Requirements for Data Mining the Decision Space

Jill L. Drury, Gary L. Klein, Scott Musman

The MITRE Corporation

Yikun Liu, and Mark Pfaff

Indiana University Indianapolis

June 2012

Research goal

Enable decision makers to attain deep levels of option awareness and choose a robust option via data mining visualizations of the decision space

Photo: Army Joint Support Team, usacac.army.mil/cac2/AJST

Problem

Decision makers need, and often have, an abundance of facts about the situation

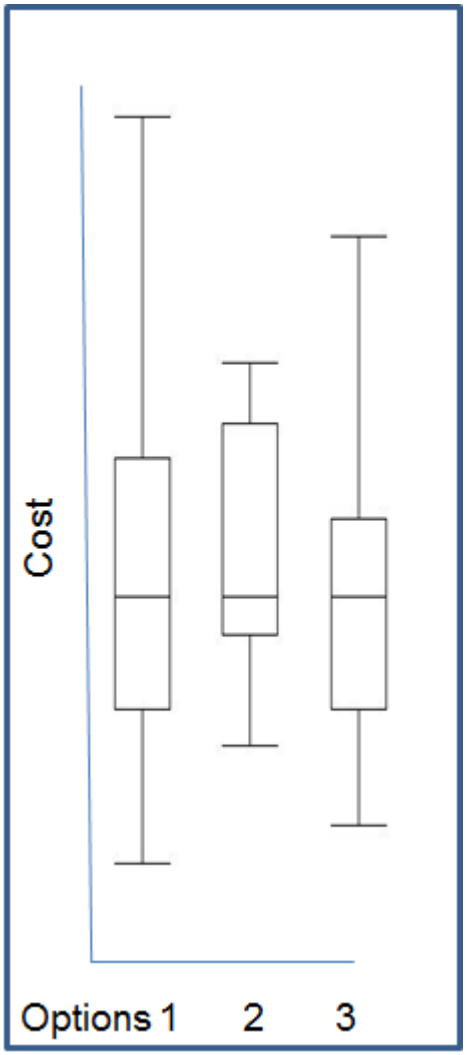
Especially under uncertainty, they need, but don't often have, information that aids comparing the relative desirability of possible options

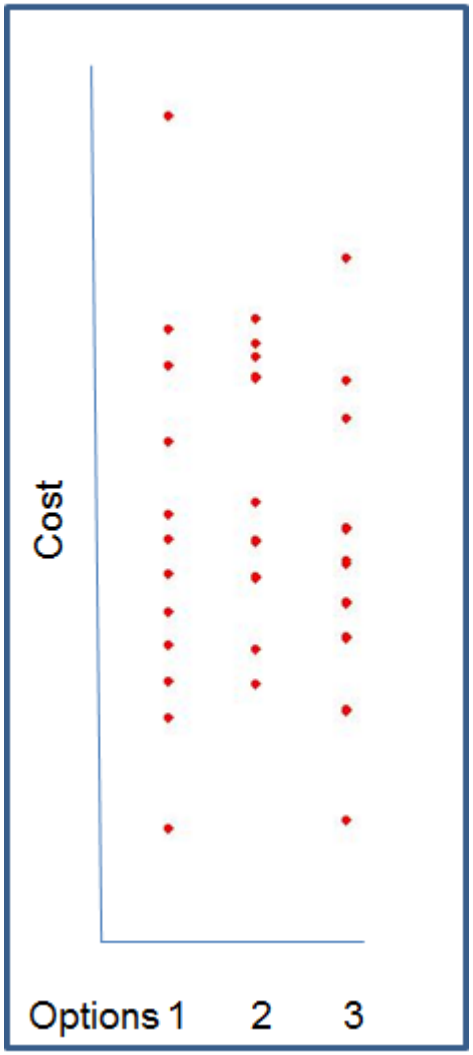
We are bridging this situation space-to-decision space gap

	Situation Space	**Decision Space** (*Hall et al., 2007)
Components	Descriptive information from reports, sensors, etc.	Potential options and possible consequences
Outcome	***Situation Awareness*** (Endsley, 1988)	***Option Awareness*** (Drury et al., 2009 Klein et al., 2010)

How decision spaces can help

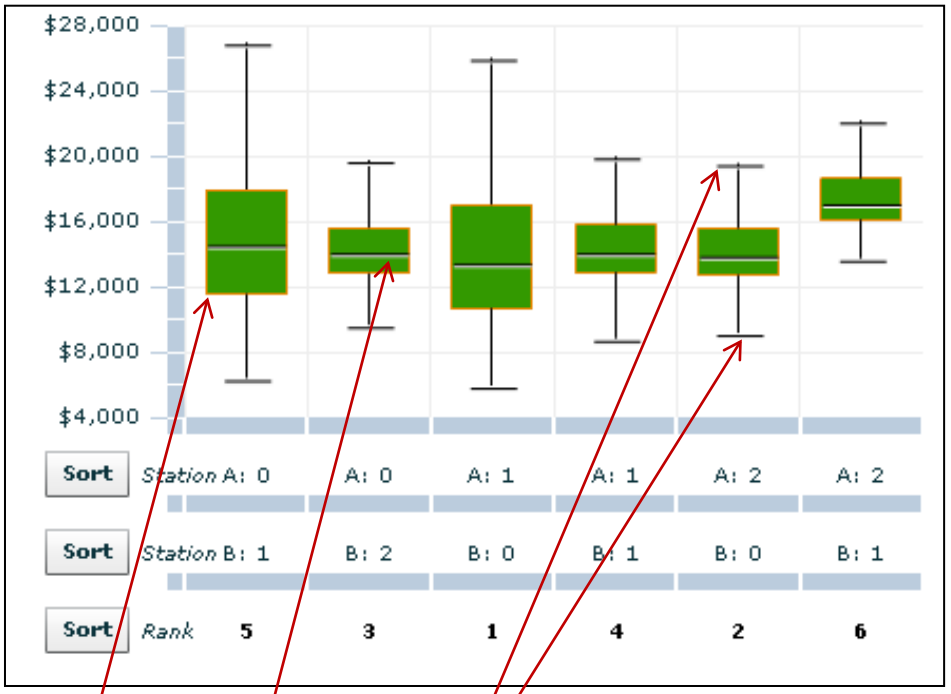
- A decision space visualization provides comparative information about options – ***Option Awareness***
- For each option, multiple conditions are tested that result in a range of plausible costs (or other measures)
- These conditions and their costs can be generated in a number of ways
- Graphics like box plots can visually clarify differences in cost distributions
- The decision space is a visual artifact that facilitates team collaboration





Using decision spaces to attain Level-1 Option Awareness

- **Showing frequencies, not probabilities, reduces decision biases**
- **Scores futures by cost or other metric; assuming an example fire scenario:**
  - **Cost of sending the trucks**
  - **The immediate damage and casualties that occurred**
  - **Damage that might occur elsewhere in the near future because the committed trucks are now unavailable**
- **Seeing the shape of the distributions yields *Level-1 Option Awareness***



Line in box indicates cost of median case

Upper and lower whiskers indicate maximum and minimum costs

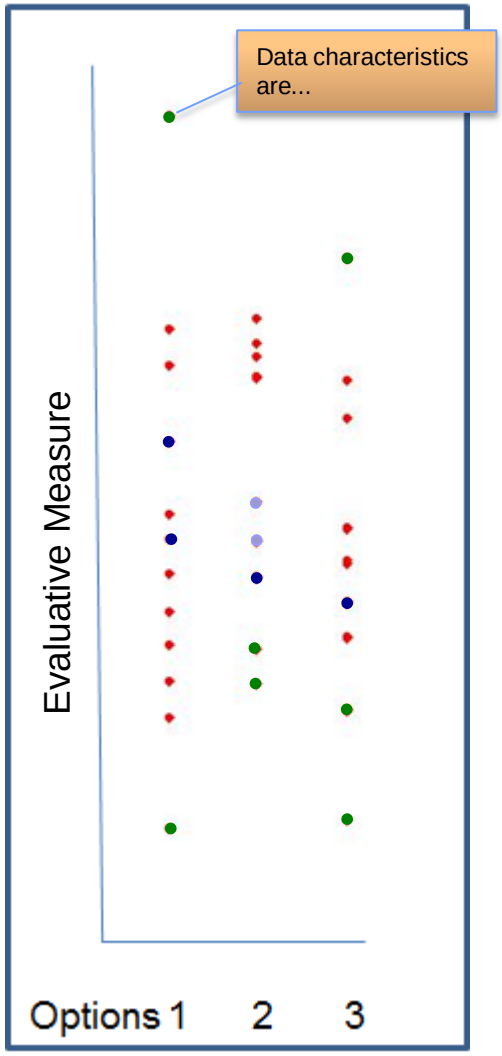
Box defines range of costs for cases in 25th – 75th percentiles

Additional Option Awareness levels

- **But there are two more levels of Option Awareness**
- **Analogous to the levels of Situation Awareness defined by Endsley (1988, 2000)**
- **To dig into the data at Option Awareness Levels 2 and 3, we are using *data mining*:**
  - **The analysis step of the process of extracting patterns and discovering knowledge in data**

Option Awareness Levels

Perception and comprehension of relative robustness of alternative options	Perception and comprehension of relationships between factors underlying the option outcomes	Projection of underlying relationships to adjusted or new options	**Decision**
Level 1	**Level 2**	**Level 3**	



Focus of this paper

- Our prior research showed that having Level-1 Option Awareness enables faster, more correct, and more confident decisions
- Now we're investigating Level-2 and -3 Option Awareness via a decision aid that enables interactive mining of the data
- As a first step, we defined requirements for a novel interactive decision support system
  - Combining interactive data mining and frequency-formatted decision space visualizations for the first time
  - Using prototype mock-ups to explore how requirements could be satisfied

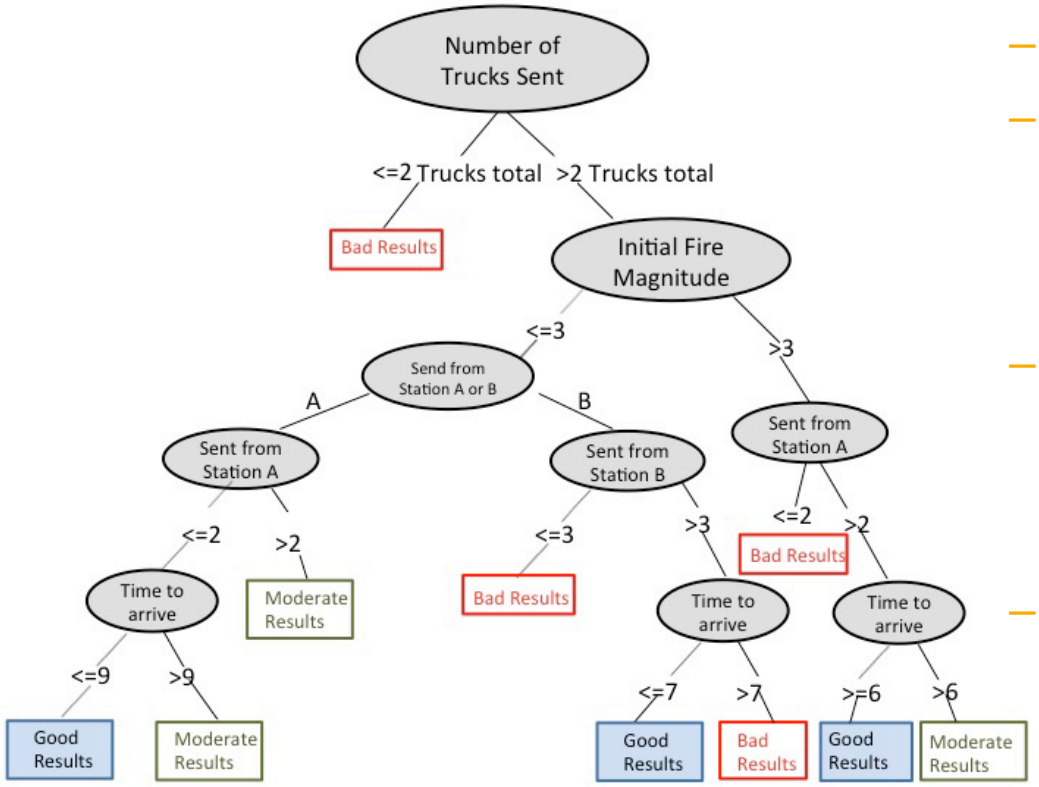
Example domain: Emergency response command and control

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

- **Used NeoCITIES simulation model developed at Penn State U.**
  - **Discrete-event and time-stepped**
  - **Team-based**
- **Includes parameters such as:**
  - **Initial fire magnitude**
  - **Location of fire**
  - **Traffic congestion en route**
- **Decision makers can choose:**
  - **Number of firefighting vehicles**
  - **Where the equipment originates**

Data mining

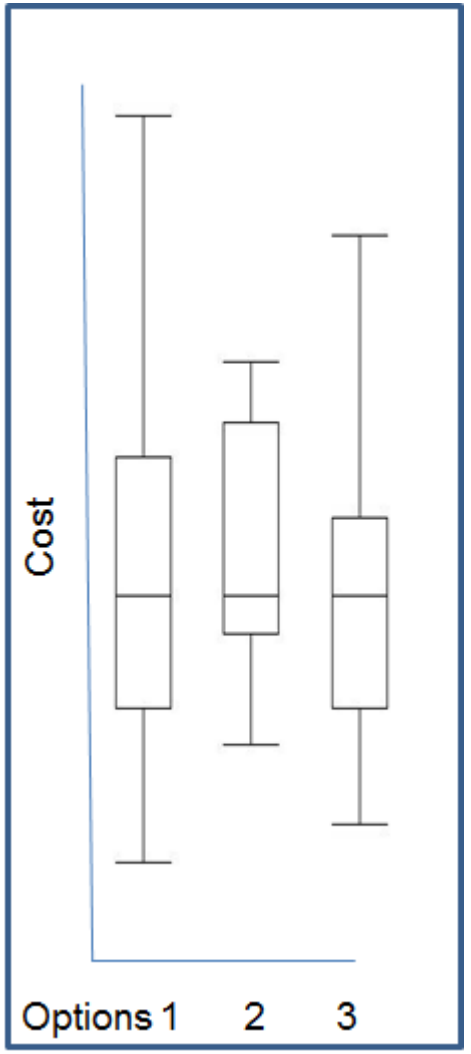
- Used Weka: open-source, general-purpose tool developed at the University of Waikato
- User enters criterion to delimit good versus bad outcomes from the simulation
- J48 algorithm in Weka produces a classifier tree



- Hierarchical structure
- Starts with one node that represents an attribute that provides the greatest discrimination among outcomes
- Each successively lower level tree branches either lead to another discriminatory factor or to a leaf node
- Current format is for research only; will likely show users another data presentation

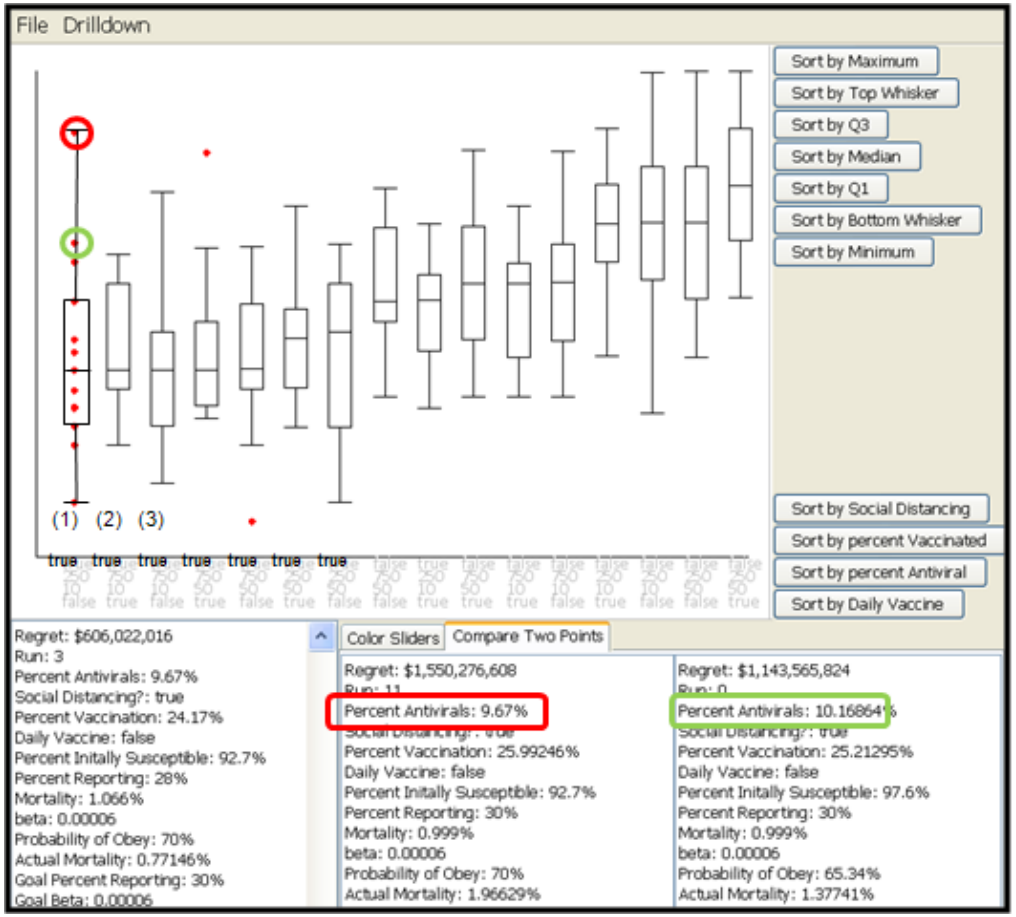
Requirements: Overall goal

Reqmt #	1.0 Overall Goal
Reqmt	The system shall enable the user to see which options have more robust outcomes
Rationale	To aid understanding decision choices, users will want to find options that have a majority of good outcomes
OA level	1



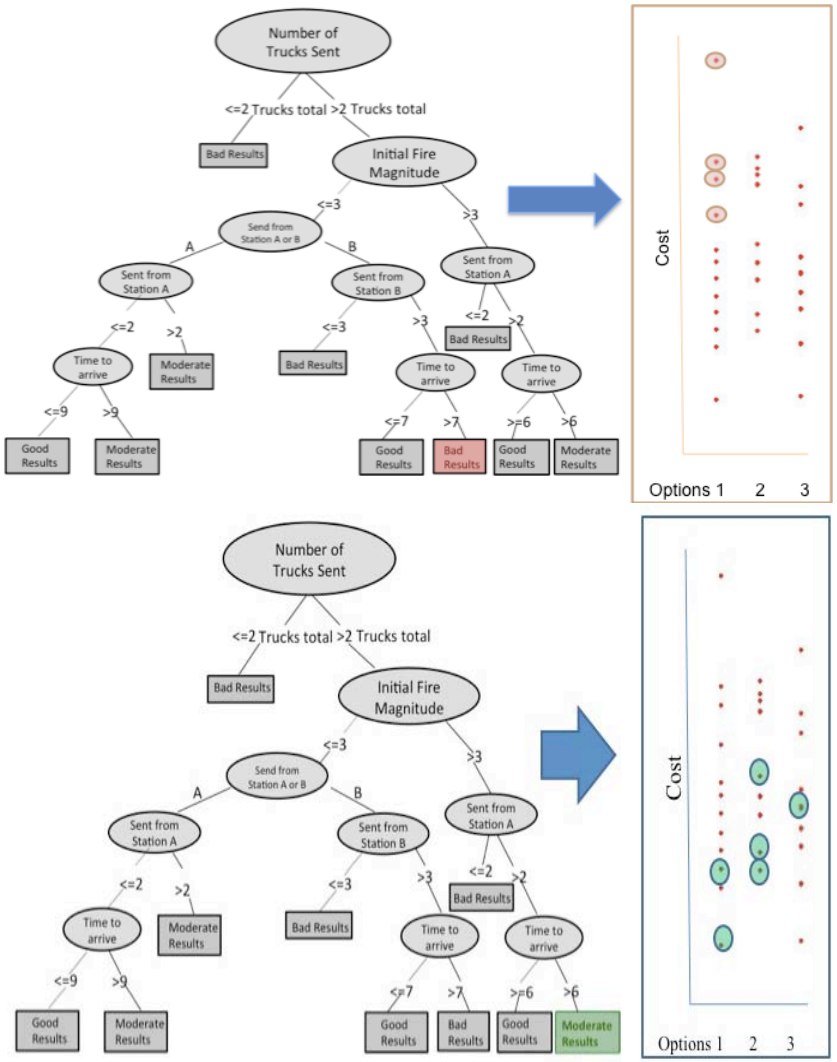
Requirements: Robustness support

Reqmt #	1.1 Robustness Support
Reqmt	The system shall enable the user to see options, when they exist, that have the possibility of being made more robust after being modified to mitigate the conditions leading to bad outcomes and/or facilitate conditions leading to good outcomes.
Ration-ale	Options that have conditions that lend themselves to being shaped are more attractive.
OA levels	2 and 3



Example of earlier work: Pandemic Disease Spread Option Analysis Simulation Model – with Jennifer Mathieu

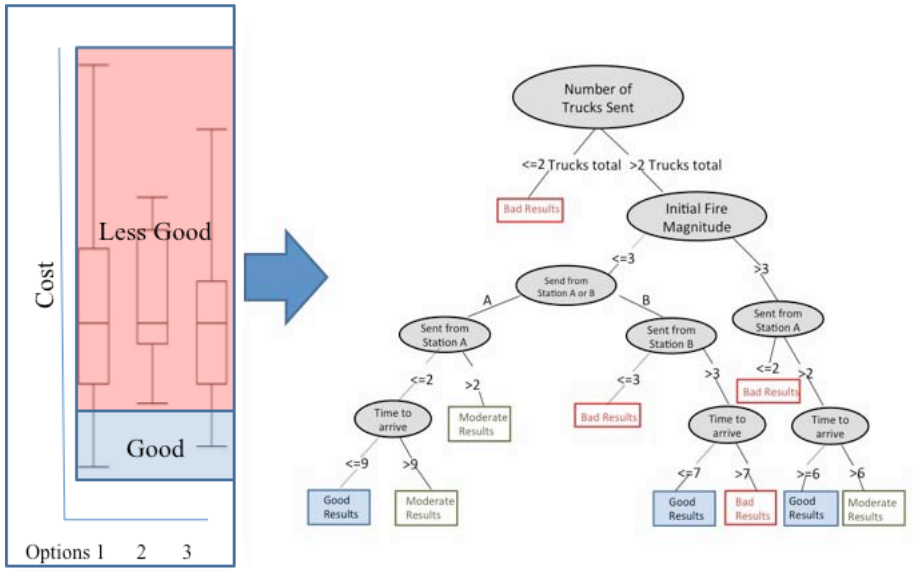
Requirements: Data exploration support



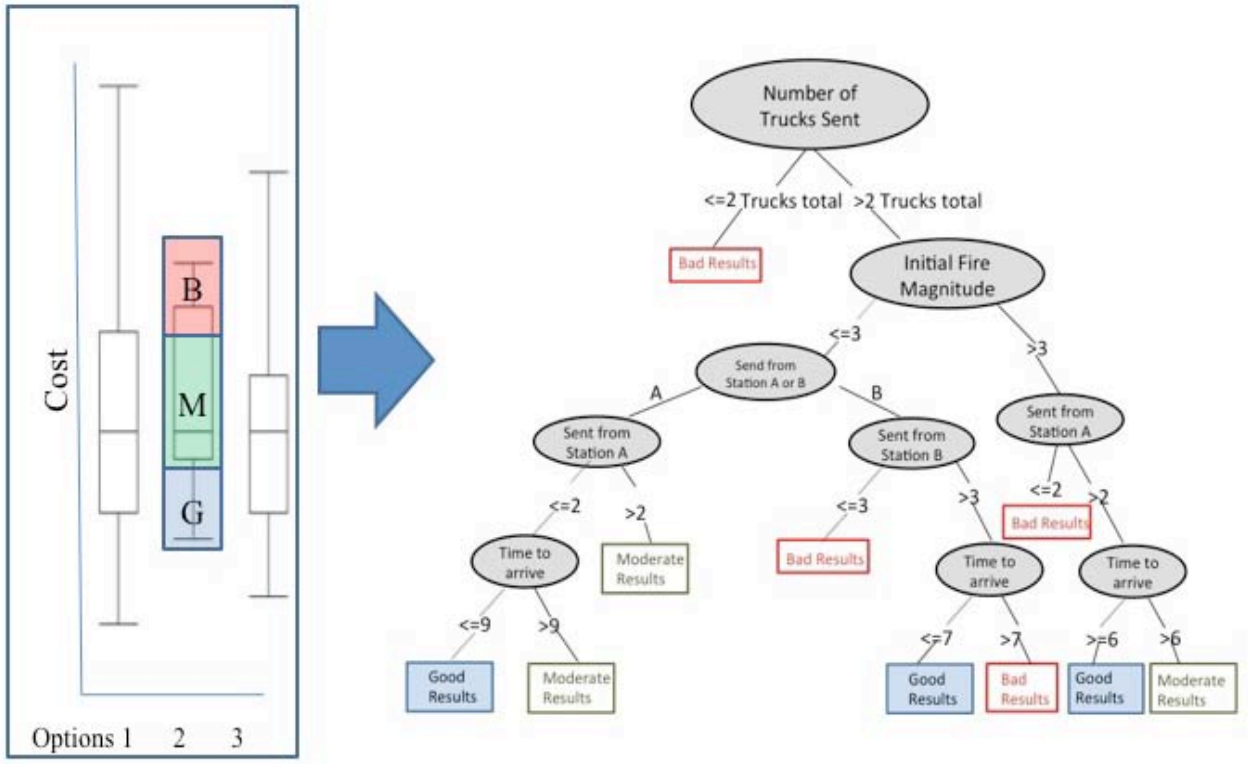
Reqmt #	1.2.1 Data Exploration Support
Reqmt	The system shall support the user in exploring the data in real-time, such that the user can immediately see the relationship between the change in their selection and the results.
Ration-ale	Research shows that performance can be improved by allowing people to make dynamically changing selections and see the results of the changes in real time.
OA levels	Applicable to all three levels.

Requirements: Data exploration support (2)

Reqmt #	1.2.2 Data Exploration Support
Reqmt	The system shall support the user in identifying which conditions yield similar outcomes.
Rationale	Users want to know what conditions will tend to lead to better or worse outcomes so that they can choose or modify an option with the result that more satisfactory outcomes will occur.
OA levels	Applicable to all three levels.

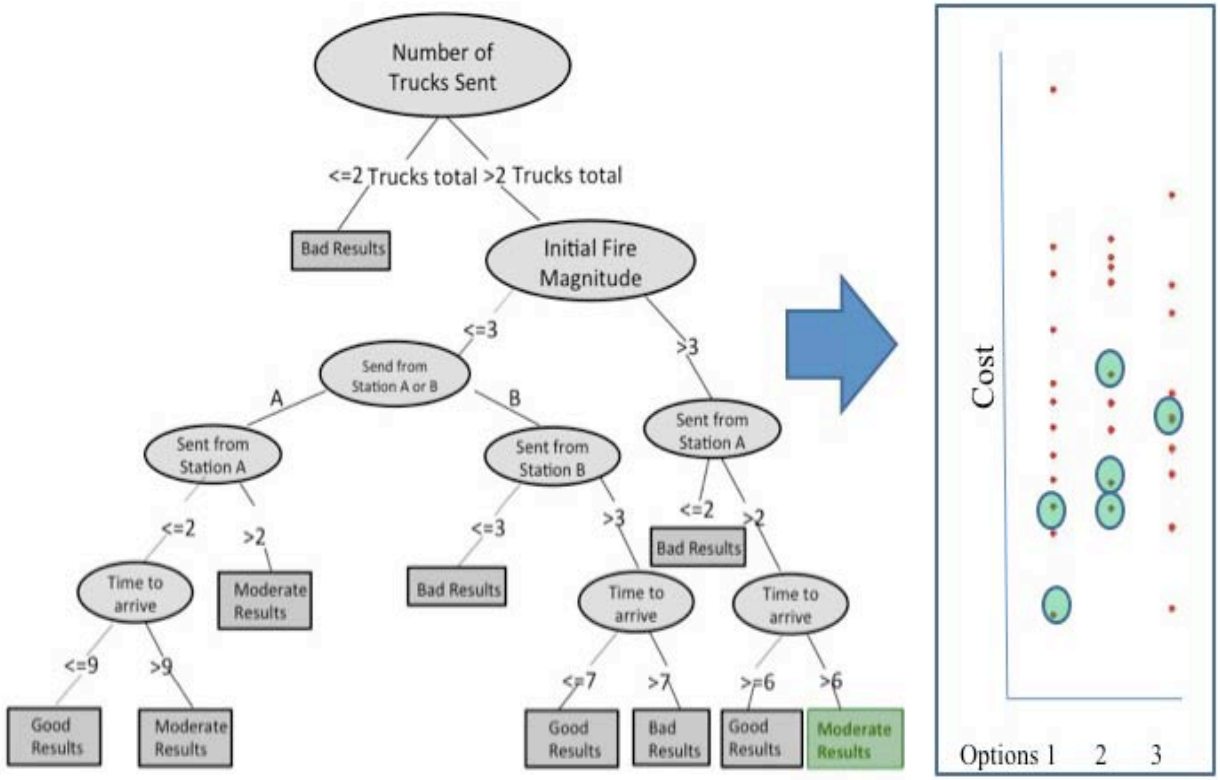


1.2.3 GUI support for data exploration



Starting Point	Number of Options Explored	
	Outcomes for Single Option → Conditions	Outcomes for Multiple Options → Conditions
	Conditions → Outcomes for Single Option	Conditions → Outcomes for Multiple Options

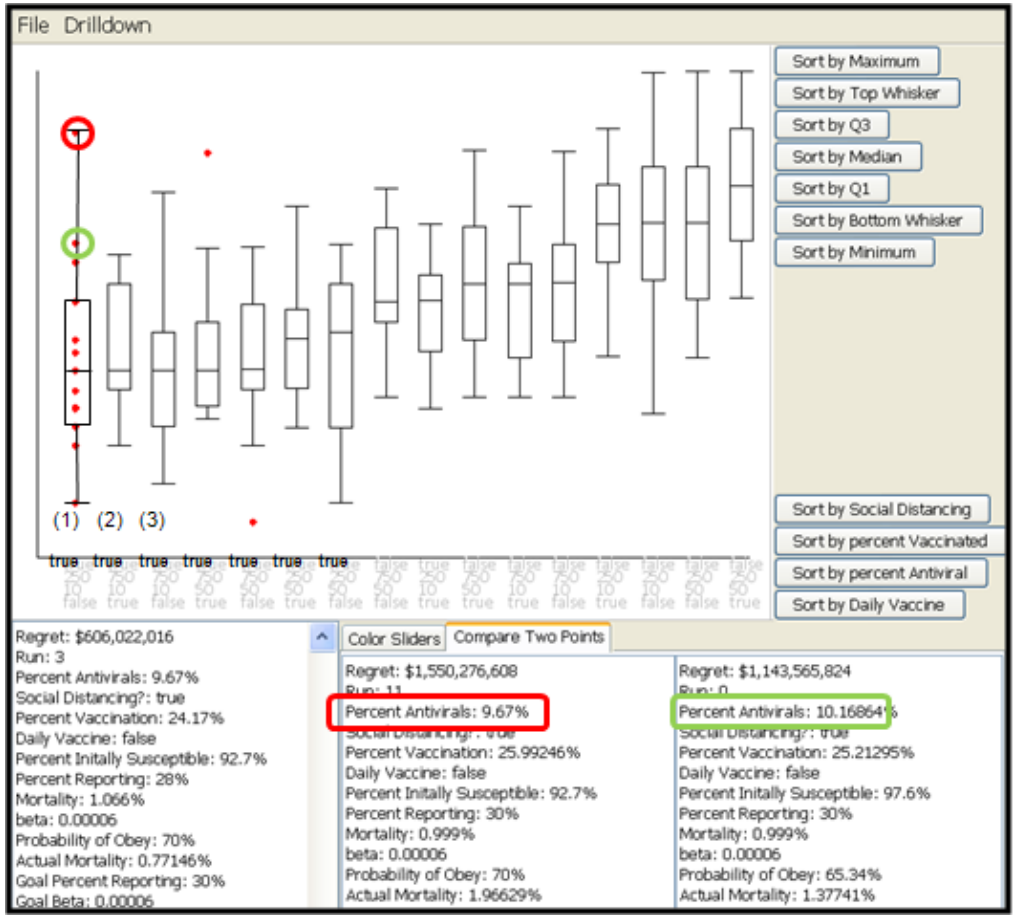
1.2.3 GUI support for data exploration (2)



Starting Point	Number of Options Explored	
	Outcomes for Single Option → Conditions	Outcomes for Multiple Options → Conditions
	Conditions → Outcomes for Single Option	Conditions → Outcomes for Multiple Options

Requirements: More GUI support for data exploration

Reqmt #	1.2.5 GUI Support for Data Exploration
Reqmt	The system shall display, upon user request, the conditions associated with at least two individual (user-specified) outcomes simultaneously.
Ration-ale	Comparing the conditions associated with outcomes can yield understanding.
OA levels	Applicable to all three levels.

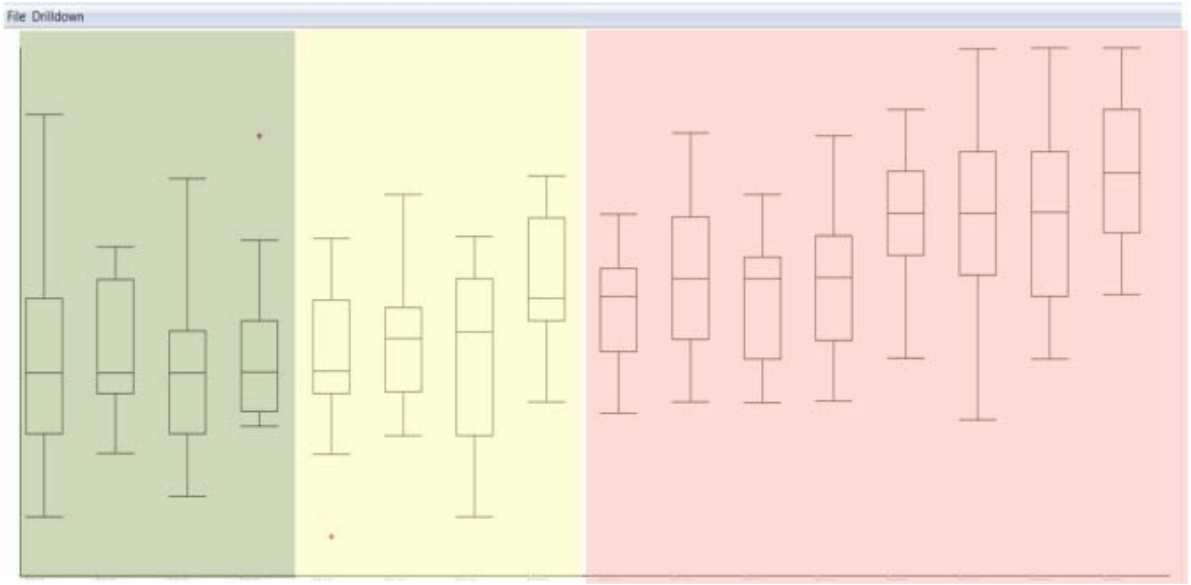


Example of earlier work: Pandemic Disease Spread Option Analysis Simulation Model – with Jennifer Mathieu

Requirements: More GUI support for data exploration

Reqmt #	1.2.6 GUI Support for Data Exploration
Reqmt	The system shall provide a visual or textual summary of the outcomes that result from the selected conditions in a manner that is comprehensible to the intended set of users.
Rationale	Summarizing this information enables choosing or modifying robust options.
OA levels	Applicable to all three levels.

Example of earlier work: Pandemic Disease Spread Option Analysis Simulation Model – with Jennifer Mathieu



Discussion and future plans

- **Discussion**
  - **Alternative visualizations could be substituted for box plots, scatter plots, and classifier trees**
  - **Requirements are not specific to a particular:**
    - **Visualization type**
    - **C^2 domain**
- **Future plans**
  - **Currently building a DSS in accordance with these requirements**
  - **Plan to run human-in-the-loop tests to determine the effect of level-2 and level-3 option awareness information**
  - **Also working with a more complex simulation model of state stability and insurgent recruitment**