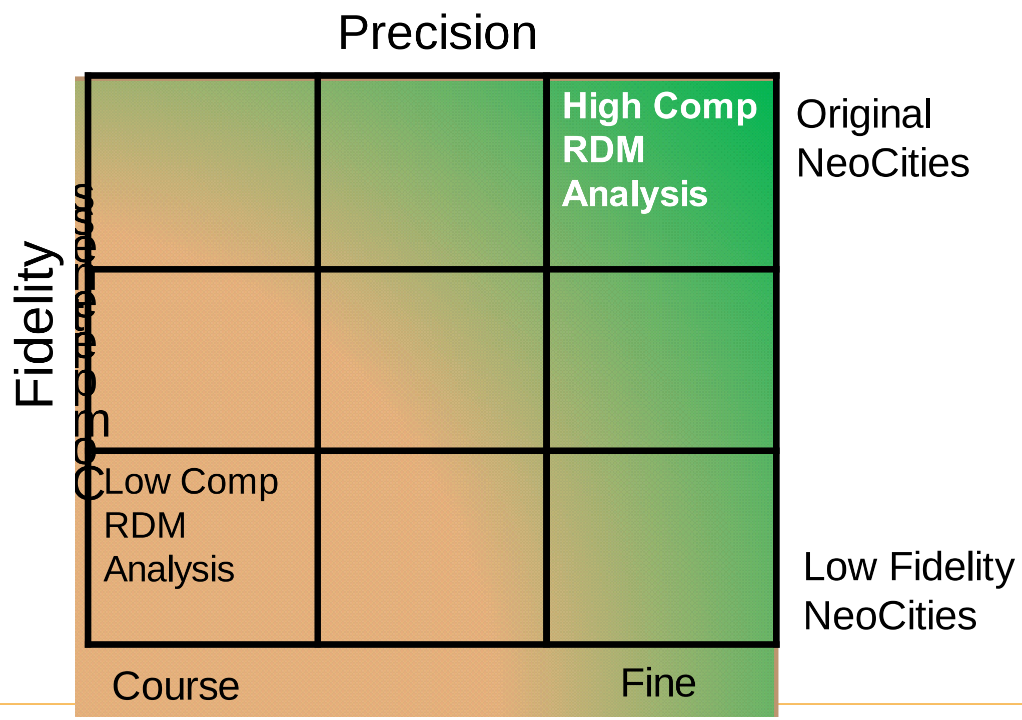
Photo: Jill Drury

- Situation space consists of the facts
  - …such as raw sensor data, map-based information, or alerts
  - Provides situation awareness
- Decision space results from analysis of options
  - Models provide bridge to projecting plausible futures
  - Provides “option awareness”

Research approach

*To make the modeling faster: more **tactical***

- **Develop guidelines for the tradeoff space between fidelity and precision of models to generate RDM displays**
- **Develop RDM visualizations that enable emergency responders to understand the robustness of any given course of action (COA)**
- **Perform human-in-the-loop testing for each visualization to assess subjects' decision-making performance**



An illustration of the trade space using Penn State U's NeoCITIES model

Explore model tradeoff space:
The NeoCITIES model

- Scaled world simulation
- Teams of emergency responders allocate resources to events
- Algorithmic/time-stepped model
- Developed by Penn State U.

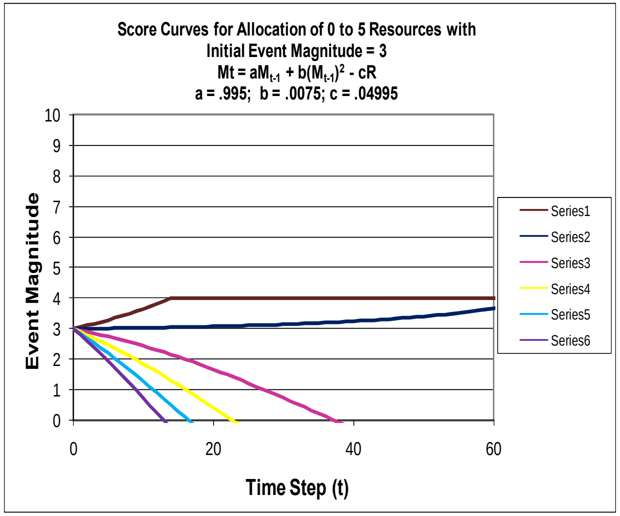
Photos of NeoCITIES lab courtesy of Penn State

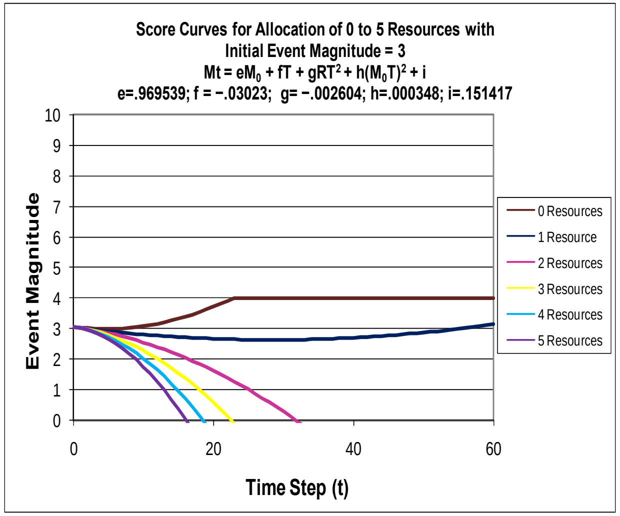
Explore model tradeoff space:
Metric for comparing decisions

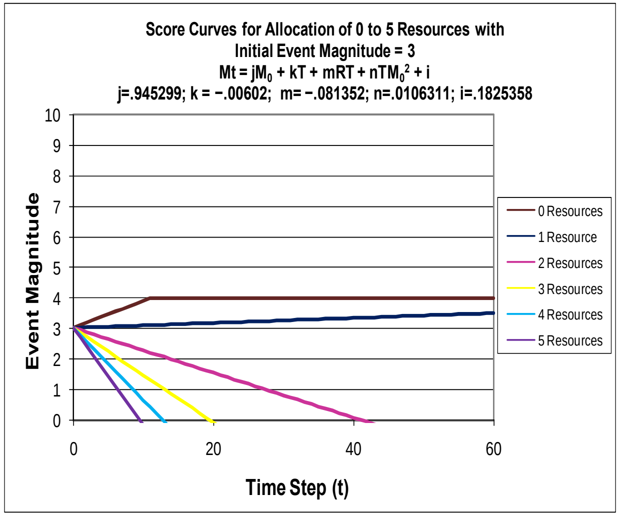
- The Decision Space construct enables evaluating models in terms of their impact on the RDM analysis
  - Order of options
  - Distance between options
  - Changes in situations underlying option performance
- Key to this analysis is developing a "cost" measure for that includes immediate and future consequences
- For NeoCITIES, cost components are:
  - Sending resources for current emergency
  - Injuries/deaths sustained in current emergency
  - Property damage incurred in current emergency
  - Injuries/deaths that resulting from negatively impacting the next emergency (e.g., if insufficient resources remain)
  - Property damage occurring in next emergency

Explore model tradeoff space
Cost model fidelity

- "Event magnitude" calculation is central to calculating the measure of costs of each course of action
- We ran models varying fidelity by manipulating the underlying equations that calculate how the magnitude of a simulated event grows and evaluated impact on COAs







Gold Standard
Incremental non-linear

"Medium" fidelity
Non-linear regression
R^2 = .87

"Low" fidelity
Linear regression
R^2 = .96

Explore model tradeoff space

Sensitivity analysis precision

	Sample Size Factors			
Precision	# Time Steps Sampled	Resources Arrived (R)	Initial Incident Magnitude (M)	# Monte Carlo Runs
High	60 {1, … 60}	6 {0, 1, … 5}	16 {1.25, 1.5 …4.75}	10K
Medium	12 {5, 10… 60}	6 {0, 1, … 5}	8 {1.25, 1.75…4.75}	1K
Low	6 {10, 20… 60}	6 {0, 1, … 5}	4 {1.25, 2, 3, 4}	100

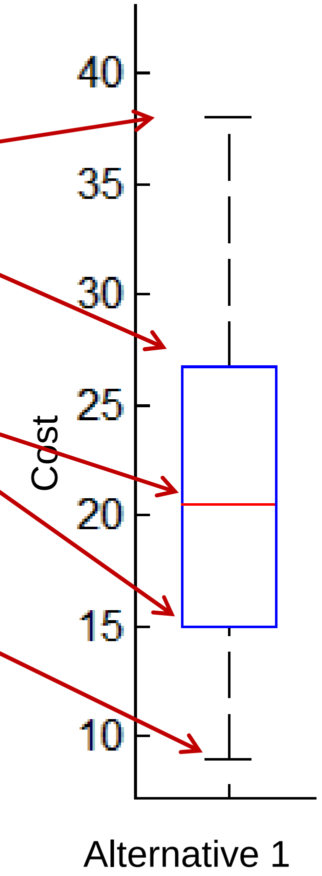
Where a chosen COA is {0, 1, 2, …,5} assigned resources (e.g., fire trucks)

For a given COA, for each simulation run, at each time point:

- **R = a random sample from a Poisson distribution around the chosen COA**
- **M = a random sample from a Normal distribution around the reported Magnitude**

About RDM visualizations
Initial visualization

- We are developing multiple ways to show ranges of costs for COAs, but…
- …starting with Tukey's box plots:
  - The highest cost of all possible futures
  - The cost of 25% of all futures fall between here and the median
  - The median cost (half cost more & half cost less)
  - The cost of 25% of all futures fall between here and the median
  - The lowest cost of a future under this alternative
  - The costs of 50% of all futures fall within the box
- The median cost and cost ranges depend on the likelihood of each possible future and how it will interact with the chosen alternative



Explore model tradeoff space

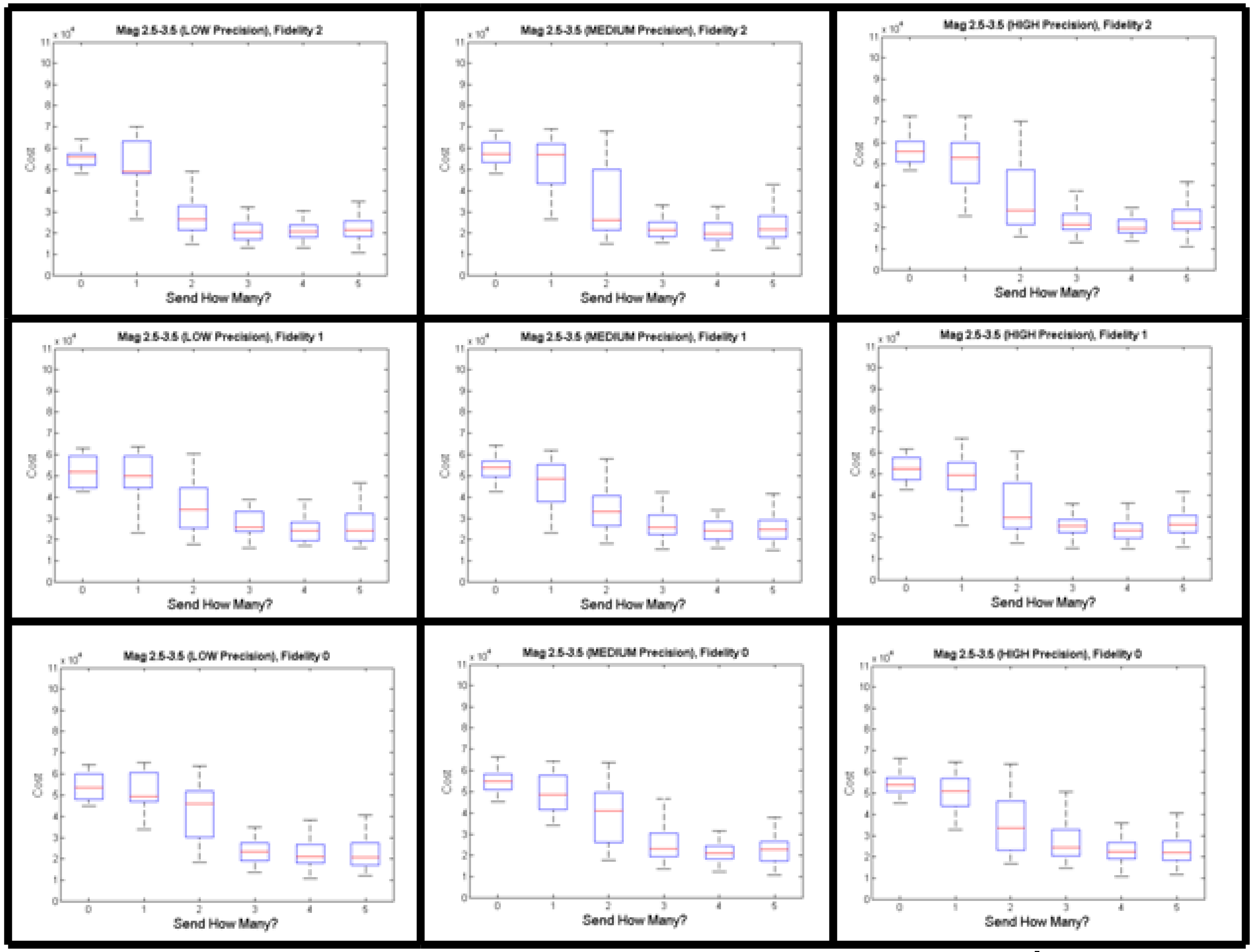
Results: Impact on order, distance, & robustness?

Precision

Fidelity

Original NeoCities

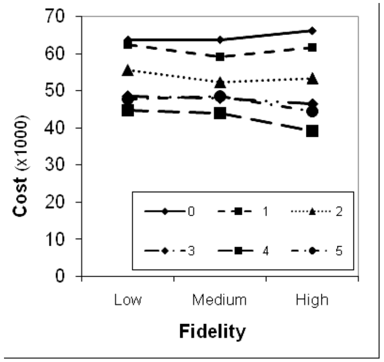
Low Fidelity NeoCities

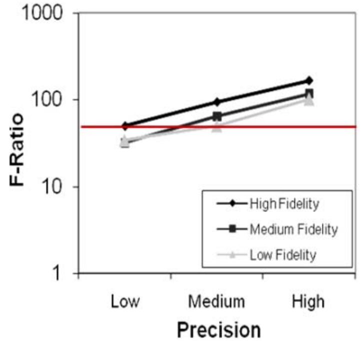


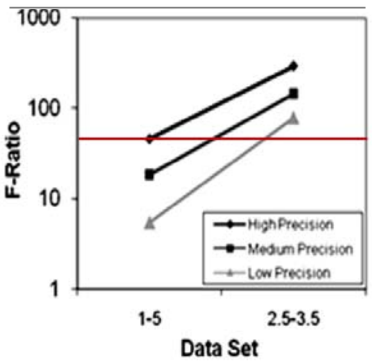
Course

Fine

Explore model tradeoff space:
Results of NeoCITIES analysis





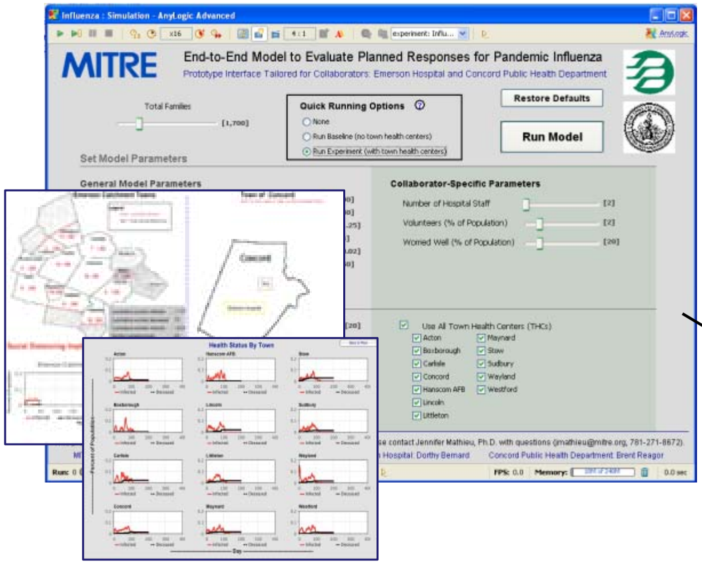


- **Changing fidelity does not change order**
- **Interaction with precision**
  - **Precision of the sensitivity analysis has a highly significant impact on the differentiation (F-ratio) among options**
  - **Fidelity of the cost model also has a statistically significant but smaller impact on differentiation**
- **Interaction with data ambiguity**
  - **High precision does better overall**
  - **All means are statistically different**
  - **The difference is greatest when the data is more ambiguous**
  - **All levels of precision do better with less ambiguous data**

Explore model tradeoff space:
Disease spread model

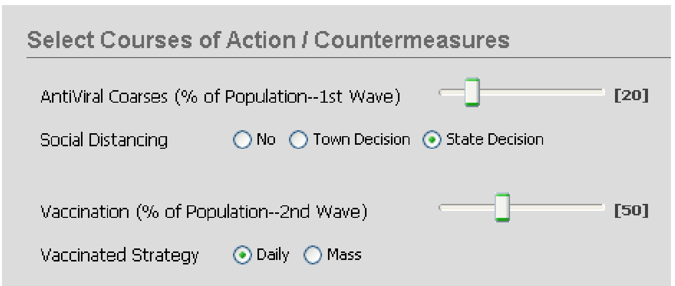
Used 2nd model to explore generalizability of results

- **Hospital catchment scenario**
- **Discrete event, agent-based model**
- **Calculates infectivity (i.e., disease spread) and the course of action-related time delays (i.e., process model)**
- **Developed by MITRE**



Emerson Hospital

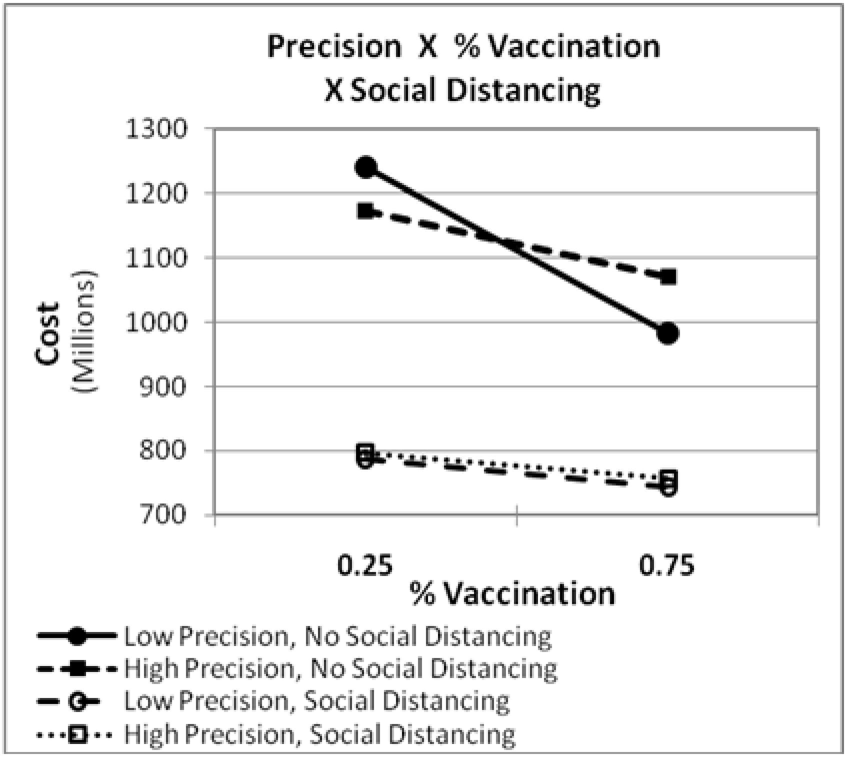
Explore model tradeoff space:
Disease spread model investigation



- **Ran at 2 levels of precision**
- **Investigated 4 courses of action**
- **2 x 2 x 2 x 2 x 2 full-factorial ANOVA:**
  - **Precision: low, high**
  - **Social distancing: true, false**
  - **Daily vaccination strategy: true, false**
  - **Level of vaccination: 25%, 75%**
  - **Level of antivirals: 10%, 50%**

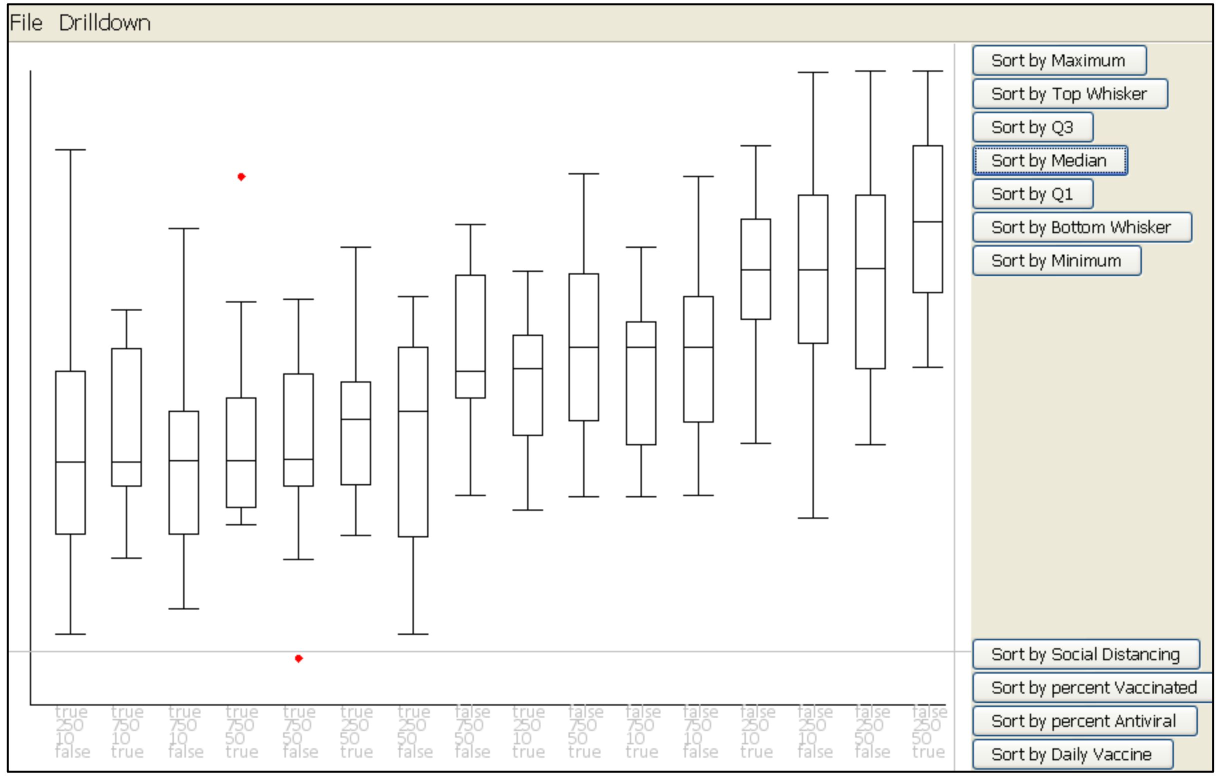
Explore model tradeoff space:
Disease spread model results

- Three-way interaction between precision, level of vaccination, and social distancing, $F(1,688) = 4.15$, $p < .05$, $\eta_p^2 = .006$
  - Interaction between precision and level of vaccination only occurs when social distancing does not occur
  - Otherwise, there is no apparent difference in the cost predictions of either the high- or low-precision levels
- Both models lead to same COA conclusions



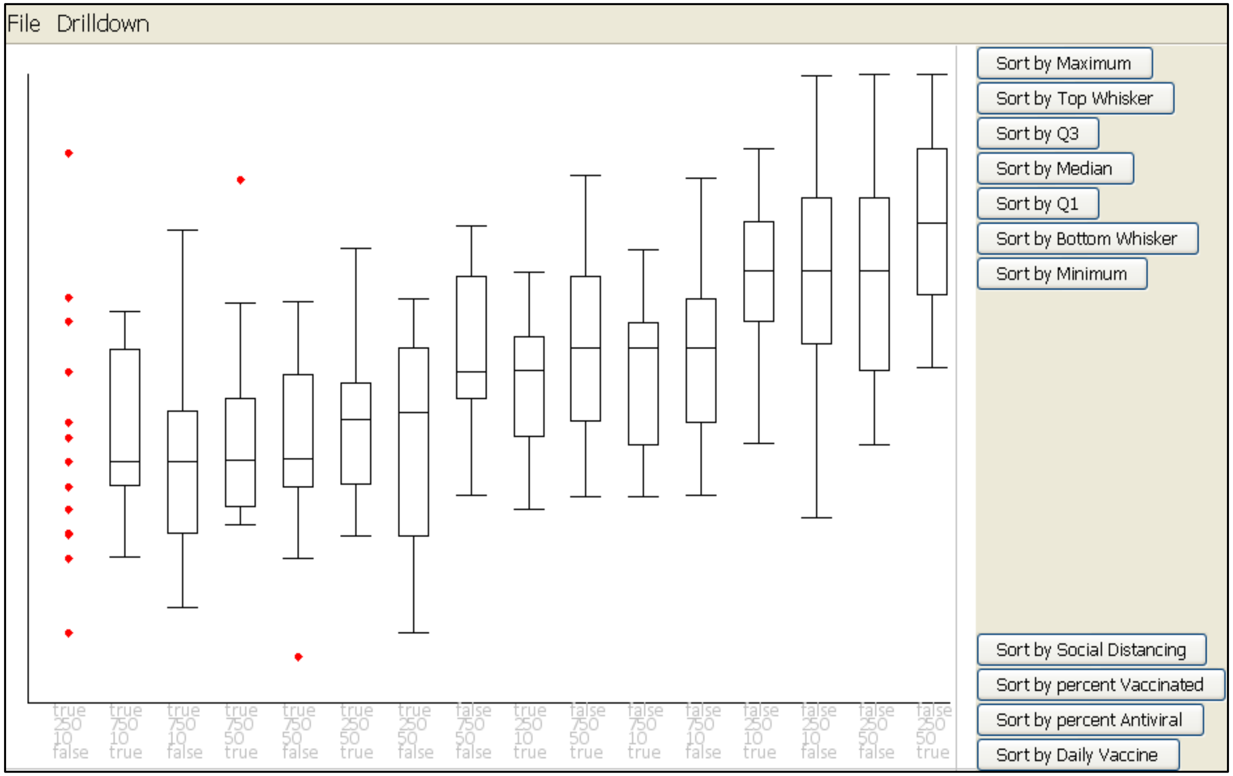
Develop RDM Visualizations
A disease spread model results

- **Social distancing clearly leads to lower cost outcomes**
- **Conflict apparent between**
  lowest median cost
  and
  lowest maximum cost



Develop RDM Visualizations
A disease spread model results

- **Social distancing clearly leads to lower cost outcomes**
- **Conflict apparent between lowest median cost and lowest maximum cost**
- **Drill down enables exploration of underlying conditions**
- **New COAs can result**



Perform human-in-the-loop testing:
The need for principled test scenario generation

Will the wind come up or die down? Will it start raining soon?

- To determine whether our decision aids result in high-quality decisions:
  - Need to test them under a representative sample of challenging decisions
  - Decisions must span the major types of cognitive challenges for a particular domain

A Huntington Beach (CA) Fire responder surveys the Cedar Fire in 2003. See http://www.surfcity-hb.org/images/users/fire/cedar_fire3a.jpg

Perform human-in-the-loop testing:
Three step process

A Huntington Beach (CA) Fire truck responds to the Santiago wildfire in 2007. See http://www.surfcity-hb.org/images/users/fire/2007_Santiago_Fire.JPG

1. **Define the decision space**
2. **Determine the cost components, e.g.:**
   - **Initial incident magnitude, number of resources used, property damage, injuries/deaths, "extra" future costs due to over-commitment of resources to current incident**
3. **Choose conflicting pairs of cost components, e.g.:**
   - **A small fire, implying low *property damage*, in a densely inhabited area, which implies high *personal injury***

Perform human-in-the-loop testing:
Test environment

Perform human-in-the-loop testing:
Validation of scenario generation process

- **Experiment involved non-ambiguous (control) scenarios and three types of ambiguous scenarios**
- **Participants made decisions significantly faster in non-ambiguous scenarios**
- **All three types of ambiguous scenarios were significantly different from one another**

Validated that we could develop scenarios in a structured fashion, controlling for types of ambiguity

Additional results from human-in-the-loop testing

Participants interact with the NeoCITIES testbed at Penn State (photo courtesy of PSU)

- The decision space information did positively impact decisions made using the box-plot decision aid
- Group with the decision space had a higher confidence in decisions than the group with only the situation space
- The decision space group felt they had greater decision support than the situation space group
  - Participants did not appear to have difficulty in understanding or making use of the plots

Conclusions

- Systematically examined the impact of reducing the fidelity and precision of mathematical models, and the precision of agent-based models
  - Led to decreases in computation time for lower fidelity/precision models
  - Resulted in statistically significant changes in the decision space
    - Changes were limited to the distance among the options, but not the ordering of the options
    - Normatively a decision maker should make the same choices under the less computational intense models as under the high fidelity/precision models….
- Visualizations of RDM result in better, more confident choices

More tactical RDM models are possible