UNLIMITED DISTRIBUTION

SYSTEMS ENGINEERING
Research Center

Enterprise Systems Analysis

Technical Report SERC-2016-TR-103

March 14 2016

Principal Investigator: Dr. Michael Pennock, Stevens Institute of Technology

Co-PI: Dr. William Rouse, Stevens Institute of Technology

Senior Researcher: Dr. Doug Bodner, Georgia Institute of Technology

Research Team:

Christopher Gaffney

John Hinkle

Christopher Klesges

Mehrnoosh Oghbaie

Contract Number: HQ0034-13-D-0004 Task Order: 010, RT138

The Systems Engineering Research Center (SERC) is a federally funded University Affiliated Research Center managed by Stevens Institute of Technology.

This material is based upon work supported, in whole or in part, by the U.S. Department of Defense through the Office of the Assistant Secretary of Defense for Research and Engineering (ASD(R&E)) under Contract H98230-08-D-0171 and HQ0034-13-D-0004 (TO 0110).

Table of Contents

Figures and Tables

1. INTRODUCTION

The overarching goal of the research in enterprise systems analysis has been to support decision makers and policy makers through the development of new modeling and decision support approaches. Enterprise systems are defined as those where there are multiple interacting organizations, but there is not central locus of control. As a result, change must often be the result of influence and incentives as opposed to command and control. To understand the potential impact of a policy option, one needs to capture the spread of potential future scenarios that may result. In particular, the long-term goals are to allow decision makers and policy makers to

- Identify the key drivers of system behavior and resulting outcomes
- Perform "what if" analyses
- Evaluate the efficacy of policy options to alter system behavior and outcomes
- "Test drive" the future
- Allow key stakeholders experience the behavior of the "to be" system

In support of these long-term goals, the objective of RT-138 was to evaluate and evolve the enterprise modeling methodology developed during RT-44 and RT-110. The primary mechanism to do so was the counterfeit parts case study. The focus of the case study was to understand the impact of hypothetical policy options to combat the intrusion of counterfeit parts into the defense supply chain. It is ideal as a testing ground for the modeling methodology because it involved all of the key features required of an enterprise systems: multiple interacting organizations (multiple government agencies, private corporations, and counterfeiters), no locus of control (the government can promulgate policy but suppliers don't have to supply, and counterfeiters can try to bypass), and the system is in the midst of a major shift in state (counterfeiting has rising dramatically in recent years).

Beyond the counterfeit parts case study, there were also a number of other subtasks that explored various issues related to the application of the methodology. Figure 1 depicts the subtasks of RT-138 and relationships among them.

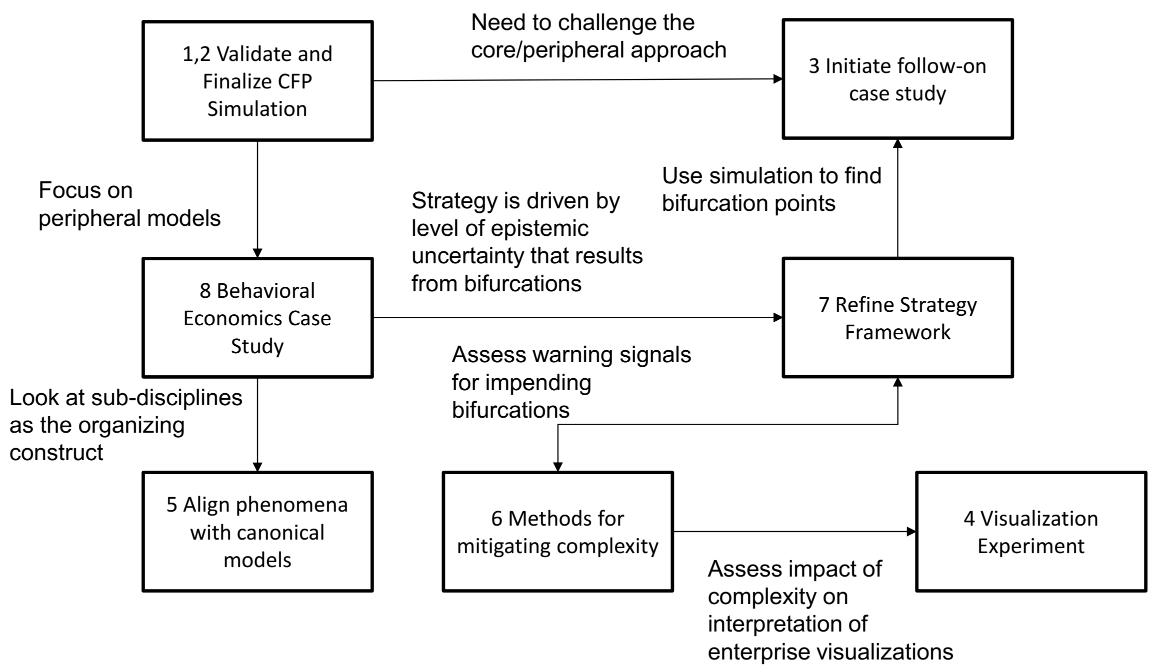


Figure 1 - Relationships among research tasks

In short, the challenge of modeling enterprise systems is that the intrinsic complexity of the underlying social systems fundamentally limits the ability to make precise predictions using models and simulations. The analysis of historical data and extrapolation from that data may be viable during periods of relative stability. However, social systems are prone to abrupt shifts behavior (sometimes referred to as bifurcations in the dynamical systems literature). This circumstance requires a different approach to employing models for decision making than that traditionally applied in engineering, which is essentially trend extrapolation. To that end, each of the subtasks were intended to explore how to address a different problem brought about by the complexity of the enterprise system. To summarize:

- The classical decision models used in engineering modeling treat humans as perfectly rationale decision makers. However, it is well known that this not an accurate representation of human behavior. Given, the central role of humans in enterprise systems, the *behavior economics case study* was intended to the apply research on modeling actual human behavior to an enterprise problem. The question was whether or not this would enable the detection of behavioral changes or if the effect would get lost in the "noise." Ultimately, the case study was conducted by examining the real world case of dynamically tolled roads for congestion management.

- Modeling enterprise systems necessarily requires the simultaneous consideration of the system from multiple perspectives. Given the nature of enterprise systems this often requires models from different scientific disciplines that were not intended to be integrated. Previous tasks (RT-44, RT-110) considered some of the challenges of composing such models. It is has been done successfully in some instances, but often times it proves difficult. Thus, the question is what allows one to reuse and compose models from different disciplines successfully. In the *Aligning Phenomena with Canonical Models* task, we considered how the nature of the phenomena in question affected such efforts.

- Assuming that one can build a model of an enterprise system that can be used to find abrupt shifts in the behavior of an enterprise system, the question remains how to use that information to support decision and policy making. During RT-110, we proposed a notional strategy framework to manage the resulting risks. In the *Refine Strategy Framework* task we examined the notional framework and linked the strategies to the source of the uncertainty.

- No model is capable of forecasting all possible scenarios. Consequently, some changes in enterprise behavior will be a surprise. Some researchers have investigated the possibilities of early warning signals of such "surprises" in biological and financial systems. In the *Methods for Mitigating Complexity* task, we critically review the literature on this topic and consider whether or not these techniques can be used to mitigate the impact of a surprise change in enterprise behavior.

- Once one builds a model of an enterprise system, there remains the concern of how the insights derived can be internalized by analysts, decision makers, and policy makers. Many have employed interactive visualizations in such situations, however, it is unclear how effective they are. Given the complexity of enterprise systems, there is a very real possibility that the decision makers will perceive spurious correlations as causal. To study this concern in more depth, the *Visualization Experiment* task involved a human subjects research experiment where subjects were asked to use an interactive interface to diagnose the contributing factors that led to an enterprise failure.

- Finally, a single data point is never sufficient to validate an approach. Consequently, while the counterfeit parts case study provided a single test case for the enterprise modeling methodology, it is not enough to validate it. It is entirely possible that the results were artifacts of unique feature of the case. Consequently, the *Initiate Follow-on Case Study* task was intended to investigate another enterprise system that could serve as a second test case for the enterprise modeling methodology in an entirely different context.

The remainder of this report discusses the results of these research tasks. First are the two case studies. Section 2 describes the counterfeit parts case study, and Section 3 describes the behavioral economics case study. Section 4 considers how models of such cases are affected by model composition and reuse issues. Section 5 critically reviews the early warning signals literature for applicability to enterprise systems. Section 6 draws together the implication of the previous sections and considers how that affects enterprise modeling and the strategy framework. Section 7 explains the visualization experiment. Section 8 revisits the enterprise modeling methodology based on everything that has been learned through RT-138. Section 9 presents the preliminary description of the follow on to the counterfeit parts case study. Finally, Section 10 concludes the report and discusses future work.

2. COUNTERFEIT PARTS SIMULATION

For a case study demonstration of the enterprise modeling methodology, the problem chosen was that of counterfeit parts in the Department of Defense supply chain. Counterfeiting, particularly of electronic components, has become a major issue over the last 10-15 years (ABA, 2012; AIA, 2011; Business Insider, 2012; DoC, 2012; Economist, 2012; McFadden & Arnold, 2010; Pecht & Tiku, 2006; SASC 2012; Stradley & Karraker, 2006; Villasenor & Tehranipoor, 2013). A number of counter-measures have been proposed (DAU, 2013; DoD, 2011; DoD, 2012; DoD, 2013; DoD, 2014; GAO, 2010; GAO, 2011; GAO, 2012a; GAO, 2012b; Guin et al., 2014; Livingston, 2007a; Livingston, 2007b; Livingston, 2014; SAE, 2014). Many relate to reviewing suppliers to determine legitimacy, applying penalties to those who pass counterfeits, and developing new methods to test counterfeits to keep up with the increasing quality of counterfeited parts.

Counterfeits come in many types. Fraudulent counterfeits can be recycled and passed as new, remarked and passed as new or different grade, defective and passed as functional. Fraudulent components may be functional, but not cared for appropriately (e.g., not contained in static-proof containers), but have forged paperwork indicating appropriate care. Components may be considered counterfeit if they were overproduced by a legitimate contract manufacturer above and beyond the limit set by the trademark holder. Malicious counterfeits are intended by an adversary to do harm. For example, a genuine component may be tampered to provide a back door or fail under certain circumstances. Cloned components use reverse-engineered designs and can be either fraudulent or malicious.

While counterfeit parts are a real problem, we must ask why this is an enterprise problem. There are a number of features that make this an enterprise problem.

- There is no locus of control.
    - Multiple agency/industry stakeholders are addressing the problem, ranging from DoD and its programs, to suppliers, to the Department of Justice (DoJ) and Customs and Border Patrol (CBP).
    - DoD can promulgate policy, but it must be cognizant of reaction from industry base (e.g., diminishing supply base).

    - Programs have methods of addressing counterfeits on their own.
    - Legislation plays a role.
- There is significant adaptive behavior.
    - Counterfeiters adapt to new technology and new policies.
    - Policy-makers must and do adapt to these new strategies.
    - Legitimate suppliers may adapt to new policies.
- There is significant complexity.
    - There is substantive socio-technical behavior (human behavior and social behavior interacting with technical system).
    - There are multiple systems interacting with unpredictable effects.

An initial model was developed previously (Bodner, 2014; Pennock et al., 2015). Here, we overview the model and recent enhancements and discuss its use (Bodner, 2015). Then we summarize several reviews of the model. These reviews were undertaken to attain a sense of the model's validity and usefulness. They involve a number of different stakeholder sets, ranging from academics affiliated with the SERC, to experts from MITRE who transition research results, to DASD(Systems Engineering), to the broader community of experts engaged in anti-counterfeiting. An underlying question is whether such an enterprise model would be useful in a general sense. Based on feedback from these reviews, we then present discussion on transition planning.

Model Development Summary

This model was developed using an enterprise modeling methodology advocated by Pennock and Rouse (2014). While the various steps in that methodology were followed more or less (Pennock et al., 2015), the model development process did not utilize the strict multi-level modeling formalism envisioned in that methodology. Rather, elements from the multiple levels were combined into a core model, with other elements structured into exogenous models. Policies were modeled somewhat separately, due to the need to engage users with an interface to explore and analyze the effects of different policies.

As reported in Pennock et al. (2015), a number of subject matter experts were involved in discussions that helped shape model development. These experts represented the following organizations and agencies.

- DASD(Systems Engineering)
- DASD(Logistics & Materiel Readiness)
- Defense Procurement and Acquisition Policy (DPAP)
- A prime contractor
- A component supplier
- Obsolete parts manufacturers via government trusted sourcing
- An electronics industry consortia group
- Customs, law enforcement and counter-intelligence

- Subject matter experts on counterfeit parts

MODEL ARCHITECTURE

The original multi-level modeling methodology envisioned four levels that constitute a conceptual model for an enterprise. These levels are the following.

- Eco-system
- System structure
- Delivery operations
- Work practice

In the counterfeit parts domain, the eco-system is the national security organizations of the government, plus the defense industrial base, plus the electronics market and counterfeiters. The system structure constitutes the relationships between the various actors, such as government contracting with primes, suppliers contracting with other suppliers, and policies from one organization impacting others. The delivery operations level consists of the flow and delivery of parts, sub-systems and systems through the supply chain to their eventual deployment in inventories and operation systems. This is governed by the system structure. Finally, the work practice level consists of individual operations with the supply chain, such as manufacture of components, inspections of component lots at a Customs station, or retrieval of a part from inventory to perform a repair.

These four levels served as the basis for visualizing various phenomena in the counterfeit parts problem, as well as their interactions, as shown below in Figure 2.

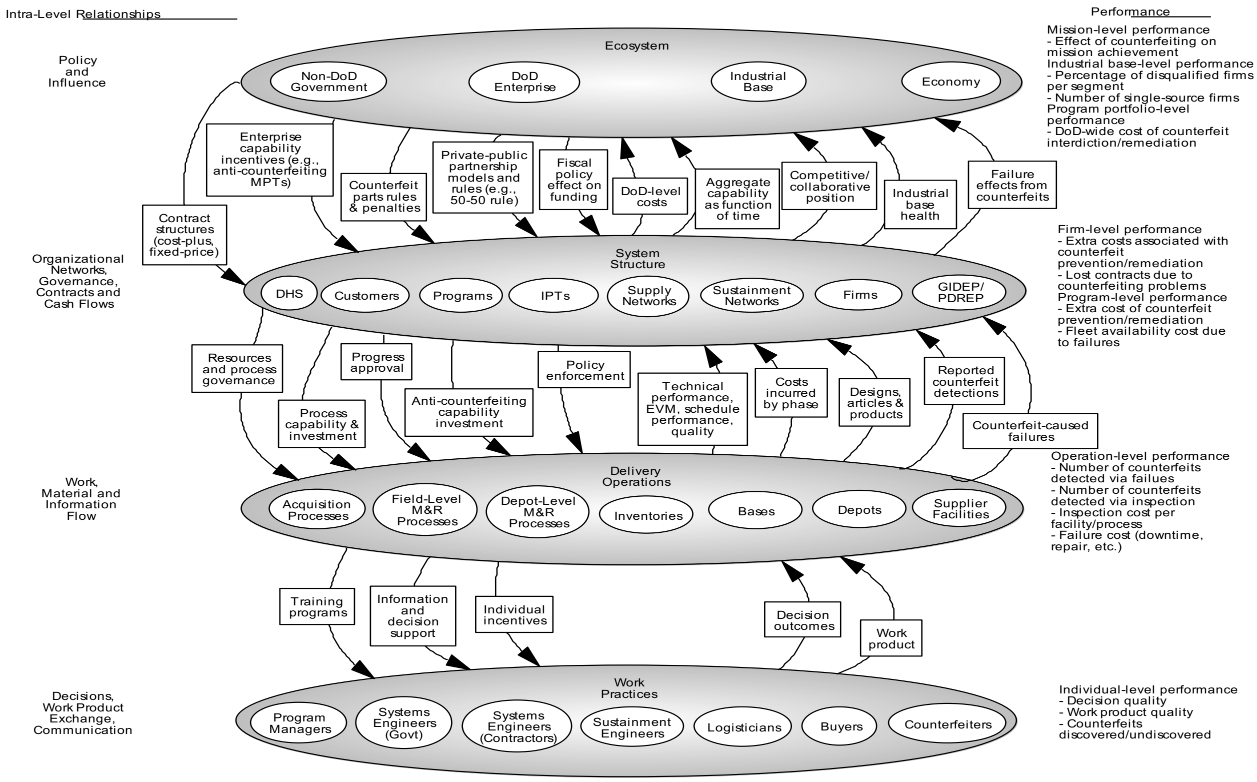


Figure 2 - Visualization of counterfeit parts in an enterprise context

However, when the counterfeit parts enterprise model was developed, it did not utilize four levels with interactions between them. Rather, it focused on a core model consisting of elements from the system structure level, the delivery operations level, and the work practice level. These levels were contextualized as a supply chain model, with the following characteristics.

- The system structure level represents the network of customers, suppliers, manufacturers and wholesalers in a large-scale supply chain.
- The delivery operations level represents the flow of components, sub-systems and systems in both acquisition and sustainment. Thus, new systems are populated with parts and components from this supply chain in acquisition, and likewise, fielded systems are repaired and maintained with parts and components in sustainment.
- The work practices level represents individual operations within the delivery operations level.

When the model was designed, it reflects these phenomena in a main, or core model, that includes the military systems subject to counterfeit infiltration (e.g., fighter planes, submarines, tanks) and their constituents (sub-systems and components), the supply chain operations that produce and deliver these systems and constituents, and the enterprise actors that direct the supply chain operations (suppliers, government agencies that order from them, and counterfeiters).

An exogenous model represents those phenomena in the eco-system level that impact elements of the core model, but are not directly part of the enterprise affected by counterfeit parts. In many instances, these phenomena are economic, technological or societal in nature.

A policy model was then designed to focus on the various policies that the model would be used to evaluate. The policy models reflect the various policy-making agencies that can influence the enterprise in addressing counterfeit parts, the policies that they can enact, and the base state of policies (or lack of policy). The policy model interacts with the core model by enabling, preventing or influencing behavior of its elements. It interacts similar with the exogenous model. At present, the policies in the policy model are user-controlled, in that a user specifies which policies are enacted when, and for certain policies at what level. Thus, a user may react to certain behaviors or results from either the core model or exogenous model and change policies. In the future, the policy model could be refined to make it similarly reactive.

Figure 3 illustrates the architecture of the counterfeit parts model.

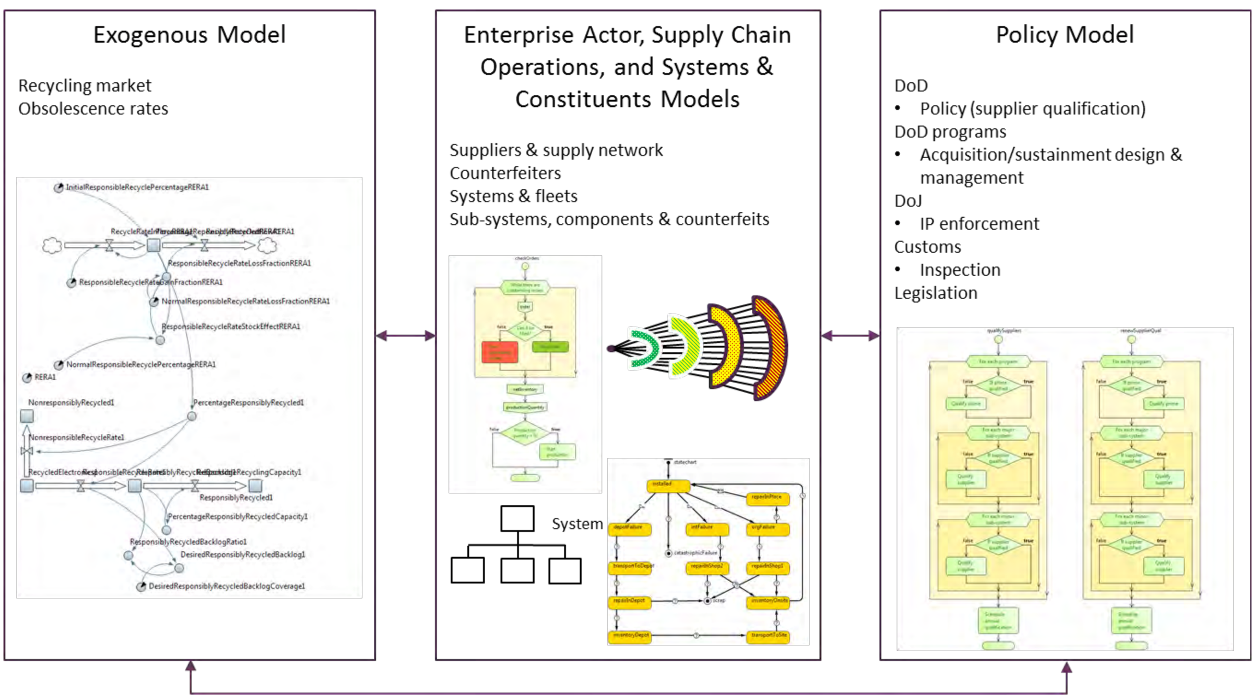


Figure 3 - Model architecture

We now briefly discuss the five sub-models.

Systems & Constituents Model

The systems and constituents model addresses the structure of DoD systems via a work breakdown structure (or bill of materials) and the behavior as a state-chart that addresses failures, maintenance and repair. Each system has major sub-systems, then minor sub-systems, and the components. The electronic components are susceptible to counterfeiting. The reliability and resulting system availability is affected by counterfeit parts. The structural and behavioral models are shown in Figure 4.

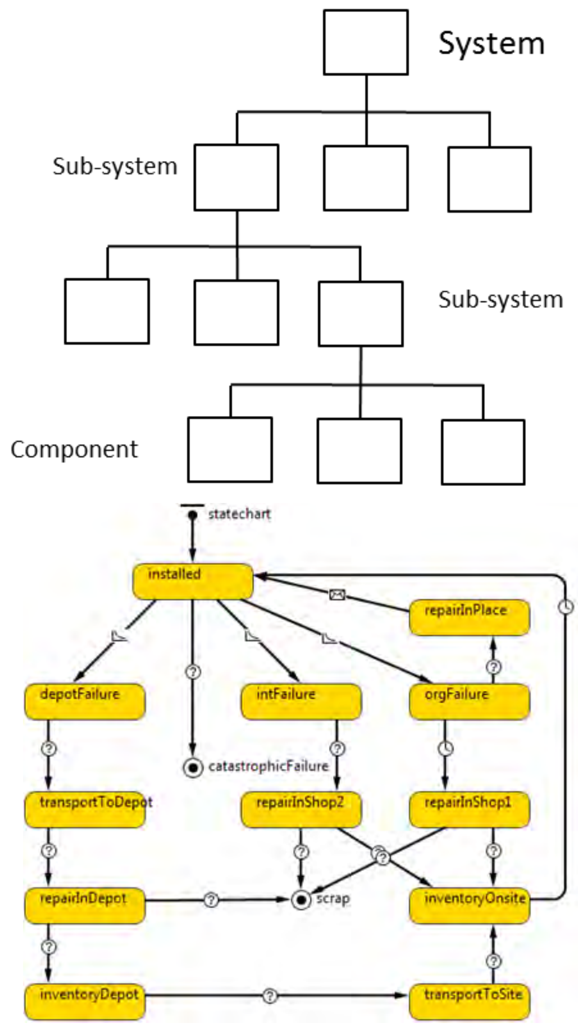


Figure 4 - Systems & constituents structure and behavior

Originally, this model represented two programs. It has recently been enhanced with an additional program.

Supply Chain Operations Model

The supply chain operations model represents the various locations in the supply chain, in terms of factories and warehouses, and their networked relationships for purposes of part flow. Component manufacturers are the starting points. They fabricate components and then feed then to sub-systems manufacturers, who assemble components into sub-systems. Minor sub-systems are sent to other manufacturers that assemble them into major sub-systems.

In acquisition, major sub-systems are sent to lead systems integrators for assembly into new systems. In sustainment, components and sub-systems are sent to depots for maintenance and repair purposes. Note that components imported from abroad are subject to inspection by Customs. DoD may employ control points where items are inspected before admission into the DoD supply chain from commercial sources.

Distributors may import components and/or source them from other domestic suppliers in case the original component manufacturer has exited the market. Distributors may also exist when the original component manufacturers are still producing.

The supply chain operations model contains typical inventory and inventory reorder models in the multiple tiers of the supply chain. Figure 5 illustrates the structure and behavior of this model.

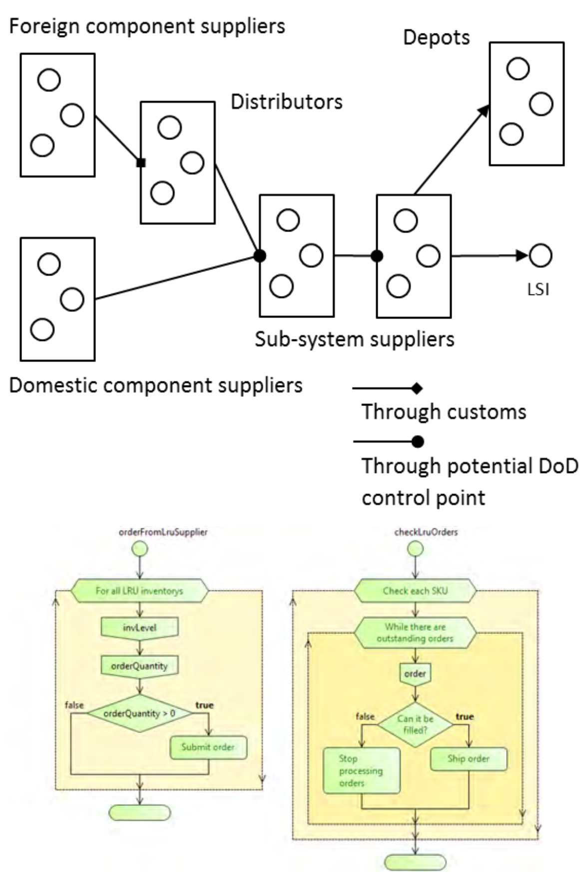


Figure 5 - Supply chain operations model

ENTERPRISE ACTOR MODEL

The enterprise actor model contains models of the various actors in the enterprise, mainly suppliers (legitimate and counterfeiter), as well as government organizations that contract with suppliers for purchasing. This model addresses the behavior of these actors over time, including reactions to changing conditions. Original component manufacturers (OCMs) and original equipment manufacturers (OEMs) may exit the mark if their margins are too low. Counterfeiters may become more sophisticated over time in the types of counterfeits that they produce. If supply of electronic waste is reduced, counterfeiters may transition from recycled counterfeits to more sophisticated counterfeits such as clones. Figure 6 shows example decision logic associated with enterprise actors.

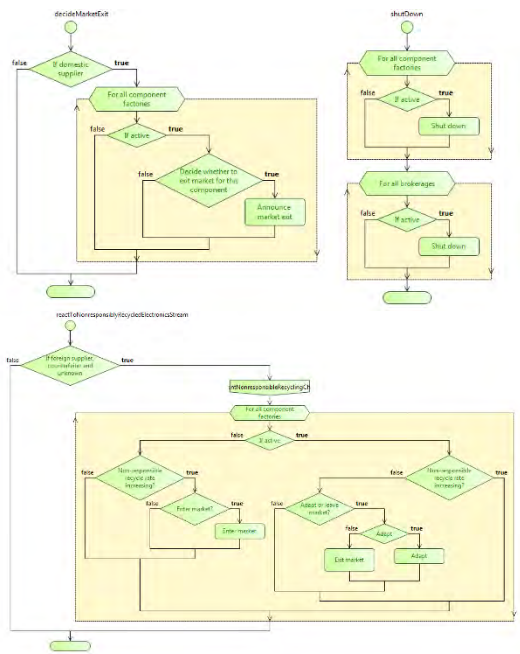


Figure 6 - Enterprise actor model decision logic example

POLICY MODEL

The policy model contains the set of actors that can promulgate policy, as well as the effects of policy enablements in terms of state changes, new business rules and delayed effects that propagate to the other models. For instance, central DoD policy-making and DoD program-level policy making are represented. In addition, other policy-making agencies outside DoD are represented, such as Department of Justice and Customs and Border Patrol. Figure 7 illustrates typical decision logic used in the policy model.

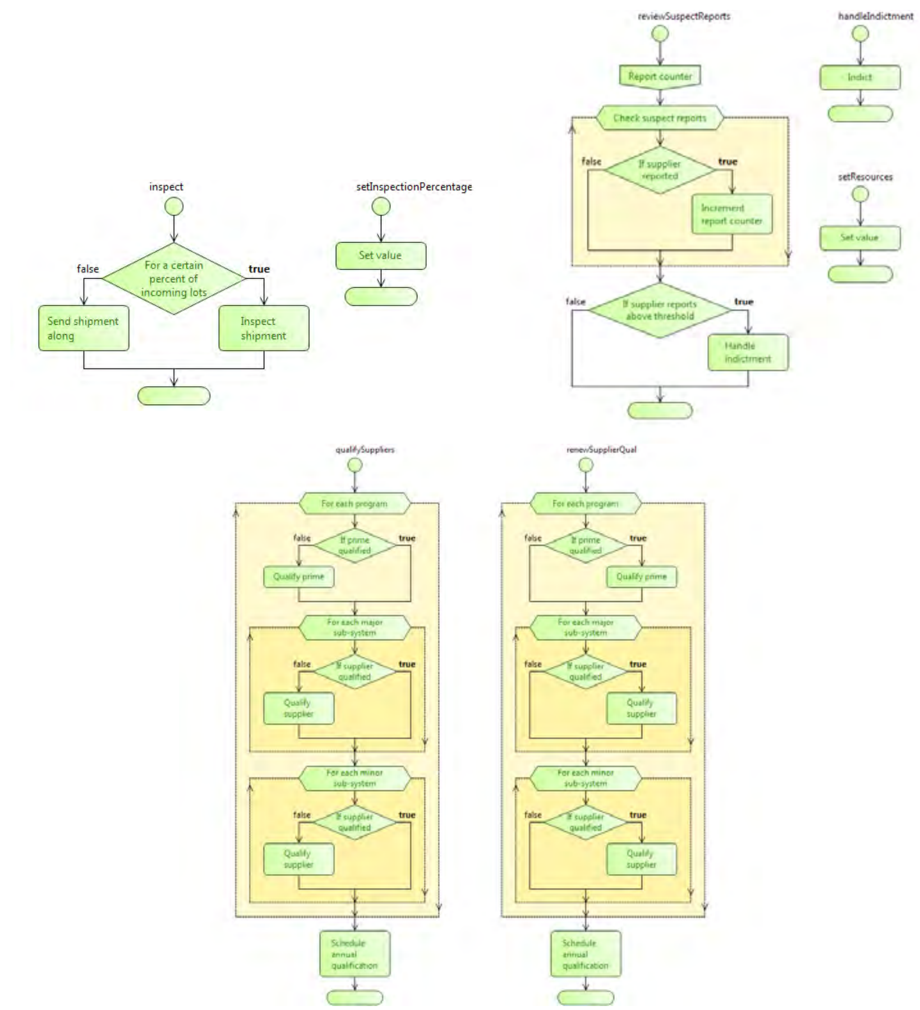


Figure 7 - Enterprise actor model decision logic example

The following anti-counterfeiting policies are implemented in the model.

- Supplier qualification with criticality levels (DoD level)
- Acquisition
  - Supply chain design via sourcing to lock down reliable long-term suppliers
  - Design refresh planning to prevent obsolescence (not implemented fully)
- Obsolescence management in sustainment
  - Design refreshes in sustainment to prevent obsolescence
  - Lifetime buy of soon-to-be obsolete parts
- Customs policies
  - Inspections frequency
  - Cooperation with IP holders to determine legitimacy of suspect counterfeits
- DoJ policies
  - Resources devoted to counterfeiting prosecutions
  - Priorities in IP prosecution

- Electronic waste export legislation

Exogenous Model

The exogenous model primarily addresses technology change rates and the effect of electronics recycling export. The electronics recycling market was added recently. It is represented in system dynamics as a supply-demand system with potential restrictions. One of the major drivers of counterfeit parts is the importation of electronic components that originally were exported as waste to non-OECD countries, then "recycled" via remarkings to pass as genuine new components.

If Congress passes legislation aimed at restricting the export of U.S. electronics, the model creates a larger U.S. market for responsible electronics recycling, with delays. It should be noted that subsidies are not present in this model, but they would be an additional policy level to quicken the process. Figure 8 shows the system dynamics model of the recycling market.

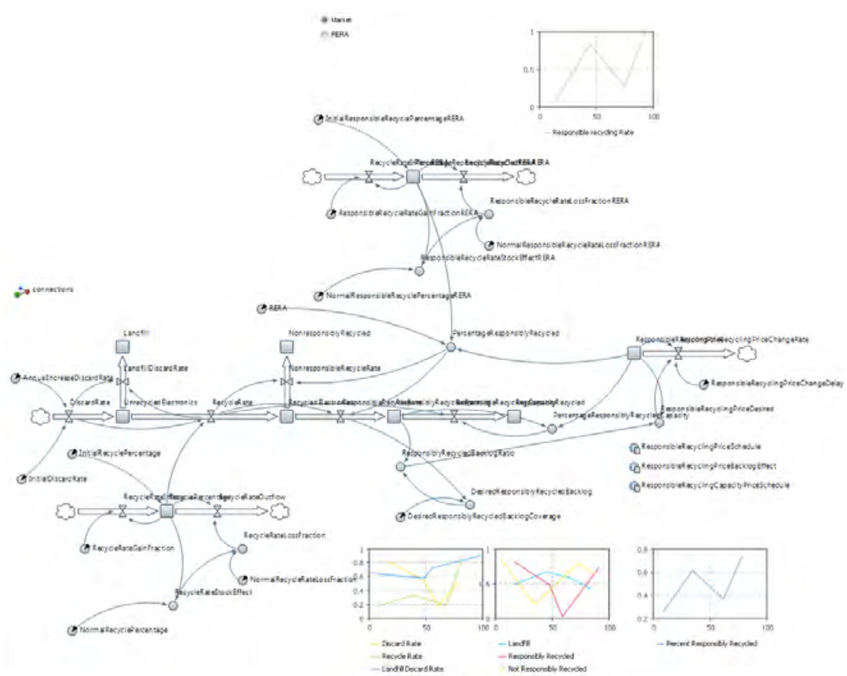


Figure 8 - Recycling market model

Implementation

The model is implemented using AnyLogic® 7 simulation software. AnyLogic supports discrete-event, agent-based and system dynamics simulation. Thus, it is useful for potential model composition. In addition, it provides an API for Java class extensions and is therefore useful for development of reusable model component libraries.

The model is implemented primarily as an agent-based model, with complex agents for enterprise actors and policy actors, and simple agents for systems and components. System dynamics models are used for external influence factors in the exogenous model.

This approach is reasonably well-suited suited to handle multi-scale enterprise modeling and organizational decision-making. However, too many agents could require excess computational resources, making the model unsuitable for small computing platforms.

CHALLENGES

An enterprise model faces a number of challenges in its design and development. The counterfeit parts enterprise model faced the following challenges.

- *Multiple scales of resolution*. This particular model includes phenomena from macro level (agencies and organizations, economics) to the micro level (electronic circuits). There are implications for computational performance and potentially for data and parameters (i.e., consistency across multiple scales of resolution). If the model is scaled up for transition, the computational performance will be an important issue. The model may not run well on a laptop due to memory limitations. In this case, it could be reconfigured to run on a server with browser client interface, for example. The consistency of data is another issue that would need full exploration in transition with a stakeholder dataset.

- *Multiple stakeholder perspectives*. It is typical to have a diverse set of stakeholders for any enterprise modeling effort in terms of their interests and perspectives. In such an environment, it is challenging to ascertain overarching important issues, derive how they interact, and on the other hand include issues of importance to each stakeholder. The anti-counterfeiting roundtables provided an effective way to determine many important issues and interactions, plus include issues of importance to the stakeholders.

- *Multiple data sources/representations*. Any enterprise model will require data from multiple sources using multiple representations. This is one of the motivations of the overall enterprise modeling and analysis research effort. Clearly, data would come from different stakeholder organizations, and once again consistency is an issue. This is mitigated by the agent-based representation used for the model. Each of the complex agents can use data specialized to the organization that it represents. The overall enterprise interaction model must be generic, though, in the sense of being independent of data representations used by particular organizations. This has been achieved at least as a first order result. Detailed transition will test this result. In addition to heterogeneous datasets from different organizations, enterprise models typically include human/organizational decision-making, processes, and technical behavior. The decomposition of the model into agents representing organizations, supply chain process behaviors, and agents representing technical behavior of systems and constituents has largely addressed this issue.

USAGE

The model is designed to be used by multiple stakeholders representing various agencies and organizations in the overall enterprise. In this sense, it would be used for scenario and what-if analysis. Since multiple stakeholder and policy options are available, one of the important aspects to consider is the timing and order of policy enablement.

In addition to scenario and what-if analysis, the model can also be used for purely experimental purposes. For instance, what are the effects of individual policies on system availability versus policy cost? What are the interaction effects between different policies? The timing and ordering of different policy enablements could also be an experimental feature. This would create a complex response surface analysis problem, but it may produce interesting results.

Another way to conceptualize model users is to consider producers of results versus consumers of results.

- Producers of results
    - Analysts in stakeholder communities (DoD systems engineering, logistics, policy; other agencies; industry)
    - Perform experiments and sensitivity analysis
    - Perform analysis for consumers
    - Primarily looking for quantitative results or relative effects
- Consumers of results
    - Policy makers in stakeholder communities
    - Test different scenarios and policies to see effects and interactions
    - Primarily looking for insight

Figure 9 depicts the current model interface. The upper left contains the policy dashboard. These items, in the grey box, allow the user or users to enable or change the level of different policies among different stakeholders. Different stakeholders include the DoD itself, individual DoD programs, the Department of Justice, Customs and Border Patrol, and Congress. Each stakeholder has various policy options.

The upper right features a scenario dashboard. These are model features that are not controllable by the stakeholders, but may be of interest as experimental variables (e.g., percentage of foreign suppliers that are counterfeiters).

The bottom two levels constitute the status dashboard (in blue boxes). The upper one is the DoD state, while the lower one is the state for enterprise actors outside of DoD. The model assumes that DoD policy cost and system availability outcomes occur at the program level. In addition, the status dashboard shows the number of counterfeit lot suspects and escapes that each program experiences. For CPB, the status dashboard shows the number of counterfeit suspects and escape experienced, plus the policy costs incurred. The DoJ status dashboard shows the number of indictments of counterfeiters, plus the policy costs incurred. The electronic waste status dashboard shows the number of toms of electronic waste being exported over time.

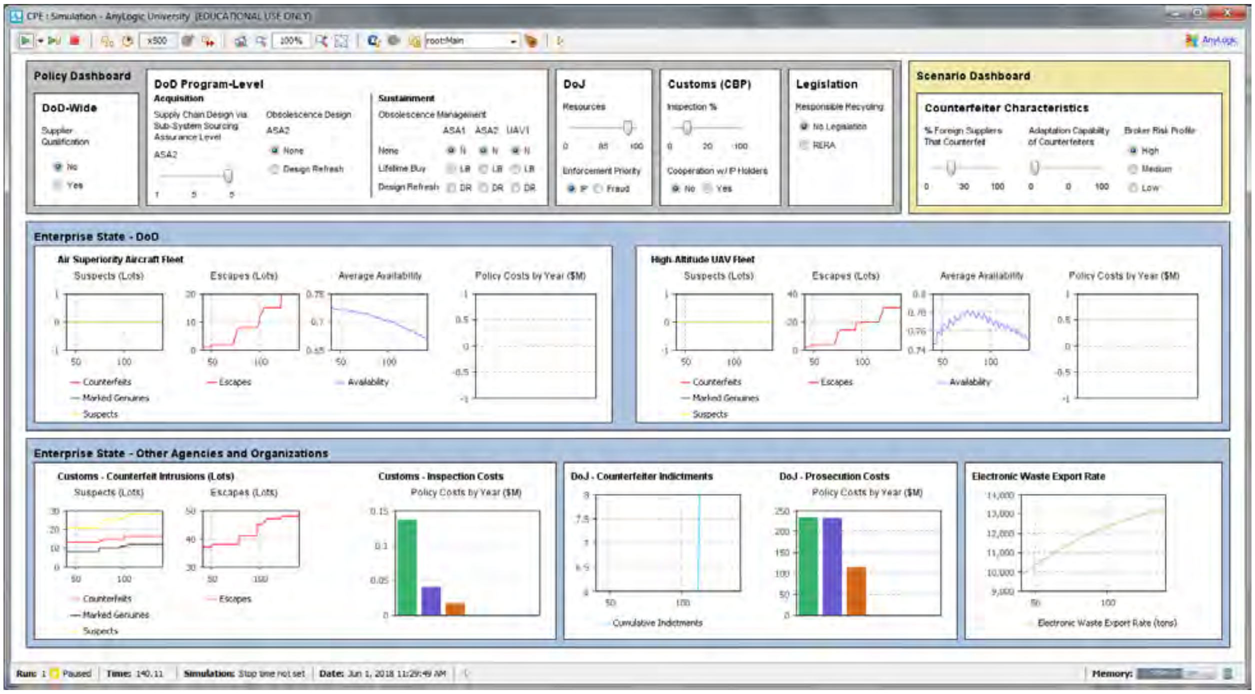


Figure 9 - Model interface

EXAMPLE ANALYSIS

In the simulation run reflected in Figure 9, the DoD and DoD program policy options are not enabled. Hence, there are no policy costs incurred. The number of escapes (counterfeit component lots passing into the program) are shown over time for each program. Since qualification/testing are not enabled, there are no suspects reported at the program level. Customs inspects 20% of incoming lost randomly, and the number of suspects and escapes are shown over time. It should be noted that many of the escapes from Customs are still in the supply chain and have not made it yet to a program. Finally, DoJ has issued several indictments during the simulated time period.

Currently, the model is populated with test data. The goal is to provide data that is realistic for a given scenario using a reasonably generic data structure. The motivation here is that real data is difficult to verify for many aspects of the model due to the sensitive nature of the problem, distributed nature of data across multiple agencies, and lack of knowledge about counterfeiters and their operations. Thus, an analyst would be responsible for populating the model with data from their scenario of interest.

As an example of the analysis that can be performed by the model, consider the four scenarios.

- Scenario 1 – Baseline scenario
    - No supplier qualification
    - No obsolescence management
    - Customs inspects 20% of incoming
    - Baseline DoJ enforcement resources
- Scenario 2 – Baseline scenario plus supplier qualification for all sub-systems
- Scenario 3 – Baseline scenario plus increased resources for prosecution (50%)
- Scenario 4 – Scenarios 2 and 3 combined

Table 1 shows average results for the various metrics over the four different scenarios for a ten year period with ten replications. It should be kept in mind that since the model uses test data, no real inferences can be made on the results. This is only illustrative of potential analysis that can be run, such as an analysis of variance to determine policy effects and interaction effects. It does illustrate that enabling supplier qualification has an effect on reducing escapes into programs.

Table 1 - Example model analysis (notional data)

Model outputs (averaged over ten replications)	**Scenario 1**	**Scenario 2**	**Scenario 3**	**Scenario 4**
Escapes – fighter jet program (lots)	56.3	13.8	53.8	12.1
Suspects – fighter jet program (lots)	0	72.3	0	70.3
Policy cost – fighter jet program ($M)	0	30.4	0	30.7
Escapes – UAV program (lots)	51.7	11.6	48.9	11.4
Suspects – UAV program (lots)	0	69.7	0	65.2
Policy cost – UAV program ($M)	0	31.9	0	32.5
Escapes – Customs (lots)	640.2	636.0	635.2	632.1
Suspects – Customs (lots)	595.3	608.7	598.8	580.5
Policy cost – Customs ($M)	56.1	55.7	57.9	57.1

Indictments – DoJ (lots)	0	0	65.4	66.5
Policy cost – DoJ ($M)	0	0	53.6	52.1

Academic Peer Review

One of the tasks for this research project was to conduct a peer review of the counterfeit parts enterprise model with representatives from SERC-affiliated universities. This review was conducted on September 25, 2015, and it included representatives from Massachusetts Institute of Technology, Purdue University and Stevens Institute of Technology.

The review consisted of a summary overview of the overall research task, then a presentation on the counterfeit parts enterprise model and a demonstration of the model's use. Then the review was conducted along the following main lines.

1. Validity — the extent to which the simulation is technically correct relative to the purposes for which it was developed.
2. Acceptability — the extent to which the simulation addresses problems in ways that are compatible with current preferred ways of decision-making and/or potentially useful new ways of multi-stakeholder decision-making.
3. Viability — the extent to which use of the simulation for the purposes intended would be worth the time and effort required.

In addition, the questions below were posed as potential secondary topics on which participants could comment.

1. Is this a useful way to model an enterprise problem?
2. Are there corrections needed in the current model?
3. Are there important aspects of the counterfeiting problems not currently modeled?
4. Are there additional decisions for which the analyst should have controls?
5. To what extent is this replicable to another enterprise model?
6. To what extent does/can methodology facilitate building these types of models?

The discussion from the review is summarized below.

Validity

- Savings from successfully reducing counterfeits should be incorporated, for example reduced repair bills, fines from successful prosecutions, etc.
- How does the behavior of the counterfeiters change over time?
- It seems like the major trade is availability vs cost.
- How do the lags or interaction effects occur when a policy configuration is changed?
- There might be cases where one would want to order policy decisions over time intentionally. For example, it may make sense to change the Customs interdiction policy before changing the sustainment policy.
- It might be a good idea to have some human players in the loop to try to represent the counterfeiters. Sometimes the interaction is more important than the numbers that come out of the simulation.
- Does the modeling methodology address the creation of phenomena? It looks like that is happening through the entry of agents.
    - The methodology is iterative, so new phenomena that are noticed can be included.
- The approach looks good, but as backup plan to getting an overall dataset for validation, a backup would be to validate behavioral model for the actors.
- What is the negative value of a counterfeit part?
- Have currency counterfeiters been examined to determine what parallels there may be?

Acceptability

- It may be useful to try different approaches with stakeholders and see how they react. Some could be told very little about how the model works; others could be told underlying model details; others could be told to expect limited outputs.
- What is the current state of the art that people are currently using in this problem space? Do DoD decision-makers have access to policy tools?
    - We are mostly aware of focused tools and methods such as methods to analyze the effectiveness of counterfeit suspect testing (Cohen & Lee, 2014).
- This model does seem useful for supporting policy analysts.
- One issue seems to be the struggle between snapshots in time versus dynamics in terms of model output. It could be useful to have thresholds on the graphs to give some sense of what is good and bad. Would a policy analyst be able to understand what a level of an output is good or bad?

- A past workshop addressed when you should make things very transparent versus a high-level view (Rhodes & Ross, 2015). A high-level person may not want to see the details but still needs to trust to the model. This report discusses the idea of use cases, which could be helpful. One conclusion is that it is challenging to transition tools. The report also discusses methods to trace back causal connections from a terminal event (e.g., part failure) to a spontaneous initiating event (e.g., a policy).
- It is a good idea to identify "hot spots" for user interaction. There may be non-policy factors that are important.
- It might be worth creating a class to let students investigate this problem

Viability

- How much effort would it take to put this on someone's desktop?
  - There's a learning curve within a particular domain such that models take less time to develop as you create more of them.
  - There's a model maintenance cost, since the world changes, and the model must be updated to reflect changes. This was an important take-away from the anti-counterfeiting roundtables, since the stakeholders see new things frequently.
  - There's a lot of value in having a tool that lets you get rid of bad ideas.
- We want to transition knowledge and insight, not just tools.
- This looks like a comprehensive effort, especially using the roundtables to get different perspectives
- The government is dealing with many similar problems. It would be beneficial to hold a review or roundtable with other potential users.
- How quick do we need to be and inexpensive to make this viable?
- Videos of recorded presentation would be useful for classes.
- One of the reasons that the Department of Defense is interested in counterfeiting is the impact on operational readiness which is hard to monetize. Is there a way to show operational readiness as an output?
  - System availability is a proxy for operational readiness. Other metrics could be developed, including trust and resilience.
- How can the tragedy of the commons be addressed? No one owns this problem. How much would it to cost to set this up in the future? That would help decision-makers assess whether to make further investments in this type of modeling.

MITRE PEER REVIEW

The model was next presented to a set of subject matter experts from MITRE on November 15, 2015. These experts work with an array of government agencies, including the Army, the Air Force, the Department of Treasury and Internal Revenue Service, and the Veterans Administration. In addition, representatives from MITRE's Systems Engineering Technology Center were present.

This review used a format similar to that of the academic peer review. However, this review focused on the acceptability of the model and its viability and transitionability due to the expertise of the participants in working with various government customers and transitioning results. The model's validity for this group of subject matter experts essentially served as a component of the acceptability, as the participants are not experts on counterfeit parts. The discussion is summarized below.

Validity and Acceptability

- It would be beneficial to have a tool that can inform policy. This would be useful to MITRE's customers. For instance, is it better to pursue enforcement or education? A design of experiments around this question would be interesting.
- The effect of randomness in the various agents should be explored, especially in terms of agent interactions and bad actors. This is another area in which a design of experiments would be valuable.
- In the Treasury domain, they use a tensions chart to model competing tensions. It would be interesting to look at trade-offs from that perspective.
- The cost of a policy is critical both for validity and acceptability. This is normally determined by a standard cost model and an economic policy model.
    - It was noted that the current model does address policy costs. A standard cost model is needed as input data.
- The Veterans Administration has complex business processes. These types of processes should be integrated into the model.
    - Similarly, a supply chain has complex business processes that are captured in the current model. There is capability to model business processes and rules.
- How are multi-scale effects addressed? Is there a common taxonomy?
    - There is a common taxonomy largely centered on supply chain related phenomena. This allows modeling at the enterprise level down to the component (i.e., integrated circuit) level.
- Is there a formal model of metrics? The formal model of metrics focuses on system availability.

- Cost is included, but needs to be enhanced via a standard cost model.

- Looking through the lens of the policy maker, is the model fidelity such that it will help craft a better policy than ones can be found via common sense? How would a policy-maker know that the implication is believable?
- In a complex system/enterprise with interacting elements, availability versus cost are good metrics.
- Validity is in the eyes of the beholder. It would almost an inductive process for each group of stakeholders to establish its own validity. It is important to work closely with this group to establish validity.
- How could a policy go bad? One valuable use of the model is to find out negative outcomes. Would a policy trigger the opposite of what is intended? What is the range of possible outcome spaces?
- If something unexpected happens, how do you diagnose it in the model, in terms of being a mistake in the model or a real effect? If you have the underlying framework to do this kind of analysis as a sandbox, that is valuable.
- Counterintuitive results are of great value, if they can be produced and validated.
- Being able to incorporate business process models is important, since most organizations have them. This would aid with validity and acceptability.
- Many decision makers in government are currently using point models but not this type of model that encompasses tiered levels and policy abstractions.
- One major area of concern in government is bridging the gap between IT and business processes (CIOs and COOs). This approach has a lot of potential.
- Sensitivity analysis both within and between models is important. The example of enforcement versus education was again discussed in terms of sensitivity analysis.

Viability/Transitionability

- We need to find out what the big steps are to make something like this work. The huge spread of scales is really interesting. Productionizing the ability to do this in different domains is important
- Visualization is huge for going to upper leadership. You need to show pictorially that policy recommendations make sense.

DASD(SE) REVIEW

The model was then presented on February 1, 2016, to representatives from the Office of the Deputy Assistant Secretary of Defense for Systems Engineering. A format similar to the two previous reviews was used. The discussion is summarized below.

- Senator Tester had a study done that would be beneficial to the model. DLA would be the source for information on that study.
- DPAP wants to use qualified suppliers, but it is not clear what the exact definition of a qualified supplier is. Industry is driving the definition of qualified suppliers. The problems are several tiers removed from prime contractors and at least one tier removed from DLA and depots.
- What are the major influencers in the model?
- The European Union and Asia also export their electronic waste to non OECD countries. That should be addressed in the model.
- GIDEP has a semi-automated response system. The enforcement of the GIDEP monitoring and reporting system would be of interest. What is the influence on diminishing suppliers and qualification?
- It would be of interest to see “under the hood” of the model to see how the major model elements interact with one another.
- There are any number of “silver bullets” that are proposed to address counterfeit parts. These may work, but may then cause unintended effects. Using the model to study those effects would be of interest.
- The limitations on data are an issue. Due to this, it would not be expected to use the model for exact numerical results, but rather to provide useful insights for policy analysis.
- The Office of the Assistant Secretary of Defense for Logistics & Materiel Readiness (L&MR) may be a customer for the model.
- The model needs more direct, detailed and sustained involvement with a community to be viable.
- What level of resources would be required to make this model operational?
- It would be a good idea to find discrete aspects within the model that have data sources and flesh them out so that the model can be tuned. The operation of GIDEP may be one example, where the effects of monitoring and reporting on counterfeit suspects, supplier diminishment and supplier qualification can be studied.

Anti-Counterfeiting Roundtable Review/Workshop

As another task in this research, the model was presented on February 5, 2016, to the group of subject matter experts who comprised the anti-counterfeiting roundtable that provided input into the model design and development. The review also focused along the lines of validity, acceptability, and viability. The discussion is summarized below.

Validity

- The Department of Defense has collected data on quality control over the years. These data might provide a reasonable proxy set. This is a follow-up item for future model tasks.
- There are direct and indirect policy costs. Policies can cause a large amount of "extra work" that is not reflected in the direct cost. How should that be captured? One example is the extra work caused by false positives. This extra work occurs both in the test process and also in the system-wide response to GIDEP alerts from false positives.
- Some parts are not testable, so a feature should be added that would allow you to make certain parts un-testable.
- The U.S. is not the only exporter of electronic waste, so electronic waste from Asia and Europe should also be considered in the model.
- It would be of interest to have a product choice lever in the simulation, where programs can leverage more non-DoD unique parts. For instance, a program may want to align select parts with the automotive industry to ensure longer term supply and reduce the obsolescence issue.

Viability/Transitionability

- One participant indicated that the model and approach show promise, but potential users need a better understanding of the dynamics in the model, a deeper understanding of the business rules in the model, and also better data to get a sense of validity of the model. Overall, it could be useful.
- One potential conclusion from this research is that we need to identify efficient ways to collect data in the future the address the data issue in complex enterprise problems.

Transition Plan Discussion

While the model has been developed primarily as a demonstration of the overall enterprise modeling methodology described in this report, it could potentially be of use in policy analysis. The various reviews have provided a rich set of feedback for the development of the following transition plan proposal for the counterfeit parts enterprise model, and there has been generally positive reception to its usefulness. We present the following points as the beginning of discussion and proposal for transitioning the model to use.

Owners and stakeholders. DoD agencies would seem to be the logical home for the counterfeit parts enterprise model. The three main agencies would be DASD(Systems Engineering), DASD(Logistics & Materiel Readiness) and Defense Procurement and Acquisition Policy (DPAP). One arrangement would be for one organization to be lead, while coordinating model use and focus with the others. This configuration most likely would have the model have a particular focus, such as acquisition, logistics or policy.

Other owners and stakeholders are possible, such as electronics industry consortia, prime contractors or other government agencies that address counterfeiting. However, the model may need substantial revision to meet the particular needs of those organizations.

One of the key takeaways from the various reviews is that the owner and main stakeholders need a deep understanding of the model's assumptions and underlying dynamics and interactions.

Model focus. The model should be focused on a particular concrete set of phenomena for which data is available. For instance, one comment from the Anti-Counterfeiting Roundtable review mentioned monitoring and reporting via the GIDEP database. This phenomena set likely has available data and could be explored in more detail. There are other phenomena sets such as the flow-down process from lead systems integrators to their suppliers of various DoD policies such as anti-counterfeiting and supplier qualification. Clearly, the focus must be on something of interest to the owners and main stakeholders, and there must be enough relevant data. This is in line with the notion of the "core model" approach to enterprise modeling; however, the intent is to focus on a smaller set of phenomena than, for example, the delivery operations level of a multi-level enterprise model.

Model refinement and enhancement. While the interaction with stakeholders during model development was useful, more detailed interaction with owner-stakeholders is needed to elaborate and refine the model. The reviews have provided numerous examples of potential refinements, ranging from modeling GIDEP monitoring and reporting to modeling the effect of electronic waste exported to countries with counterfeiting recyclers by Asian and European nations. In addition, there may be other phenomena that would be relevant to include, such as new policies (e.g., penalties, program notifications to DoJ on counterfeit suspects, test selection for different types of inspections to be done on incoming parts), behaviors (e.g., supply chain adaptation by counterfeiters in response to government actions), and alerts/indicators for phenomena such as supplier diminishment.

Model validation. There is a strong need to validate the model with the owner-stakeholders beyond what has initially been accomplished with the various reviews. This would likely occur across four lines.

- Validation of detailed model behavior using available data.
- Validation of aggregate model behavior under a variety of different scenarios via comparison to subject matter expert expectations.
- Detailed validation of individual model components and agents.
- Identification and investigation of any counter-intuitive results to determine if they may be real effects of model error artifacts.

3. Behavioral Economics Case Study

The use of behavioral models such as prospect theory may be helpful in describing enterprise systems and perhaps in identifying the bifurcation points that exist in these systems. In particular, those bifurcations that stem from human decision making, as opposed to physical or mechanical bifurcations, may be better identified by behavioral modeling. In this section we discuss prospect theory and investigate its use in this problem. As an example, we consider the problem of modeling driver response to congestion pricing. We conclude that prospect theory can be used to accurately model this situation, and also discuss how a simpler, utility-based, model could be constructed so as to capture the same phenomenon. In light of this, we argue that the true benefit of the behavioral approach is that it lends itself more naturally to the identification of the bifurcations that occur within this problem.

Behavioral Economics and Prospect Theory

Behavioral Economics describes a broad set of economic modeling tools that attempt to describe economic phenomena from the perspective of those engaged in it, as opposed to the perspective of the perfectly rational agent that is commonly assumed in the classical economic literature. It serves as an explanatory methodology rather than a prescriptive one in the sense that it seeks to model the actual behavior of people, and not the behavior that people should engage in to perfectly respond to a given set of circumstances. For a synopsis of behavioral economics see Camerer and Loewenstein (2004).

The inclusion of a behavioral element seems necessary for any descriptively accurate model of a socially-based phenomenon. As hard as they may try, humans do not act in the purely rational way assumed by many models. From a computing perspective, humans are simply incapable of the instantaneous processing and calculation needed to determine the optimal action for a given situation, and even if they had this capacity, there are sometimes physical barriers to implementing a perfectly rational strategy. For example, a driver may be prevented from choosing the optimal route simply because an exit is unreachable due to heavy traffic or speed.

In their seminal works, Kahneman and Tversky (1979, 1992) show cases where expected utility is faulty and introduce Prospect Theory as a way of surmounting these faults. It is characterized by risk aversion with respect to high probability gains, risk seeking with respect to high probability losses, and for low probability events, risk seeking with respect to gains and risk aversion for losses. The theory defines a two stage decision process. The first stage, editing, orders the different possible outcomes and selects a reference point which defines gains and losses. Evaluation, the second stage, provides a value and probability for each outcome, and additionally, a probability weighting function. Combining these values, one can determine the expected prospect of a choice. Cumulative prospect theory is much the same, with the exception that the probability weighting function is applied to the cumulative probability distribution instead of to the probability of individual events. As an example of prospect theory being put to practical use, Rasiel, Weinfurt, and Schulman (2005) describe an example in which medical decision making departs from expected utility theory, and show how the situation might be modeled by prospect theory. Barberis (2013) discusses the history of prospect theory and the difficulties associated with its application.

More technically, prospect theory utilizes a concave value function for gains and a convex function for losses, where the loss function is steeper than the gain function. Kahneman and Tversky suggest the following value function:

Equation 1

$$v(x) = \begin{cases} (x - x_0)^{\alpha} \text{ if } x > x_0 \\ -\lambda(-x + x_0)^{\beta} \text{ otherwise.} \end{cases}$$

Here x represents a particular outcome and x_0 represents the reference point. The probability weighting function overweights small probabilities and underweights high probability events.

The authors also name five phenomena that occur regularly in the real world, but are not accounted for by expected utility theory: framing effects, nonlinear preferences, source dependence, risk seeking behaviors, and loss averse behaviors. All of these phenomena are described by prospect theory.

Prospect theory can thus be viewed as a response to the failures of utility, i.e., it arose as a method of correcting the predictions that one would make based a utility model. Put in another way, there are situations where a bifurcation separating reality and utility predictions can arise. Prospect theory can bridge this gap providing predictions that better match reality. The difference between utility theory and prospect theory is thus the ability to accurately model these bifurcations.

Application to Enterprise Systems

Bifurcation Modeling

In Pennock and Gaffney (2016), the authors introduce the idea that certain types of model inadequacy are caused by bifurcations or phase transitions. These can occur either within reality or between reality and the model used for its description, and usually involve a qualitative shift in some phenomena. The existence or possibility of these bifurcations adds an extra element to be captured by a model and thus increases the likelihood of model error. Although some bifurcations are physical in nature, and can thus be modeled based on known facts (e.g., the various phases of water), others arise because of social factors, and as such, are much more difficult to model.

We argue that a behavioral approach to modeling may yield insights with respect to the existence of phase transitions and bifurcations. Due to the shape of the value function used in prospect theory, bifurcations are actually quite natural to model. Indeed, a basic feature of the value function is that gains and losses are evaluated by separate functions. Hence, the point where we switch from one function to the other is the reference point at which the bifurcation occurs. Prospect theory is thus a natural choice for bifurcation modeling.

Congestion Pricing

As an example of a phenomena containing bifurcations, we consider the problem of congestion pricing. Highway tolling has a long history, with tolled roads existing throughout the world. Dynamic congestion pricing is a typical strategy used to reduce demand on roads. The typical equilibrium pricing policy is one that imposes upon the driver the marginal cost to society of their trip. As Xu et al. (2011) explain

The premise is that being charged the marginal external costs their trips impose to the society, users will voluntarily change their travel behaviors in such a way that traffic congestion is minimized or social welfare is maximized. As the marginal external costs vary over time, space or vehicle type, "theoretically-optimal" tolls will be highly differentiated and fully dynamic.

From the perspective of those who are actually in charge of road management and toll implementation, the main assumption likely involves an assessment of road demand under different toll prices. It would be quite natural to write demand as a decreasing function of toll price, for example,

Equation 2

$$D(p) = cp^{-z},$$

where p is the toll price, $D(p)$ is the demand for the tolled road at price p, and c and z are constant parameters. Such an assessment could be based on stated preference surveys or perhaps even the basic intuition that higher prices lead to lower demand. We contend that, if the traffic manager gives in to the temptation to immediately accept such a model and use its results regarding toll road demand as an input to some larger scale model, serious problems are likely to emerge.

The source of the problem is that it is unlikely that drivers will respond "correctly" to the disincentive of the toll. Recent studies, for example, have shown evidence that driver response to tolled lanes is not always accurately described by classical economics, i.e., higher tolls do not always result in lower demand. In particular, there is evidence from both Minneapolis (Janson and Levinson, 2014) and San Diego (Brownstone et al., 2003) that show cases where the demand for the tolled lanes actually increases when the toll price is increased. A purely classical model such as (Equation 2), with demand decreasing as price increases, is thus inadequate for capturing the full complexity of the situation.

The missing factor is the benefit felt by the driver with respect to time savings. If a change in the toll was not indicative of a change in traffic volume, then the effect of the toll would be the straightforward loss of utility due to paying a higher toll for no benefit. However, if a change in toll is accompanied by (at least in the mind of the driver) an increase in traffic in the non-tolled lanes, then the change in utility is not so clear; it could increase or decrease depending on the results of a comparison between utility loss due to toll and utility gain due to time savings.

PROSPECT THEORY MODEL

Traditional models of driver response to congestion pricing center around utility maximizing individuals, although there is some recent work that takes a behavioral approach to the problem. In a sense, the behavioral approach to congestion pricing modeling is the most obvious and natural way to go, as we are trying to model the response of drivers to a toll in reality, and not their behavior in an idealized world. Furthermore, the choice of a driver is likely a split second decision for the infrequent toll road driver, or an ingrained and nearly automatic response for the everyday driver. The decision of whether or not to utilize a toll road may therefore be divorced from optimality considerations and have more to do with these practical concerns. There are several papers in the literature that apply prospect theory to congestion pricing, for example, Pan and Zuo (2014) and Xu et al. (2011). The goal of such models is generally to formulate an equilibrium price, and not to match the unexpected demand behavior that occurred in Minneapolis and San Diego.

Our purpose here is to show how the unexpected driver behavior might be anticipated or noticed *a priori*, i.e., by looking at a model of the situation instead of through experimental analysis. While we contend that prospect theory is well suited to find the particular type of bifurcation involved in this situation, other behavioral models may be of use in finding other types of bifurcations. For instance, paradigms such as information economics for signaling or game theory for non-cooperative behavior might be used to identify alternative types of bifurcations.

In this section we discuss how a prospect theory based model of toll road demand could be formulated. We assume that driver utility is derived from a comparison of their expectations regarding trip length and cost and the actual trip length and cost incurred by the driver. This conforms to the idea that drivers view their driving experience relative to their own particular frame of reference. To a certain extent, these frames of reference are likely similar from person to person. For instance, a driver crossing the George Washington Bridge is likely to expect a more costly and slower trip than a trip of the same length in a rural region. The prospect theory approach also agrees with the thought that small losses are viewed more negatively than a correspondingly small gain is viewed favorably. Put in another way, the good we feel by being slightly early is far outweighed by the bad that we feel by being slightly late.

The driver must choose between an unreliable (with respect to trip length) non-tolled lane and a reliable (i.e., constant trip length distribution) toll lane. We can imagine a sort of indifference curve spanning different toll/time pairs which serves as a boundary between preferring the tolled and non-tolled lanes. We acknowledge that many drivers have their minds made up regarding toll road utilization well before the choice is given, but assume that there is a population of drivers who can be swayed to one road or the other based on the present conditions.

We also assume a range of possible outcomes regarding trip length in both the tolled and non-tolled lanes, and an associated probability distribution describing the likelihood of these outcomes. When the toll is viewed, the values of trip length are updated for the non-tolled lanes (increasing with toll values), while it is assumed that trip length ranges in the tolled lanes do not degrade further.

To develop the model, let t = the toll value, where $t \geq 0$, M_t = expected trip length in the non-tolled lanes when the toll price is t, and M_X = the reference trip length for driver X. M is determined by the driver, while we assume that M_t is linearly dependent on and increasing with t, We define the utility (or prospect) for driver X paying toll t as

Equation 3

$$U(X_t) = p(t) + f(M_t - M_X),$$

where $p(t)$ is a penalty function for paying a toll of t dollars and $f(\cdot)$ is the value function of prospect theory.

We assume here that the penalty function is of the form $p(t) = \mu \cdot g(t)^{\rho}$, where $g(t)$ is a function converting a monetary value t to a time value, and $\mu < 0$ and ρ are given parameters. Adapting (Equation 1), we use the value function

Equation 4

$$f(M_t - M_X) = \begin{cases} (M_t - M_X)^{\alpha} \text{ if } M_t > M_X \\ -\lambda (M_t - M_X)^{\beta} \text{ otherwise,} \end{cases}$$

where $\lambda > 1$.

We also require that $g(0) = 0$. Then, $U(X_0)$ gives the utility of the free lanes, which can be compared with $U(X_t)$ for any driver X and toll $t > 0$. If $U(X_t) > U(X_0)$, then the tolled lanes are preferred, while the free lanes are preferred if the reverse inequality holds.

With proper parameterization, this model can be used to show how the demand for toll roads can shift in the way that has been found in Minnesota and San Diego. Let $M = 25$ minutes, $\mu = 1.1$, $\rho = 0.25$, $\alpha = \beta = 0.25$, and $\lambda = 1.5$. To slightly generalize the situation, we assume that M_0 can take one of five values, 23, 24, 25, 26, or 27, with respective probabilities of 0.05, 0.2, 0.5, 0.2, and 0.05. We assume that each of these five values increases at a rate of $0.75 \cdot t$, so, for example, if the toll is \$1, the five possible values of M_1 are 23.75, 24.75, 25.75, 26.75, or 27.75. We assume that the associated probabilities are the same as above. The time/dollar conversion function is assumed to be $g(t) = 2t$.

Figure 10 shows the difference in utility between tolled and free lanes. Notably, we see that the utility of the free lanes is higher than the tolled lanes for tolls below \$6 (approximately). The free lanes would thus be preferred to the tolled lanes until the toll reaches \$6. At this point, driver X would switch from free to tolled lanes, exhibiting the type of unexpected behavior noted above.

Figure 10 – Difference in utility between the tolled and free lanes

OBJECTIONS

We have shown how prospect theory can be used to model the decision making of drivers faced with congestion tolling. The question is if this strategy is better than, or preferable to, other reasonable models such as utility theory. To this end, we describe how utility theory could also be used to give a descriptively accurate model, and we conclude with a discussion of why, if both models can adequately capture reality, the behavioral model may be preferred.

As we have discussed, paying a toll will decrease one's utility, *ceteris paribus*. However, paying the toll will result in a benefit to the driver, as a time savings will occur by taking the toll road as opposed to the non-tolled road. We can imagine a utility function depending on two parameters, toll price t and trip time m. If m is fixed, we would have $\frac{\partial}{\partial t}u(t,m) < 0$, and if t is fixed we have $\frac{\partial}{\partial m}u(t,m) < 0$. However, we contend that t and m are not independent, at least in the behavioral reaction of the driver to new information (the toll price). Namely, it seems reasonable to assume that, as t increases m should decrease. Thus, there will be a gain in one aspect of the utility function and a loss in the other aspect. The magnitude of these derivatives will therefore govern the favorability of the toll or non-tolled road for a given toll value.

The entire question regarding the suitability of a utility model could thus be restated as a question of how the developers could have anticipated that the monetary utility function and the time utility function had derivatives that would sometimes lead to the unexpected behavior. While functions can certainly be written that will generate the desired outcome (the unexpected behavior), we see no reason why utility functions with these properties would be chosen *a priori*.

Behavioral models, on the other hand, are explicitly meant to capture situations where a clear change in behavior takes place. The main idea is that what is needed to `forecast' a bifurcation is exactly what prospect theory adds to utility theory. Even if there is a simpler representation of a phenomenon (i.e., a simple utility model), that representation may not be useful with respect to bifurcation identification, whereas prospect models are ideal in representing bifurcations based on reference points which differentiate gains from losses and risk aversion from risk seeking.

Finally, many of the basic benefits of prospect theory (e.g., risk aversion and risk seeking under different circumstances, framing effects, and losses being viewed more harshly than gains) are all clearly present in traffic situations, and so we would expect that a modeling strategy explicitly focused on such issues would be superior to one with no such focus. Although utility theory could be used to model such things once we know that they are present, behavioral models are useful for showing the *possibility* of such events, and so are useful for the traffic engineer who must, without empirical data, determine the possible response of drivers to fluctuating tolls.

CONCLUSIONS

While the original intent of this task to increase the fidelity of agent decision models in enterprise models, the results took us in a different direction. Rather we concluded that the issue was no so much the fidelity of the model as the ability to detect bifurcations in system behavior. In enterprise systems, the uncertainty is often so high that only substantial shifts in model behavior are meaningful. Unless the increase in model fidelity results in such a detection, any other improvements in accuracy are likely to be lost in the noise. This result has strongly influenced the whole enterprise systems analysis effort. The implications are discussed in greater detail in Sections 6 and 8.

4. PHENOMENA AND CANONICAL MODELS

Modeling enterprise systems necessarily requires the simultaneous consideration of the system from multiple perspectives. Given the nature of enterprise systems this often requires models from different scientific disciplines that were not intended to be integrated. Previous tasks (RT-44, RT-110) considered some of the challenges of composing such models. It is has been done successfully in some instances, but often times it proves difficult. Thus, the question is what allows one to reuse and compose models from different disciplines successfully. This section is divided into two parts. First, we discuss the problem of phenomena and reuse and a conceptual level and draw conclusions based on past approaches to model composition and reuse. Second, we develop a mathematical approach to analyze the conceptual view and assess the resulting necessary conditions for composition and reuse. Finally, we consider which research directions show the most promise to facilitate enterprise modeling.

A Conceptual view of Model composition and reuse for enterprise modeling

When we attempt to create models of complex systems and enterprises, we often intend to reuse existing models as components of the overall formulation. For example, an overall model of healthcare delivery (e.g., chronic disease management) would need component models for the incidence and progression of hypertension, type 2 diabetes, and heart disease. In this case, the acceptability of the overall model would depend on the component models having been vetted and published in reputable medical journals.

Such component models might be in the form of parameterized equations, representations encoded in commercial software tools, or legacy software code developed for previous purposes. Integrating such component models into an overall model presents several challenges at syntactic, sematic, and pragmatic levels (Tolk, 2003, 2013; Pennock & Rouse, 2014). These challenges range from assuring compatible variable definitions, units of measure and coordinate systems, to being consistent about assumption concerning independence, conservation, and continuity.

This section addresses such reuse in the context of an overall framework for creating and assembling models. This framework starts with the nature of the problem of interest and proceeds as follows:

- The nature of the problem of interest and the questions associated with this problem should drive the development of models and simulations.
- The nature of the problem and questions strongly influence the extent to which reuse of component models and simulations can be justified.
- Problems can be defined in terms of the phenomena (e.g., physical, human, economic, and social) associated with addressing the questions of interest.
- The variable predictions needed to address the questions should inform the choice of modeling paradigms and representations.
- Paradigms and representations have associated typical assumptions that may or may not be warranted and should be consistent across component models.
- Reuse can occur at five levels: paradigms, representations, standard problems, solution software packages, or legacy software code; risk typically increases with level of reuse.

Archetypal Problems

Rouse (2015) discusses six archetypal problems that provide test cases for the overall modeling and visualization methodology presented in the book.

- Deterring or Identifying Counterfeit Parts
- Financial Systems and Bursting Bubbles
- Human Responses and Urban Resilience
- Traffic Control via Congestion Pricing

- Impacts of Investments in Healthcare Delivery
- Human Biology and Cancer

Table 2 characterizes these six problems in terms of the historical narrative whereby these problems emerged, the overall ecosystem characteristics, the organizations and processes involved, and the people or other basic elements of the system.

These six problems have several common characteristics:

- All involve behavioral and social phenomena, directly or indirectly
- All involve effects of human variability, both random and systemic
- All involve economics (pricing) or financial consequences
- All include both designed (engineered) and emergent aspects

There are also important distinctions:

- Counterfeit Parts and Financial System involve deception by a subset of the actors
- Healthcare Delivery and Human Biology involve aberrant functioning by a subset of the actors
- Congestion Pricing and Urban Resilience involve aggregate consequences (e.g., traffic) of all actors

Another important distinction is between two classes of problems:

- Bottom-Up: Detection and remediation of aberrant actors involves stratifying actors and exploring behaviors of each stratum in different ways
    - Aberrant actors tend to react to remediation strategies, eventually undermining their effectiveness
- Top-Down: Economic strategies, e.g., pricing, payment models, procurement practices, based on aggregate behaviors
    - Individual actors tend to react to aggregate strategies, often undermining the desired consequences

Table 2 - Characterizations of Archetypal Problems

Levels of Phenomena	Counterfeit Parts	Financial System	Urban Resilience	Congestion Pricing	Healthcare Delivery	Human Biology
Historical Narrative	Evolution of aerospace / defense ecosystem in terms of decision processes and incentives	Evolution of financial ecosystem in terms of investment instruments, regulations, etc.	Evolution of urban ecosystem in terms of social development, communities and neighborhoods	Evolution of transportation ecosystem in terms of technologies, demographics & expectations	Evolution of healthcare ecosystem in terms of ends supported and means provided	Evolution of humans in terms of genes, proteins, cells, tissues, organs, systems and signaling
Ecosystem Characteristics	Aerospace / Defense ecosystem – norms, policies, values and supplier economics	Financial ecosystem – what is assumed, allowed, illegal, and enforced	Urban ecosystem – norms, values and elements of social resilience	Transportation ecosystem – norms, values & expectations of convenience	Healthcare ecosystem – norms, values and resource competition	Human biological ecosystem, including factors such as lifestyle and environment
Organizations & Processes	System assembly and deployment networks and controls; test and evaluation	Commercial and investment banks, mortgage companies, and regulatory agencies	Urban infrastructure networks and flows -- water, food, energy, and people	Transportation infrastructure networks and flows, and control systems	Provider, payer and supplier organizations – investments, capacities, flows, outcomes	Cardiovascular, pulmonary, digestive, nervous, reproductive, et al. systems
People or Basic Elements	Flow of parts in supply chain to assembly and deployment	Investors, financial engineers, traders, and homeowners	Peoples' evolving perceptions, expectations and decisions, as well as shared beliefs	Individual vehicles and driver decision making in response to flows and controls	People's health and disease incidence, progression and treatment	Cellular processes and signaling mechanisms; therapy decisions

Considering how the phenomena associated with these problems might be represented, three common features should be noted. First, the set of phenomena associated with a problem can be represented at different levels of abstraction, e.g., individual instances of counterfeiting versus macroeconomic policies that motivate counterfeiting. Second, each problem has phenomena of interest that emerge within each layer of abstraction. This would suggest that a different representation of the complex system would be relevant for each layer. Third, each problem exhibits feedback loops that cut across two or more layers. For example, the incentive to counterfeit increases with declining supplier profit margins. High-level policies designed to combat counterfeiting could raise costs at the lower levels. This could further erode profit margins and actually increase the incentive to counterfeit. Thus, the counterfeiting problem cannot be addressed without considering the relationships between the different layers of the complex system. The phenomena associated with these six problems are later discussed.

MODELING PARADIGMS

A scientific paradigm is a distinct set of theories, research methods, postulates, and standards defining legitimate research. Examples include Newtonian mechanics, Einsteinian relativity, and quantum mechanics. In contrast, a modeling paradigm is a class of formalisms, typically mathematical or computational, for representing a phenomenon of interest. Examples include control theory, queuing theory, and network theory.

Variable Predictions of Interest

The choice of a modeling paradigm depends on the variables one needs to predict, as well as the assumptions one is willing to accept. Listed below is a range of variable that one might need to predict.

- Response magnitude
- Response time
- Stability of response
- Control errors
- Observability
- Controllability
- State estimates
- Estimation errors
- Number and time in queue
- Number and time in system
- Probability of balk or renege
- Shortest distance
- Shortest time
- Propagation of sentiment
- Choice selected

- Game equilibrium
- Election results
- Impacts of incentives
- Time until problem solved
- Steps until problem solved
- Problem solving errors
- Net present value
- Net option value
- Net capital at risk

Predictions, Paradigms, Representations, and Assumptions

Table 3 maps these predictions of interest to paradigms, representations, and assumptions. Eight different modeling paradigms are summarized in terms of the predictions they typically provide, the most common representations, and assumptions usually associated with each paradigm.

There are, of course, many more types of variables that might be of interest, as well as quite few other modeling paradigms that might be employed. Thus, Table 2 is meant to be representative rather than exhaustive. In particular, it suggests a line of reasoning to map from problems and questions to candidate component models for inclusion is an overall model of a complex system or enterprise.

It is also useful to note that there are domain-specific versions of most of these modeling paradigms. For example, there are a variety of models of disease incidence and progression available in the healthcare delivery literature. As another example, there are several traffic congestion models available in the transportation literature.

Table 3 - Predictions, Paradigms, Representations and Assumption

Predictions of Interest	Modeling Paradigm	Representation	Typical Assumptions
• Response magnitude • Response time • Stability of response	Dynamic Systems Theory	• Differential or difference equations	• Newton's Laws • Conservation of mass • Continuity of transport
• Response time • Stability of response • Control errors • Observability • Controllability	Control Theory	• Differential or difference equations • Stochastic processes, Markov processes	• Known transfer function or state transition matrix • Stationary, Gaussian stochastic processes • Given objective function of errors, control effort
• State estimates – filtering, smoothing, prediction • Estimation errors	Estimation Theory	• Differential or difference equations • Stochastic processes, Markov processes	• Known dynamics of process • Known ergodic (stationary) stochastic process • Additive noise inputs
• Number and time in queue • Number and time in system • Probability of balk or renege	Queuing Theory	• Differential or difference equations • Stochastic processes, Markov processes	• Known arrival and service processes • Future state only depends on current state • Given service protocol, e.g., First Come, First Served, priority
• Shortest distance between any two locations (nodes) • Shortest time between any two locations (nodes) • Propagation of sentiment among actors	Network Theory	• Graph models • Agents as nodes • Relationship arcs	• Discrete entities, e.g., agents • Decision rules of entities, e.g., agents • Typically binary relationships • Relationships only via arcs or edges
• Choice selected • Game equilibrium • Election results • Impacts of incentives	Decision Theory	• Utility functions • Payoff matrix • Social choice rules	• Known utility functions • Comparable utility metrics • Known payoff matrix • Given voting rules
• Time until problem solved • Steps until problem solved • Problem solving errors	Problem Solving Theory	• Neural nets • Pattern recognition • Production rules	• Known human mental model • Known information utilization • Known repertoire of patterns • Known troubleshooting rules
• Net present value • Net option value • Net capital at risk	Finance Theory	• Time series • Stochastic processes	• Projected investments • Projected operating costs • Projected revenues and costs

Phenomena And Paradigms

Table 4 summarizes eights classes of phenomena employed to characterize the six archetypal problems at deeper levels. These characterizations can be used to map the phenomena associated with each problem to modeling paradigms likely of use for addressing the problem.

Table 4 - Eight Classes of Phenomena

Class of Phenomena	Example Phenomena of Interest
Physical, natural	Temporal and spatial relationships & responses
Physical, designed	Input-output relationships, responses, stability
Human, individuals	Task behaviors & performance, mental models
Human, teams or groups	Team and group behavior & performance
Economic, micro	Consumer value, pricing, production economics
Economic, macro	Gross production, employment, inflation, taxation
Social, organizational	Structures, roles, information, resources
Social, societal	Castes, constituencies, coalitions, negotiations

Table 5 maps selected phenomena identified for the six archetypal problems to modeling paradigms. The full list of phenomena associated with these six problems is discussed in Rouse (2015).

Table 5 - Archetypal Phenomena and Modeling Paradigms

Category	Phenomenon	Modeling Paradigm
Physical, Natural	Flow of Water	Dynamic Systems Theory
Physical, Natural	Disease Incidence/Progression	Statistical Models, Markov Processes
Physical, Natural	Cell Growth & Death	Network Theory, Biochemistry
Physical, Natural	Biological Signaling	Network Theory, Biochemistry
Physical, Designed	Flow of Parts	Network Theory, Queuing Theory

Physical, Designed	Assembly of Parts	Network Theory, Queuing Theory
Physical, Designed	Flow of Demands	Network Theory, Queuing Theory
Physical, Designed	Traffic Congestion	Network Theory, Dynamic Sys. Theory
Physical, Designed	Vehicle Flow	Agent-Based Models
Physical, Designed	Infrastructure Response	Dynamic Sys. Theory, Network Theory
Human, Individual	Diagnosis Decisions	Pattern Recognition, Problem Solving
Human, Individual	Selection Decisions	Decision Theory
Human, Individual	Control Performance	Dynamic Systems Theory, Control Theory
Human, Individual	Perceptions & Expectations	Pattern Recognition, Bayes Theory
Human, Team/Group	Group Decision making	Decision Theory, Social Choice Theory
Economic, Micro	Investment Decision Making	Decision Theory, Discounted Cash Flow
Economic, Micro	Operational Decision Making	Network Theory, Optimization
Economic, Micro	Risk Management	Decision Theory, Bayes Theory
Economic, Micro	Dynamics of Competition	Game Theory, Differential Equations
Economic, Macro	Dynamics of Demand & Supply	Dynamic Systems Theory, Optimization
Economic, Macro	Prices, Costs & Payment	Discounted Cash Flow, Optimization
Social, Info. Sharing	Social networks	Network Theory, Agent-Based Models
Social, Organizations	Domain social system	Network Theory, Decision Theory
Social, Values/Norms	Domain values & norms	Network Theory, Decision Theory

PARADIGMS AND STANDARD PROBLEMS

Table 6 summarizes representative "standard problems" associated with the modeling paradigms. If the problem of interest can be represented as a standard problem, then one can potentially adopt the standard solution to this problem. A central concern in deciding to do this is the extent to which the standard assumptions associated with one of these solutions is acceptable.

Table 6 - Modeling Paradigms and Standard Problems

Modeling Paradigm	Standard Problems
Dynamic System Theory	Response Time & Stability
Control Theory	LQG Optimal Control
Estimation Theory	Kalman Filtering
Queuing Theory	(G/G/c):(FCFS/N/∞)
Network Theory	Spanning Tree, Shortest Path
Decision Theory	ROC Models, MAUT
Problem Solving Theory	Recognition Primed DM
Finance Theory	DCF, CAPM, Real Options

Table 7 lists representative software packages, equation solvers, and codes employed to compute solutions to the standard problems in Table 5. The commercial software packages and equation solvers are usually quite well supported and often have active user groups that can help with questions. A primary difficulty, however, is the extent to which the overall model, of which the standard problem is just one component, can be embodied in the package or solver that provides the standard solution of interest.

This difficulty is lessened if one can access legacy code, perhaps in C++, Java, or Python, that computes the solution to the standard problem of interest. Unfortunately, most legacy code is poorly documented and supported, if at all. As a consequence, one may be able to assure syntactic compatibility, but not semantic and pragmatic compatibility. In other words, you can get the pieces to execute together, but you really have no assurance that the variable flow is consistent and the computed results are meaningful.

Reuse Of Solutions

This leads to the overall issue of reusing solutions. The foregoing emphasized the use and reuse of computational and/or solution methods, e.g., for system dynamics, discrete event, or agent-based representations. The formalisms, computational methods, and typical visualizations are fairly well understood for this type of reuse, although not without challenges.

Table 7 - Problems and Solutions

Problems	Packages	Solvers	Code
Response Time & Stability	Stella, Vensim	Matlab, Mathematica	C++, Java, Python
LQG Optimal Control	DIDO, GROPS-II	Matlab, Mathematica	C++, Java, Python
Kalman Filtering	R	Matlab, Mathematica	C++, Java, Python
(G/G/c):(FCFS/N/∞)	Simio, Arena, AnyLogic		C++, Java, Python
Spanning Tree, Shortest Path	GraphTea		C++, Java, Python
ROC Models, MAUT	DMS, GRIP	Excel	C++, Java, Python
Recognition Primed DM			C++, Java, Python
DCF, CAPM, Real Options	Real Options Valuation	Excel	C++, Java, Python

Another type of reuse addresses domain/problem representations. Good examples are military operations, supply chains, and highway traffic. Some of these types of simulations are available in, for example, AnyLogic and Vensim. People familiar with these software packages can usually readily "reverse engineer" legacy models to gain insights into assumptions and representations.

Many of these types of simulations are legacy software codes that were developed years ago for particular purposes. The original developers are often long gone. Documentation is often lacking because reuse was not anticipated – or budgeted. For those cases where reuse was anticipated, documentation is usually much better, but the code can tend to be rather opaque.

The approach to model creation and assembly outlined earlier has the following implications for reuse:

- Paradigm: Reusable if paradigm matches phenomena and questions of interest
- Representation: Reusable if typical assumptions are acceptable for questions of interest
- Standard Problem: Reusable if typical assumptions are acceptable for questions of interest
- Software Package: Reusable if representation and assumptions can be instantiated in package
- Legacy Code: Reusable if variables of interest and key assumptions match problem formulation
- Principle: Variables and assumptions should be consistent and compatible across component models

There are also implications of reuse. Typically time and money are saved if inconsistencies and incompatibilities are minimal. Success stories include computational fluid mechanics, semiconductor design, and supply chain management (Rouse, 2015). These examples involve representations and solution approaches that have been developed, used, and refined by active communities of users.

For modeling endeavors that are closer to "one off" projects, assumption management across legacy component models can be quite difficult. When there are many components of legacy code, inconsistencies and incompatibilities can be difficult to identify and manage. Such problems can lead to "model-induced design errors."

Rouse (2012) reports on an interview study of eight companies; two each in the automobile, building systems, commercial aviation, and semiconductor industries. All interviewees were keenly aware of the issues associated with model reuse. They reported great caution with regard to model reuse and the risks of model-induced design errors.

Their caution was also influenced by their intent to manufacture hundreds, thousands, or millions of units of their systems, with the stark understanding that their companies would be responsible for the consequences of any model-induced design errors. They were unanimous in the importance of learning from past modeling efforts, but avoided any off-the-shelf adoption of past models.

Mathematical Analysis of Phenomena, Reuse, and Composition

In light of the previous discussion, it should be fairly obvious that every modeling effort involves some degree of reuse. So when practitioners and researchers speak of the challenges of model composition and reuse, what do they mean?

In this section we will employ a mathematical tool known as commutative diagrams to explore the necessary conditions for a valid multi-model under different circumstances. What we will find is that most of the emphasis on reuse and composition is not in the area that is causing the problems. The work in ontology and conceptual interoperability is on the right track, but there are likely fundamental limitations.

BASIC SETUP

Before we consider model reuse, we need to frame the language of the discussion. In order to discuss an arbitrary system of interest to us, we will leverage Rosen's approach to viewing systems as developed in the monograph "Fundamentals of Measurement and Representation of Natural Systems." (Rosen 1978). Rosen's concern was how we measure and model natural systems and its associated implications for physics and biology. His approach provides a natural starting point for a discussion of modeling systems and model reuse.

First, we will briefly describe Rosen's setup. We frame it entirely in terms of set theory as most engineers have a working knowledge of set theory, but it is extensible to category theory. (In fact, Rosen does so in his monograph). Rosen frames the general problem as follows: We have a system with a set of states, S. However, we do not directly interact with the state space, S. Rather we measure observables such as position, temperature, voltage, and so on. These observables are essentially functions that map the state space, S, to another set such as the real numbers. A given set of observables, F generates an equivalence relation, R_F, on S.

Let $R_F = \{(s, s'): s, s' \in S \ \wedge \forall f_i \in F[f_i(s) = f_i(s')]\}$ be the equivalence relation induced on S by the set of observables F. For example all states of S that have the same temperature and pressure would be viewed as equivalent for that set of observables, temperature and pressure.

The quotient set S/R_F is the reduced set of system states that result from the set of observables, F. (i.e., it is a partition of the set S.) For example, if the set of observables are the x and y coordinate in a Cartesian coordinate system, then the state space of the system is reduced to the xy-plane.

It is important to note that the reduced state space of the system, S/R_F, is a consequence of which observables are collected. How we look determines what we see. Of course, we do not measure the observables directly. Rather, we use specially configured systems called meters that are designed to dynamically interact with the system of interest and asymptotically approach a value we take to be the measurement of the observable. An example would be using a thermometer to measure temperature.

What we are generally interested in are changes in the state of a system. We can capture state transitions in S through an automorphism on S. (i.e., a bijective mapping from S onto itself). An example for a finite set of states would be a state transition matrix.

Let T be an automorphism on S. If T is compatible with R_f then T induces an automorphism on the reduced set of states S/R_F. Let us call this automorphism T_F. This is a description of the state transitions for states defined by the set of observables F. If we introduce the composition operator, we can generate a group of automorphisms from T_F that can be used to define trajectories in the state space. For example, if we have a discrete state transition matrix, this would be equivalent to repeatedly applying the state transition matrix to itself to generate a sequence of states that would result for each possible starting state. If we index resulting elements of the group by $t \in \mathbb{Z}$ or $t \in \mathbb{R}$, then we can describe changes in system state versus time. We call this the system's dynamics.

Of course, we do not perceive observables directly. We measure them via meters which are specially configured systems that interact with the system of interest via the observables. The system of interest and the meter induce dynamics on each other. The meter is configured so that its dynamics asymptotically approaches a value (typically a real number) that we use to "measure" the observable.

For example, we do not directly perceive the temperature of a glass of water. We introduce a specially configured system called a thermometer that dynamically interacts with the water. The net result of the dynamics of their interaction is a temperature reading on a real number scale. The temperature reading asymptotically approaches the "temperature" of the water.

If we are interested in measuring changes in the state of the system over time, we obtain a diagram like the one below (Figure 11).

$$\begin{tikzcd} S/R_F \arrow[r, "T_F"] \arrow[d] & S/R_F \arrow[d] \\ M_F & M_F \end{tikzcd}$$

Figure 11 – Measuring a System over Time (Adapted from Rosen 1978)

A few characteristics to note: If, as we assumed above, T_F is a bijection (one to one and onto) then the dynamics is deterministic and reversible. Of course, the dynamics, T, on the system states, S, does not have to be compatible with the reduced state space, S/R_F. For many realistic problems, it will not be. The result is that T_F will split equivalence classes of R_F. This has the interesting result of enabling us to discriminate among more states of S then we could with F alone, but it also makes the system appear stochastic and/or irreversible.

For example, if T_F is onto but not one-to-one, the dynamics is stochastic. If T_F is a function, but not onto, then the dynamics is irreversible. Repeated applications of T_F generate a set of trajectories. If we index these by an integer we have a discrete time view of the system and if we index them by a real number, then we have a continuous time view of the system. Finally, note that the reduced state set could be discrete or continuous or both depending on the spectrum of S in the functions of F.

Now we are interested modeling the system depicted in Figure 11. This means that we need to introduce another set of states, X, with a corresponding set of dynamics (or state transitions), D_X.

$$
\begin{tikzcd}
S/R_F \arrow[r, "T_F"] \arrow[d, "\alpha"'] & S/R_F \arrow[d, "\alpha"] \\
X \arrow[r, "D_X"] & X
\end{tikzcd}
$$

Figure 12 – Necessary condition for $(\boldsymbol{X}, \boldsymbol{D_X})$ to be a model of $(\boldsymbol{S}/\boldsymbol{R_F}, \boldsymbol{T_F})$

If the diagram (Figure 12) commutes (i.e., starting with any given element in the upper left, following all possible paths will lead to the same element in the terminal set on the lower right), (X, D_X) could be viewed as a model for the dynamic behavior of the system under the set of observables F. Note that (X, D_X) could represent a physical analog (e.g., a scale model of an aircraft in a wind tunnel) or a mathematical model (e.g., a computer program that simulates the aerodynamic behavior of an aircraft). If this diagram commutes, repeated applications of D_X for a particular starting state yields the predicted trajectory of the system through the state space.

The reader may note that if the dynamics of the system is stochastic, (i.e., T_F, is not one-to-one), then the diagram will not commute. Of course, if the diagram does not commute, then (X, D_X) is not particularly useful as a model. This is addressed in stochastic models by making the observables exhibiting stochastic behavior probability distributions. That is the observable of interest is converted from a point value on the real number line to a function. This restores of the commutativity of the diagram and makes the model deterministic over this adjusted set of observables. For example, if a weather model predicts temperature, we would want the model to generate the same probability distribution of temperature as is observed in the real weather system of interest. Another example is quantum mechanics, the propagation of the wave function is completely deterministic. It is the specific point measurement that is probabilistic (i.e., collapsing the wave function).

Rosen spends an entire monograph exploring the implications of this setup for physics, biology, and science in general. However, our interest here is far more modest. As we will subsequently see, this setup is also quite useful for considering the nature of model composition and reuse for engineering models and simulations.

Defining a Model

As a term, model has many different uses in many different contexts. Consequently, we must define what we mean by model in the context of this discussion acknowledging that this definition is not universal. In the discussion that follows, we will limit the scope to models that we use for prediction as that is chiefly the motivation behind the model composition efforts in engineering.

In short, prediction is the ability to determine the state of a system of interest under circumstances not experienced (different time, location, context, etc.). One way we could do this is if we knew all possible state transitions for a system of interest. In terms of the setup we developed in the previous section, we would want to know the group of automorphisms that describe the dynamics of the system. If we knew that, we could determine what the next state would be given the current state. Mathematically, we would know the function:

$$S \overset{T}{\rightarrow} S$$

Of course, we have two problems here. First, we don't always know what S is, and we interact with it indirectly via meters. Second, even if we knew what S was, for any non-trivial system, determining all the state mappings is effectively impossible since we will not or cannot experience all possible states $s\epsilon S$. So what are we left with? As discussed in the previous section, we can achieve a reduced description of the state space S through observables. So the next best thing is if we could identify an automorphism (T_F) over the reduced state space for a set of observables, F, that we are interested in.

$$S/R_F \overset{T_F}{\rightarrow} S/R_F$$

So we are trying to infer the dynamics of the reduced set of states of the system by taking successive readings with our meters. One way we might characterize the above diagram generically is as a set of *phenomena*. (e.g., water flowing, heat flowing, objects moving etc.). Note that our most basic meters are our five senses, but over time we have built additional meters to measure "new" phenomena.

Again we are faced with the problem that predicting the future state of a system of interest involves knowing all possible state transitions for the reduced state space, and for any non-trivial system this will be impossible. (As discussed above, we are also making a huge assumption that a compatible T_F exists. This is typically not true. Hence the probabilistic behavior of most "real" systems. However, for many real world systems we can get close enough to be useful.)

One way we could address this problem is if we could find a relationship among the observables that is invariant over the dynamics. This is sometimes referred to as a symmetry or similarity relationship. A symmetry allows us to compress the mapping by dropping redundant relationships. They can be reconstructed from the similarity relationship when needed.

We can use a similarity relationship to construct an equation of state for a system. Thus, we no longer have to "know" everything. If we can establish the equation of state, we can instantiate it with a state of interest and extrapolate the resulting trajectory through the state space.

Symmetry relationships are critical to the practice of science. We often refer to these symmetry relationships as *laws*. E.g., Newton's laws, Ohm's law, Coulomb's law, and the laws of thermodynamics. In fact, we could regard science as search for observables that allow us to construct symmetry relationships that are useful to us. When we cannot explain phenomena that are of interest to us, we are having trouble finding exploitable symmetries. By changing or introducing observables, we can sometimes find symmetry relationships and "explain" phenomena.

By selectively applying these symmetry relationships, we build a new system (physical or mathematical) that can serve as a compressed representation of the target system's behavior. We call this new system a model for the target system. In other words, we use the symmetry relationships to reconstruct the target system's dynamics on demand via the execution of experiments for physical models or computation for mathematical models.

INTERPRETATION

Now in light of our discussion on the intent of modeling, let us revisit our necessary condition for modeling (Figure 12) and add another layer of interpretation. In particular, we will consider how we can fit the discussion of phenomena, paradigms, representations, and solutions. As noted above, we can regard

$$S/R_F \xrightarrow{T_F} S/R_F$$

as a phenomenon that we observe and we would like to model. (e.g., fluid flow, planets orbiting, economic activity, etc.). The selection of observables in the set F is the chosen representation of the system (e.g., flow rates, positions of point masses, prices and wages, etc.). More precisely the selected set of observables constitute of an abstraction of the system.

The question then becomes how do we build a model of a phenomenon given a chosen abstraction of the system? We know we need to exploit symmetries and that these are often captured by scientific laws, but where do they come in here? For that we need one more concept, the concept of linkage relationships.

Let us assume that we have two observables $f(s)$ and $g(s)$. Each generates an equivalence relation, $R_f$and R_g respectively. If we apply both at the same time, we obtain the equivalence relation, R_{fg}. What is the relationship among these three equivalences relations? If every class of R_f intersects every class of R_g, and vice versa, then the observables f and g are completely unlinked. That means that knowing the value of one observable provides no information on the value of the other. In other words, we describe the reduced state space of S/R_{fg} as the Cartesian product of the reduced state spaces generated by f and g. ($S/R_{fg} \rightarrow\ S/R_f \times S/R_g$)

On the other hand, if every class of R_f intersects exactly one class of R_g, and vice versa, then the observables f and g are completely linked. (This means that given any element s that belongs to an equivalence class of R_f, it belongs to exactly one equivalence class of R_g) That means that knowing the value of one observable determines the value of the other. This substantially reduces the possible state space as now $S/R_{fg} \subset\ S/R_f \times S/R_g$. This is a symmetry relationship that allows us to compress our representation and perform prediction. Of course, this concept is generalizable to more than two observables. A real example of this would be Ohm's law ($V = IR$) which assumes a complete linkage among the observables voltage, current, and resistance in an electrical circuit. Knowing values of two of the observables enables us to determine the third.

More generally, two or more observables may be partially linked where knowledge of the value of one observable provides incomplete information on the state of another. Perhaps more, importantly, the strength of a linkage relation among observables may vary over different subsets of the state space, S. This true of most if not all of the scientific laws observed to date. Thus, we must always circumscribe when a given law or symmetry relationship does and does not apply. This in turn has implications for the application and reuse of models, which we will discuss further in subsequent sections.

The introduction of linkage relationships allows us to complete our interpretation of Figure 12. The set, X, is the encoding via the mapping α of a subset of the state space $\prod_{f_i \in F} S/R_{f_i}$. This reduction is achievable because of the identified linkage relationships (or symmetries) among the variables. In the case of a mathematical model, X, captures the equation of state. We should note that any observables in the original set, F, that are completely unlinked with the observables of interest are typically omitted. Mathematically, this is equivalent to replacing these with constant observables. In the case that we also restrict the state space, S, such that it falls entirely within a single equivalence class of each of the unlinked observables, these observables can be viewed as parameters of the model.

In the case that a set of observables and associated assumed linkage relationships are a widely accepted explanation for a set of phenomena, we might view $S/R_F \xrightarrow{\alpha} X$ as an application of a scientific paradigm or theory. When we focus on more specific systems, we can view it as a model of that system or a class of similar systems.

The interpretation of the final component of Figure 12, D_X, depends on whether X is a physical analog of the target system or a mathematical model. Of course, it is intended to reproduce a relevant subset of the dynamics, T, in the form of a state transition, but in the case of the former, we induce some physical analog of the dynamics (e.g., operating a wind tunnel). In the case, of the latter, D_X takes the form of computation. Thus, when X is a mathematical model, D_X is interpreted as the application of a modeling paradigm to X.

For example, let us assume that all of the observables (F) of interest are continuous and the dynamics (T_F) is indexed by a real number (continuous time), then we would use a differential equation based approach to generate the mapping D_X. (Note that differential equation based models with analytic solutions still require some computation, just a lot less). If other on the other hand, the observables take discrete values and the dynamics is continuous time but stochastic, we might employ a discrete event simulation.

The leads us to the connection with the previous discussions of phenomena, representations, and paradigms. The phenomena of interest lead us to a subset of scientific or representational paradigms. The questions of interest allow us to further restrict this subset by reducing the set of observables to only those that share linkage relationships with those that are necessary to answer the questions of interest. Based on characteristics of the resulting set of observables and associated exhibited dynamics, we are led to certain modeling paradigms that fit with those characteristics. Of course, this a fairly idealized view, as the real process is more iterative. For example, one may discretize continuous observables and dynamics to facilitate computational tractability at the cost of losing the ability to discriminate among some states $s \in S$ by enlarging the equivalence classes. (i.e., a loss of accuracy)

IMPLICATIONS FOR MODEL COMPOSITION

Most models of real world systems are composites. Why? The scientific laws we work with, whether Newton's Laws or the law of one price, are only applicable under a specific set of circumstances or assumptions. For example, Newton's law of gravitation tells of the strength of the gravitational force between two point masses. What happens if we have more than two point masses? The presumption is that we can reduce the system to pieces where the law or symmetry relationship applies, then put the pieces back together again to the behavior of the whole system. (hence, reductionism)

In terms of our setup, this means we break the observables up into groups and work with the groups separately. This works when the observables are completely unlinked. An example is when we can break the dynamics of a mechanical system up into the x, y, and z components, propagate them separately and then put them back together and obtain the correct position. In this case the model is technically a composite, but the composition is fairly straightforward.

Of course, it is well known that this is not generally the case, which is why most modeling is a little more complicated than this. To understand why, let us look at what happens when we fractionate a system. As explained in the previous section, modeling a subset of the observables implies that the omitted observables are constant. If we leave them in the model, they function as parameters. If these omitted observables or parameters are unlinked with those retained in the model, then it is not a problem.

However, when the dynamics of a set of observables is not totally unlinked, the composition of the fractional models yields a state space that does not completely correspond with the state space of the real system. Mathematically, the state space of the composed model will be larger than that of the real system. For example, for two sets of observables F and G, the state space of the composed model is actually $S/R_F \times S/R_G$, of which S/R_{FG} is only a subset. (Rosen 1978) We have no way to know for sure which states of this enlarged space are real and which are artifacts of the model.

Under these circumstances, we need to build the lost linkage relationships back into the model. Let's consider a very simple example. Assume that we have a bank account that earns compound interest continuously. If we assume that interest rate is fixed, we have a very simple equation for determining the accrued value of the bank balance with two observables: balance, B and interest rate r

$$\frac{dB}{dt} = Br$$

$$B(t) = B_0 e^{rt}$$

The introduction of dynamics links the observables, but in this case, the interest rate is a parameter. But as we know, interest rates change over time. One way we could do this is to create a dynamic model $r(t)$ and compose the two. This results in the composition

$$\frac{dB}{dt} = Br(t)$$

One could imagine many other variations to increase the realism of the model by adding back in broken linkage relationships (e.g., include withdrawals and deposits, make $r(t)$ a stochastic process, etc.). However, for this discussion we are interested in considering the circumstances under which this composition is acceptable. If you will recall the discussion of linkage, we were careful to specify that complete linkage between observables involves each equivalence of one observable intersecting the other and vice versa. It is also possible to have an observable f linked to another observable g, but g is not linked to f. In this example, the model assumes that B is dynamically linked to r, but r is not linked to B. In other words, we assume that no matter what amount of money we deposit in our bank account, it will have no effect on the interest rate. This is a standard assumption for an individual small investor, but what happens when the depositor is a large government, or what if we scale up to model all deposits in the country? Then the balance will have an impact on the interest rate. In this case, the assumption of a one-way linkage is no longer valid.

This example naturally leads into a discussion of the standard cases of model composition and the associated requirements.

Necessary Conditions for Model Composition

If we start with the condition that the diagram in Figure 12 must commute for (X, D_X) to be a model of the system $(S/R_F, T_F)$, then we can view model composition as decomposing $(S/R_F, T_F)$ into pieces $\left(S/R_{G_i}, T_{G_i}\right)$, modeling pieces individually (X_i, D_i), and then composing those pieces to yield (X, D_X) while preserving the commutativity of the diagram. In other words, we make the diagram increasing more complicated as we add structural requirements.

To facilitate the discussion, let us define some notation. First, let us partition the set of observables F into n subsets G_i. Applying any one set of observables to the system yields the reduced state space S/R_{G_i}. To capture the dynamics under this subset of observables, we need to project the dynamics T_F into the subspace S/R_{G_i}. We will call this projection, T_{G_i}. This results in the reduced description of the system $\left(S/R_{G_i}, T_{G_i}\right)$. If we want to model the system, we need to build a diagram that commutes using the model (X_i, D_i).

$$
\begin{tikzcd}
S/R_{G_i} \arrow[r, "T_{G_i}"] \arrow[d, "\alpha_i"'] & S/R_{G_i} \arrow[d, "\alpha_i"] \\
X_i \arrow[r, "D_i"] & X_i
\end{tikzcd}
$$

Figure 13 - Subsystem Diagram

The challenge then becomes, can we assemble a set of diagrams like this one and still preserve commutativity over $S/R_F \overset{T_F}{\rightarrow} S/R_F$? Of course, this all depends on the nature of the linkage relationships among the subsets of observables. Note that we make no assumptions about the nature of the linkage relationships within any given subsystem. Since we are interested in reuse and composition, we are concerned with the case where we already have a scientific paradigm or model that accounts for the behavior of S/R_{G_i}. In other words, if we totally isolate S/R_{G_i}, then we have a model of its behavior already (X_i, D_i). This is essentially the what the scientific method does. But what happens if this subsystem is not totally isolated from other subsystems? We will consider four cases that depend on the nature of the linkage relationships among the subsystems:

- *Unlinked* – there is no relationship among the subsets of observables of the various subsystems
- *One-way dynamic linkage* – Any combination of states among the subsystems is allowable, but the state of one of the subsystems affects the state transition behavior of another, but not vice-versa
- *Two-way dynamic linkage* – Any combination of states among the subsystems is allowable, but that combination affects the state transition behavior of all subsystems
- *Two-way static linkage* – Not all combinations of states among the subsystems is allowable.

For each of cases, we will consider how we will have to modify the necessary condition for a model (Figure 12) to accommodate (X, D_X) resulting from the composition of multiple sub models. For the sake of illustrative simplicity, we will only consider two subsets of observables for each case, but it should be obvious to the reader how they can be extended to more than two subsets. We should also note that these diagrams are not intended to be representative of how one would actually build the model. Rather they express the mathematical conditions that must be met if one wanted to build a model.

Unlinked Observables

The most straightforward case for model composition is the case where the subsets of observables are totally unlinked. This effectively allows us to model each subset independently and still yield correct answer. If this is the case for two subsets of observables, then the diagram show in Figure 14 commutes.

$$
\begin{tikzcd}
& S/R_F \arrow[dl, "\pi_1"'] \arrow[dr, "\pi_2"] \arrow[ddddd, bend left=80, "T_F"] & \\
S/R_{G_1} \arrow[d, "\alpha_1"'] & & S/R_{G_2} \arrow[d, "\alpha_2"] \\
X_1 \arrow[d, "D_1"'] & & X_2 \arrow[d, "D_2"] \\
X_1 & & X_2 \\
S/R_{G_1} \arrow[u, "\alpha_1"] & & S/R_{G_2} \arrow[u, "\alpha_2"'] \\
& S/R_F \arrow[ul, "\pi_1"] \arrow[ur, "\pi_2"'] &
\end{tikzcd}
$$

Figure 14 – Necessary condition for a composite model for unlinked observables

Again, a good example of this is when we can decompose the dynamics of an object into x,y, and z components, propagate them separately and obtain the correct position. Obviously this is the simplest form of model composition.

One-way Dynamic Linkage

Now we will loosen the assumption that the observables are dynamically unlinked aand consider the case where subset G_2 is dynamically linked to the subset G_1 but not the other way around. This situation requires that the diagram shown in Figure 15 commutes.

$$
\begin{tikzcd}
 & S/R_F \arrow[dl, "\pi_1"'] \arrow[ddr, "\alpha"] \arrow[ddddd, bend left=80, "T_F"] & \\
S/R_{G_1} \arrow[d, "\alpha_1"'] & & \\
X_1 \arrow[d, "D_1"'] & & (X_1, X_2) \arrow[d, "D_2"] \\
X_1 & & X_2 \\
S/R_{G_1} \arrow[u, "\alpha_1"] & & S/R_{G_2} \arrow[u, "\alpha_2"'] \\
 & S/R_F \arrow[ul, "\pi_1"] \arrow[ur, "\pi_2"'] &
\end{tikzcd}
$$

Figure 15 – Necessary condition for a composite model with one way dynamic linkage

At first glance, it might seem that there is no gain from decomposition since the whole state space gets encoded in the model for G_2. The gain comes when we observe that we were able to decouple the dynamic propagation of G_1. What this allows us to do is generate the state space trajectories for G_1 first, then compute the state space trajectories for G_2, using the precomputed trajectories of G_1 as an input. In physics-based modeling, this case enables what is called serial multi-scale modeling. A non-physics based example would be the bank account balance model described in the previous section where we could generate a time varying trajectory of the interest rate then use that in our model of the bank balance to generate a trajectory of the bank balance over time.

Two-way Dynamic Linkage

Now we will consider the case where G_1 and G_2 are mutually dynamically linked. That is the current state of G_1 affects the dynamics of G_2, and vice-versa. This diagram requires that the diagram shown in Figure 16 commutes.

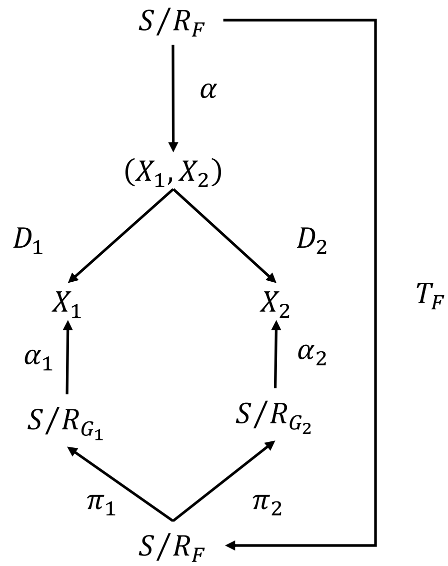


Figure 16 – Necessary condition for a composite model with two way dynamic linkage

Here the situation is slightly more complicated. We need to know the states of both subsets in order to know the dynamics. However, there is a still a computational advantage. What this situation allows us to do is to have separate dynamic models for each subset of observables. Examples would be a cellular automata model that allows us to update the state of each cell sequentially, or a model of two bodies under Newtonian gravitation where we can determine the position change for each body independently for an infinitesimally small time interval. All else being equal, this requires more computational coordination than previous case, but it is still tractable in many circumstances.

Two-way Static Linkage

Now we will consider the case where G_1 and G_2 are statically linked. That means there are infeasible combinations of classes of R_{G_1} and R_{G_2}. There are two common situations where this is encountered. First, there is a symmetry relationship among observables. For example, if the observables voltage, resistance, and current are separated into different sub models, Ohm's law would mean that there is symmetry relationship that is statically linking the submodels. Second, when the same system is represented using different abstractions there may be implicit linkage relationships. An example would be representing the same system using both quantum mechanics and molecular dynamics. This latter case arises when some aspects of system behavior are better captured with one abstraction than another. The composition is then an attempt to exploit the advantages of each.

To compose two models successfully for a two-way static linkage, the diagram show in Figure 17 must commute. This diagram requires some explanation because as we will see, this is a very challenging condition to satisfy.

$$
\begin{tikzcd}
 & S/R_F \arrow[d, "P_F"] \arrow[dddddd, bend left=90, "T_F"] & \\
 & S_1/R_{G_1G_3} \times S/R_{G_2G_3} \arrow[dl, "\alpha_1"'] \arrow[dr, "\alpha_2"] & \\
X_1 \arrow[d, "D_1"'] & & X_2 \arrow[d, "D_2"] \\
X_1 & & X_2 \\
S_1/R_{G_1G_3} \arrow[u, "\alpha_1"] & & S/R_{G_2G_3} \arrow[u, "\alpha_2"'] \\
 & S_1/R_{G_1G_3} \times S/R_{G_2G_3} \arrow[ul, "\pi_1"] \arrow[ur, "\pi_2"'] & \\
 & S/R_F \arrow[u, "P_F"'] &
\end{tikzcd}
$$

Figure 17 – Necessary Conditions for a composite model with a two-way static linkage

First, we need to introduce some additional notation. In this setup, we consider the case where we have $S_1 \subseteq S$. This means we want to model some subset of S with a different abstraction than S as a whole. An example would be we went want to model some part of the system in a higher resolution than the rest. To determine the subset S_1, we need some set of observables to define it. We will call this subset, G_3. A typical example of this set would be the Cartesian coordinates. If we were modeling the behavior a block of material, we might model the whole block using the abstraction G_2, but also model a piece of the block using abstraction G_1. Since we are trying to model each abstraction separately, what we are really modeling over is a the state space $S_1/R_{G_1G_3} \times S/R_{G_2G_3}$. (Note that G_3 is left in as a reminder that it is the determinant of S_1. This is a very different space than the original S/R_F. We introduce the mapping P_F to get us from S/R_F to $S_1/R_{G_1G_3} \times S/R_{G_2G_3}$. Note that P_F involves a substantial loss of information. If (X_1, D_1) and (X_2, D_2) are standard representations for the abstractions G_1 and G_2 in isolation, then the linkage information is completely lost. Thus, only way this diagram could commute is either 1. If there is no linkage relationship, which violates our starting assumption, or 2. If we somehow build the linkage relationships back into the models (X_1, D_1) and (X_2, D_3). This would mean that we would likely need some or all of the state information from the other abstraction in each model as well as knowledge of the linkage relationships. In other words we have to put the lost information back in somehow. In the most general case, this means that have to build a new model that accommodates all of the observables and linkages. This means we lose any computational gain we may have gotten from breaking the representation into sub models.

To put it another way, this would be analogous to breaking up Ohm's law into separate models for V and IR. Propagating them separately is likely to violate the required relation $V = IR$. So we are forced to include Ohm's law in each model as well as the omitted states. The result is that we end up building the same model twice. There is really no gain from decomposition.

There are some cases where we can compose multiple models with a static linkage among observable sets, but they require some modifications to this general case. The first is when there is a refinement relationship between G_1 and G_2. If G_1 refines G_2 then each equivalence class of G_1 intersects exactly one equivalence class of G_2, but any given equivalence class of G_2 may intersect more than one class of G_1. This is an aggregation relationship between G_1 and G_2 which is equivalent to a one-way static linkage. So the two abstractions are compatible, and the linkage relationship is known, but G_1 allows us to resolve more system states that G_2. Thus, we could view it is a higher resolution model. Under these circumstances, we can run the model (X_1, D_1) first, then use it to parameterize (X_2, D_2), which we run second. This illustrates case of multi-fidelity modeling where we conduct limited number of runs of the high fidelity model to calibrate a lower fidelity model that we used to explore a larger space.

The second case where we can compose multiple models with static linkages among them is explained by parallel multi-scale physics-based modeling. The example described by Winsberg (2010) is one such instance. This example is discussed in greater detail in Section 6, but to summarize briefly here, it involves the modeling of crack propagation using the simultaneous application three different abstractions: continuum mechanics, molecular dynamics, and quantum mechanics. The reason is that different aspects of crack propagation are best represent using different abstractions. However, unlike the previous case, there is a two-way linkage. This means that the abstractions each affect each other.

To make this work, we need to modify Figure 17. We do this by dividing the state space S into multiple subsets, S_1 and S_2 by employing the subset of observables G_3. Again a spatial division is a good example, but not strictly required. The idea is apply abstraction G_1 to S_1 and abstraction G_2 to S_2. Now if there were absolutely no overlap between S_1 and S_2 then this would imply that we could accurately model the system with no interaction among the subsets. This devolves to the unlinked case and violates the assumption that the regions affect each other. Consequently, there must be at least some overlap among the two subsets, but we would like to keep it to a minimum. For instance, it might only be a shared boundary. Mathematically, $S_1 \cup S_2 = S$ but $S_1 \cap S_2 \neq \emptyset$.

However, the overlap introduces a problem. We are now trying to represent a subset of the space using two different abstractions. This would seem to put us in the same situation as Figure 17. As Winsberg notes, this is handled by introducing fictions. In our setup, these would be observables that we invent but cannot really measure. An example provided by Winsberg is the introduction of fictitious "Silogen" atoms on the boundary between the region modeled using molecular dynamics and the region modeled using quantum mechanics. There is no such thing as a Silogen atom, but it serves to link the two different abstractions by capturing the effect each has on the other. In short, we get a more accurate model of the whole system at the price of lost information about overlap region. As long as we keep the overlap region small, this can be acceptable price to pay. When we introduce the "fictions" X_3 and X_4 to our setup, we get the diagram in Figure 18. If this diagram commutes, we can use a parallel multi-scale model (or something analogous) to capture the behavior of the system.

$$
\begin{tikzcd}
 & S/R_F \arrow[d, "P_F"] \arrow[dddddd, bend left=60, "T_F"] & \\
 & S_1/R_{G_1G_3} \times S_2/R_{G_2G_3} \arrow[dl, "\alpha_1"'] \arrow[dr, "\alpha_2"] & \\
(X_1, X_3) \arrow[d, "D_1"'] & & (X_2, X_4) \arrow[d, "D_2"] \\
X_1 & & X_2 \\
S_1/R_{G_1G_3} \arrow[u, "\alpha_1"] & & S_2/R_{G_2G_3} \arrow[u, "\alpha_2"'] \\
 & S_1/R_{G_1G_3} \times S_2/R_{G_2G_3} \arrow[ul, "\pi_1"] \arrow[ur, "\pi_2"'] & \\
 & S/R_F \arrow[u, "P_F"'] &
\end{tikzcd}
$$

Figure 18 – Necessary condition for a composite model with a two-way static linkage with minimal overlap

A few things to note. First, we still lose the linkage information among the two groups of observables when we instantiate the two fractionated models. The fictions, X_3 and X_4 are introduced to compensate, but this means that they need to be determined empirically through trial and error. This is consistent with Winsberg's observations. So the way this would typically be employed is that the structure of the fiction is determined empirically, and then the particular value is determined by a combination of the states of each of the abstractions. The important observation to make here is that if (X_1, D_1) and (X_2, D_2) are two off the shelf representations we cannot compose them "on the fly" in the event that there is a two-way static linkage. Rather some work is necessary to create the properly calibrated fictions.

IMPLICATIONS

What we can draw from this analysis, is that multi-modeling is not so straightforward as it is sometimes portrayed. In many cases, it is not a simple matter of linking up inputs and outputs. If we consider where much of emphasis for federating models has gone (e.g., DMMF, HLA, SPLASH, etc.), the focus is on coordinating the computational aspects. That is the $D_i's$ in our diagrams. When we are running two or more models in parallel we would need them to be synched in terms of time steps and observable scales. (Otherwise the diagram won't commute.)

When models are drawn from off the shelf, there is no guarantee that this is the case. Thus, there has been a great deal of work put into achieving this synchronization.

However, what these frameworks implicitly assume is that they are operating the in the case of two-way dynamic linkage in our setup. That is there are no static linkages among the observable subsets allocated out to the different models. This is a pretty big assumption, particularly when models are drawn from off the shelf. It is not clear just what assumptions were made during their development. If there are static linkages among the represented observable subsets in the real world, then a composed model that ignores these may very well produce incorrect results. For complicated or complex systems, the required level of decomposition among observables is unlikely, particularly when the simulation is intentionally trying to represent the same system at different levels abstraction as was the case for DMMF. In that case, the static linkages are intrinsic.

Even in the case where each model represents a different “subsystem” there is likely some overlap among the models because each model will need to consider its environment or bounds. When multiple models each attempt to represent the same environment or bounds differently, there is a static two-way linkage. As we saw in the final case, it is possible to manage this, but it requires empirical investigation and calibration to get it to work. That is it cannot be accomplished “on the fly” Thus, we see the reasons why frameworks such as HLA have run into trouble. Unless the static linkages among the models are well understood, the simulation will not produce accurate results. Syncing scales, times and data exchange is not enough. This also explains why multi-modeling efforts seem to require human subject matter experts in the loop.

So while computational coordination is important, successful multi-modeling requires an understanding of how the simulated entities actually relate to each other in the real world. Thus, those who have been working on referential ontologies to support modeling and simulation seem to be moving in the right direction. While these ontologies will not make the multi-modeling process automatic, they will certainly facilitate efforts to build multi-models. In particular, efforts to capture multi-scale or multi-abstraction ontologies would be particularly useful.

CONCLUSIONS

Given the above analysis, what are implications for the enterprise modeling methodology? In the general sense, building a multi-model with different layers of abstraction is possible, but it requires some work. Also, certain assumptions must be satisfied and overlap among the layers should be minimized. More importantly, the interconnections among the layers require “fictions” that can only be determined empirically through trial and error. It is this latter aspect that is the most problematic for modeling enterprise systems.

For a pure physical system, the circumstances maybe stable enough that one can conduct experiments and collect enough data and determine a valid implementation of the "fictions" to handle the overlaps among the layers. However, for an enterprise system, this is challenging for two reasons. First, is often difficult or impossible to run controlled experiments on enterprises. Second, enterprises are constantly evolving. So even if one could conduct experiments on an enterprise, it is like chasing a moving target. Since the fictions aren't "real," the right fiction may very well change as the enterprise evolves. This changes the role of multi-modeling for enterprise systems versus when it is applied to physical systems. We can experiment with different hypothesized linkage relationships to establish the possibility of relevant behaviors, but we cannot be certain that they are real. The implications of this situation are explored in greater depth in Section 6.

5. Review of Complexity Literature on Warning Signals

As discussed previously, no model is capable of forecasting all possible scenarios. Consequently, some changes in enterprise behavior will be a surprise. Some researchers have investigated the possibility that there are early warning signals of such "surprises" in biological and financial systems. In this section, we critically review the literature on this topic and consider whether or not these techniques can be used to mitigate the impact of a surprise change in enterprise behavior.

Much of the work in this area is based on the notion that surprises are bifurcations in the system of interest's dynamic behavior. Consequently, researchers have leveraged work in dynamical systems theory and complexity science to identify warning signals that one should expect to see prior to the actual bifurcation. To operationalize this concept, they propose statistical measures and experiments that should enable detection of these warning signals. The hope is that that these signals are generic in the sense that they should be present in any type of complex system whether physical, ecological, or social. In this section we will briefly review the literature with regard to several proposed statistical techniques.

Random Variable Moments - Skewness, Kurtosis

The most general of the warning signal measures is the analysis of higher order moments of random variables. While the first two moments of the variable (mean and variance) are used for model estimation, the third and fourth moments, skewness and kurtosis respectfully, capture more complex measures within the variables. It is then inferred that these capture more complex dynamics and thus drive critical transitions.

$$Standard\ Moment = \frac{\mu_k}{\sigma^k}, \qquad \mu_k\ is\ kth\ moment\ about\ the\ mean$$

$$k = 1\ \parallel \frac{\mu_k}{\sigma^k} = 0 \qquad\qquad k = 2\ \parallel \frac{\mu_k}{\sigma^k} = Variance$$

$$k = 3 \ \| \frac{\mu_k}{\sigma^k} = Skewness \qquad k = 4 \ \| \frac{\mu_k}{\sigma^k} = Kurtosis$$

While these measures are easy to use, they are limited by their granularity. Consequently, their role is often more of a starting point for analysis than a complete solution.

METRIC BASED CORRELATION – AUTO-CORRELATION, PEARSON CORRELATION

Moving from a single random variable to the comparison of multiple random variables, we next consider correlation metrics. The rationale behind the use of correlation metrics is that complex systems exhibit various forms of correlated behaviors when in a critical state prior to a bifurcation. Different transformations analyze changes in correlation profiles which correspond to potential non-linear effects driving the system. These are exploratory measures that analyze correlation profiles, and two common methods are described below.

Autocorrelation looks at the correlation profile among variables over time periods. This is done by analyzing single correlation coefficients (usually the first coefficient) from a particular model. Other higher fidelity measures are available but require process model assumptions (Scheffer, et al. 2009).

$$\rho_1 = \frac{\mathbb{E}[z(t-\mu)(z(t+1)-\mu)]}{\sigma(z)^2}, \qquad \rho_1 = First\ Correlation\ Coefficient$$

The Pearson correlation coefficient is defined as the ratio of the covariance of the variables to the product of the standard deviations of the variables. In the context of complex systems, it is used to show grouping behavior amongst functional elements.

$$\rho_{x,y} = \frac{Cov(x,y)}{\sigma_x \sigma_y}, \qquad \rho_{x,y}\ is\ Pearson\ Coerrelation\ Coefficient$$

This metric has been shown to have a strong relationship to geometric changes (Kéfi, et al. 2014), but it requires data on model vectors for convergence. It may be useful for determining geometric changes in dynamics on network. For example Chen, et al. (2011) use this metric as an early warning signal for abrupt deterioration during the progression of a complex disease.

STATE SPACE ESTIMATOR ANALYSIS – AR(P) MODEL METRICS

Assuming that there is an accepted state space model(s) of a system, one would like to know if there is a change in model factor significance. Such changes may be a signal of non-linear dynamics. The factor model (and its assumptions) can determine the particular transformations and metrics (Ives & Dakos 2012, Bartholomew et al. 2011). However, as the system changes, the relative factor eigenvalues will begin to deviate. Metrics that trace changes in eigenvalues can serve as a means for transition identification. The eigenvalue behaviors can indicate loss of independence, changes in stability, flickering among states, etc. Ecosystem models often use auto-regressive models which vary over different time periods [AR(p)], and have found flickering as a precursor to critical transitions that move the ecosystem into eutrophic states (Wang 2012).

The benefit and limit of state space estimators is their inference of dynamics given what is assumed to be the system structure. Kalman filtering is a common statistical modeling paradigm that takes initial conditions on the system state space and does inferences based on system changes over time periods. The method fits a parametric model with assumptions and then continuously infers state change components from statistical difference estimators. This can be run through additional hypothesis testing of system structure. This method is common as it allows inference that model parameters are reaching assumption boundaries, and different null hypotheses can be tested to ensure other non-linear effects are not present.

RESIDUAL ANALYSIS – CONDITIONAL HETEROSCEDASTICITY

Even as the factor model may not vary significantly until the critical transition is determined, analysis of the residual terms can show the changes in underlying unmeasured or latent factors. Similar to higher order moments on the random variable, the higher order analysis on residuals should capture potential changes in the underlying regressed model assumptions. There are a variety of tests available, but common ones are residual factor models, conditional heteroscedasticity, and various transformations for variable testing.

$$R = \mathrm{X} - \bar{\bar{\mathrm{X}}}, \qquad R = Residual\ Matrix \ \ \bar{\bar{\mathrm{X}}} = \text{ Factor Model}$$

$$r_n = x_n - \bar{\bar{\mathrm{x}}}_n, \qquad \Lambda = \begin{bmatrix} Cov(r_1, r_1) & \cdots & Cov(r_1, r_n) \\ \vdots & \ddots & \vdots \\ Cov(r_n, r_1) & \cdots & Cov(r_n, r_n) \end{bmatrix}, \qquad \Lambda = \text{Residual Covariance Matrix}$$

In area of ecology, one such technique is called conditional heteroscedasticity. Ecology can often fit regular factored, autoregressive models, but latent populations determine changes in ecological states (Seekell, et al. 2011, DeYoung, et al. 2008). Thus, the existence of different population groups makes the models susceptible to heteroscedasticity. An example is when the make-up of the population changes due to an invasive species. The approach works by taking the residuals at each time step and auto-regressing each residuals matrix. Then the regressed residuals are compared for positive or negative variation between time steps. When there is correlation among residuals, it is assumed that there are higher order factors influencing the system.

These changes in residuals have been shown to identify new, uncaptured factors that can lead to system transitions. For example BDS testing (Brock, et al. 1996) captures the residuals and tests factors for general independence and determines whether their distributions are identical. Other measures test for hidden non-linearities within the residuals that may have significance (Maydeu-Olivares & Joe 2005, 2008; Reiser 1996, 2008).

There are also general and specialized statistical treatments of residual values for estimating latent values. These measures have use in detecting changes in latent structures which is thought to present themselves in critical transitions within social systems (Vallacher & Nowak 1997, Geels 2005).

GLOBAL STOCHASTIC MEASURES - GRANGER CAUSALITY

Another common approach to identifying critical phenomena is Granger causality. This measure is based on information theory and the variation in informational entropy. Shannon and Renyi entropy which identify the conditional mutual information shown below are often used. The idea is that information conditionality corresponds to causality as defined within this framework. This allows one to infer potential future behavioral from system signals (Bossomaier, et al. 2013).

$$H(X) \equiv -\sum_{x} p(x) \log_2 p(x)$$

$$I(X{:}Y) \equiv H(X) + H(Y) - H(X,Y) = \sum_{x,y} p(x,y) \log_2 \frac{p(x,y)}{p(x)p(y)}$$

Information rates known as Kolmogorov-Sinai entropy can be computed within dynamical systems. What can then be shown is that for stationary stochastic processes, each process has a bijection to a "measure-preserving dynamical system" (Hlavackova-Schindler, et al. 2007). This allows a system to be sub-divided into separate stochastic processes. Through entropy measures, one can trace the change conditionality among them. Any decomposable stochastic system can be analyzed in this framework. The measure then becomes useful in identifying transition pathways for the system, but then limits its detection of smaller scale changes as these are assumed to be stochastic.

$$F_{Y\to X} \equiv I(X_t{:}Y_{t-1}, Y_{t-2}, \dots | X_{t-1}, X_{t-2}, \dots), \qquad F_{Y\to X} = \frac{1}{2\pi}\int_{-\pi}^{\pi} f_{Y\to X}(\omega)d\omega$$

$$X_t = A_1X_{t-1} + A_2X_{t-2} + \cdots + B_1Y_{t-1} + B_2Y_{t-2} + \cdots + \varepsilon_t$$

There are multiple applications for partial Granger systems (Guo, et al. 2008). Hysteresis effects can be estimated by analyzing the information mutual to elements through data analysis means such as k-nearest neighbors (Kraskov, et al. 2004). Granger causality has been shown in multiple instances where measured change in mutual information corresponds to impending changes within the system. This has thus been used as a critical transition metric within multiple contexts such as socio-economic systems (Bossomaier, et al. 2013), financial systems (Barnett & Bossomaier 2012), and neurology (Guo, et al. 2008).

Since discretized variables and spectral decomposition are needed for convergence, there are different variations based on: the choice of variables elements [multivariate Granger], the type of relationships [linear vs partial Granger], and the potential changes in latency and ordinality [conditional Granger]. For this reason, there are underlying structural decisions made before running Granger type measures about the nature of stochastic variables (Bossomaier, et al.

2013) and about the information model within the computational estimator (Kraskov, et al. 2004).

Alternate Exploratory Statistical Measures

Transfer entropy measures are used in conjunction with Granger causality and other stochastic measures. The level of mutual information can be analyzed itself as a means for exploratory analysis of hysteresis effects (Schreiber 2000). Shannon and Renyi entropy are commonly used for this.

Dimensional analysis is also used as a means to identify major latent changes in system function. This has been applied in the financial domain to identify periods that violate the efficient market hypothesis. There is an example of using a log-periodic measure to identify higher order fluctuations (Yan, et al. 2010) and discrete scale invariance (Sornette 1998). This is thought to show periods where second order dynamics (traders trading on trade decisions rather than information) occur. This adds dimensions to market information. The claim is that these findings presage periods of market volatility and subsequent crashes. It is thought that identifying periods where markets become heteroskedastic can identify critical transitions in the market.

Perturbation Experiments

As a system approaches a critical transition, the strength of the current attractor begins to weaken. Consequently, one would expect a loss of responsiveness to disruptions. This is the rationale behind perturbation experiments. The idea is that one intentionally perturbs the system and then observes the result. If the system is near a bifurcation point, the response to the perturbation should display unique, complex phenomena. The time series of the response is analyzed to look for a critical slowing down behavior. The idea of intentional perturbations also presents as a potential management practice to increase system resilience (Walters 1997, Wilson 2002).

Limitations

While the idea of using warning signals to predict an impending shift in system behavior has an intuitive appeal, there are some practical limitations.

First, a system approaching a phase transition is expected to show more complex behavior. The most obvious condition is that the system deviates from model predictions. This results in a trade-off between modeling and warning signal usage. The significance and reliability of a warning signal depend on a tradeoffs between minimizing model ontology, and having enough noise reduction to give significant or convergent results (Ditlevsen & Johnsen 2010, Perretti & Munch 2012).

Second, within the area of statistical measures, the central difficulties encountered were those around inference. Most notably as metrics seek to trace underlying structure that produce observed statistical profiles the Bertrand Paradox presents itself. One example is an attempt to identify epilepsy through neural signal processing. There is a study trying to identity the bifurcation within neural models that leads to seizures that can be seen using state space estimators (Rodrigues, et al. 2010). Yet the same deviation seen by the estimator preceded benign neural spikes, and the state space model had to be expanded to find differing profiles for these “false bifurcations”. This open inference problem makes determining singular root causes to particular bifurcations difficult. This means for instance that although a critical slowing down may be observed this does not guarantee a fold bifurcation as it is possible that another underlying system structure can produce this same dynamic. Although there is contention in the area, there is still no method for resolving it other than mixing structural information. In short the existence of a warning signal is not a sufficient condition.

Third, complexity can be driven by latent factors. This creates difficulty in responding to a warning signal. For example, with invasive species, metrics can identify the system irregularity but not identify the new species; it usually takes matching with existing observations of the ecosystem (Wang 2012). A particularly challenging set of latent factors that are relevant for enterprise systems those associated with human behavior such as beliefs and culture that significantly drive behavior. This presents additional challenges to responding to phase transitions in enterprises as driving factors not only may be unidentified factors but also humans invented factors.

Fourth, warning signals are impacted by boundary decisions. Depending on how the boundary is set, critical behaviors may be driven by exogenous rather than endogenous factors. This means that the warning signals could be missed and the transition perceived as exogenous shock. For instance, a primary goal of analyzing warning signals for market crashes is determining which of the market movements can be attributed to external shocks versus internal market forces thought to be part of the natural business cycle (Fry 2012).

CONCLUSIONS

While the idea statistical warning signals of impending system behavioral or structural shifts is appealing, it is not clear from the literature that these could be effectively applied to support decision making with regard to enterprise systems. More research is required.

There are natural limits to information that can be gathered a priori. Parameterized warning signals depend on known model structure or defined assumptions. Non-parameterized approaches are more flexible but are less sensitive. Perhaps the biggest issue is that one has to know where to look. Which metrics should one be looking at to spot a warning signal? This may be easy to determine after the fact, but not so easy before the event. This is especially true for systems that can be viewed using many different abstractions such as enterprise systems. Which abstraction should one be looking at?

In short, statistical warning signals may be useful for cases where we don’t have a well-defined causal mechanism to explain the bifurcation point, but we have experienced the event and the warning signal before enough times before to have some confidence. Of course, given the adaptive nature of human beings, even those cases could be of little predictive value.

6. Implications for Enterprise Modeling and the Strategy Framework

The previous sections of this report detailed efforts to evaluate different aspects of the enterprise modeling methodology via counterfeit parts case study, the behavioral economics case study, and a review of the complexity literature. As a consequence of these efforts, it was recognized that the development of enterprise models and subsequent strategy decisions are driven by the type of uncertainty that a decision maker faces. In particular, we determined that epistemic uncertainty is type of interest, and multi-modeling is a means to explore certain sub-types of epistemic uncertainty.

This recognition resulted in an expanded taxonomy of uncertainty that informs both enterprise modeling efforts and the resulting enterprise strategy. In this section, we will consider the literature on managing uncertainty for multi-models, assess the limitations of the proposed methods with regard to enterprise systems, and finally consider the implications for enterprise modeling and the strategy framework developed during RT-110.

Background

It has long been recognized within the systems community that to fully capture many real world systems, it is necessary to represent them using multiple models (Haimes 1981, Hall 1989). Depending on the bent of the researcher, this is sometimes called multi-resolution, multi-scale, multi-method, multi-discipline, or multi-level modeling. In this paper, we will simply refer to all such approaches as multi-modeling. Regardless of the term employed, the intent is the same. Different models each lend different advantages and disadvantages in terms of phenomena captured, computational burden, accuracy, etc. The idea behind multi-modeling is to leverage the advantages of each model by analytically or computationally composing them into a single model that addresses one or more questions of interest. The composed model is then used to perform trade studies or support decision making.

Such approaches are fairly common in engineering design, for instance in aerospace engineering where it is called multi-disciplinary optimization (MDO). These approaches certainly have their challenges, and there is a rich literature that addresses the management of uncertainty in this domain that falls under the title uncertainty quantification (UQ). Our concern here is the additional risk that results when we extend the multi-modeling paradigm to enterprise systems in an effort to capture the impact of behavioral and social issues on engineered systems (and vice-versa). While the natural approach would seem to be to computationally integrate behavioral and social models with engineering models and then optimize, we assert that fundamental epistemic limitations in the modeling of enterprise systems limit the efficacy of this approach.

In particular, we will argue that the complexity of behavioral and social systems increase the prominence of three additional sources of epistemic uncertainty above what is typically encountered during physics based modeling: phase, structural, and ontological. In this paper we characterize these uncertainties as bifurcations that can be difficult for models to capture. The consequence is that a component model in a multi-model is inadvertently pushed out of its zone of validity. This is not to suggest that these uncertainties are not present in physics-based models, but rather it is a matter of degree. It is likely that for many enterprise systems these sources of epistemic uncertainty become so overwhelming that building a very detailed multi-model becomes counterproductive. Rather it is more likely that decision makers would prefer to use simple models that capture the essential features of the problem and then employ an adaptive or hedging strategy to address the epistemic uncertainties in bulk.

During RT-44 and RT-110, we considered a few examples of the use of multi-modeling to explore enterprise systems. We would now like to reconsider these examples from an uncertainty perspective. To that end, we will briefly summarize each of these examples. First, Park, et al. (2012) developed a multi-level simulation to examine policy alternatives for an employer-based prevention and wellness program. Their model consists of four levels: ecosystem, organization, process, and people. Second, SPLASH is an effort by IBM Research to develop a framework for loosely coupling models from different domains to support decision making regarding complex socio-technical systems (Barberis et al. 2012). SPLASH mainly focuses on achieving a methodologically valid combination of different modeling formalisms by developing methods to coordinate inputs and outputs, synchronize time, etc. Third, the Dynamic Multilevel Modeling Framework (DMMF) was an effort by the US Department of Defense to leverage its enormous investment in models and simulations across four levels: campaign, mission, engagement, and engineering (Mullen 2013). The objective was to create a framework to allow simulations from each of these levels to interoperate, but the DMMF effort ended at the feasibility study stage.

Generally speaking, the motivation behind such modeling efforts is to account for instances where the decision variable under the control of the decision maker is far removed from the effect of interest, necessitating a chain of models to link the decision variable to the outcome.

To illustrate the point let us consider the DMMF example. What if an air force is interested in whether or not it should perform an upgrade on its fighter engines to increase speed? Increasing speed is just a means to an end. The reason why they would consider such an upgrade is to improve their chances of winning a war or some other similarly high-level objective. Consequently, decision makers would want to know what the impact of this upgrade would be on the probability of winning a war. So the thought is by linking the engineering simulations to the mission simulations to the engagement simulations, and so on, one could determine what the impact of this low-level engineering change would be on the ability to win a war. However, unless this change is near a tipping point to magnify its impact, the effect will be effectively undetectable due to epistemic uncertainty inherent in models of enterprise systems. If every small change in the low-level model produced a noteworthy effect in the high-level model, we would not have the high-level model. It would be too unstable to be useful. The rationale behind these assertions will be explained in the subsequent sections.

Historically, developing such multi-models of enterprise systems has proven extremely challenging. To examine why this might be, let us reconsider the dynamic toll road example from Section 3. Explicitly or implicitly when a department of transportation creates a dynamically tolled road, it is simultaneously considering two different system views from two different disciplines with an explicit feedback between the two: physical traffic flow from civil engineering and economic decision making from economics. The presumption is that by adjusting the price in the economic view that traffic flow will be affected in the civil engineering view. Billions of dollars have already been spent on this presumption.

If one wanted to model such a situation, the approach seems straightforward. Traffic modeling is well established using both differential equation based and agent based simulation approaches. To model the effect of a dynamic toll, one needs to add the economic response model by adjusting either the entry flow or agent entry decision based on price. Unfortunately, as considered in Section 3, any such composition may not be a correct as higher tolls resulted in higher usage.

The point here is not whether or not such behavior is explainable through existing economic theory. Rather it is that the thinking behind dynamically tolled roads is perfectly reasonable, but potentially incorrect. There is more than likely a toll level at which drivers would be dissuaded from entering the tolled lanes. For instance if the toll were $1000, it seems unlikely that many would use the toll road. Thus, there is likely a zone in which the classical demand model composed with a traffic model would be the correct representation. Unfortunately, at least in the case of Minneapolis, the operating toll range does not seem to be within that zone. The issue is how would one know that a priori? This example highlights the issue of model risk when representing enterprise systems.

EXISTING APPROACHES TO ASSESSING MODEL RISK

There is certainly nothing new about model risk. Thus, the question is how the model risk described above would be handled using existing approaches. Consequently, we will briefly review contemporary approaches to doing so. These approaches tend to focus on quantifying model uncertainty to facilitate threshold-based decision making and robust optimization. The approach to identifying uncertainty in physical multi-models is effectively described by Oberkampf, et al. (2002). While there are several different taxonomies of uncertainty in the literature, we will focus on a fairly common one that breaks uncertainty out into aleatory uncertainty, epistemic uncertainty, and error (Agarwal et al. 2004).

Aleatory uncertainty is considered to be irreducible variation. No amount of new information can remove it. An example would be the outcome of a coin flip. Epistemic uncertainty is considered reducible through the acquisition new information. For example, uncertainty in a measurable model parameter such as mass could presumably be reduced through measurements. Finally, error relates to a deficiency in the modeling and simulation implementation effort itself. An example would be applying inconsistent units of measure in the simulation code.

While minimizing error is certainly important, the problem described in the motivational example is not necessarily an error in implementation. There may be a zone where the model is correct. Thus, dealing with error is outside of the scope of this paper. This leaves us with aleatory and epistemic uncertainty.

There has been a long-running philosophical discussion as to whether or not there is a fundamental distinction between aleatory and epistemic uncertainty. From a modeling standpoint, the distinction is purely a practical one. Once we treat an uncertain event as aleatory, we have essentially asserted that within this model framework, the uncertainty is irreducible and we will represent it using a probability distribution. Consequently, the traffic modeling issue we described does not fall under aleatory uncertainty because it is not something we capture as a probability distribution within the model. Rather it is an uncertainty in the structure and/or parametrization of the model itself which can be reduced by the collection of additional information. In this case, the Janson and Levinson study did exactly that. Thus, it is epistemic.

There are a number of approaches to assessing epistemic uncertainty. Yao, et al. (2011) provide an extensive survey of the literature in this area with regard to multi-disciplinary design optimization. Typical approaches involve using double loop approach where the outer loop is used to vary the epistemically uncertain parameters. As far as characterizing the uncertainty itself, approaches include but are not limited to Bayesian, Dempster-Schafer, and possibility theory (Agarwal et al. 2004). Roy and Oberkampf (2011) provide a thorough discussion of how parametric, numeric, and model form uncertainties can be quantified in an integrated fashion.

Agawal, et al. (2004) illustrate these principles using an aircraft design problem involving three coupled models: one for weight, one for aerodynamics, and one for performance. Essentially, such approaches endogenize the epistemic uncertainty within the model in the sense the second loop provides a systematic way to vary the epistemically uncertain parameters. Once that is in place, optimization approaches can be used to find a Pareto optimal set of design options that balance uncertainty with performance. Of course, actually propagating all of the uncertainties through the models can be computationally challenging, and many researchers are developing approaches to address these challenges.

These types of approaches are sometimes referred to as uncertainty quantification (UQ). Sandia National Laboratory has been particularly active in this area via the Dakota Project (Sandia 2015). Eldred, et al. (2011) provide a detailed description of Sandia's approach to uncertainty quantification. While efforts like Dakota are certainly critical to robust engineering design, we will argue that the problem highlighted in the motivating example would not be addressed by such approaches. In particular, their emphasis is on physical systems where the phenomena are relatively well characterized but measurement error and associated issues can be problematic. If we return to the dynamically tolled road example, varying the uncertain parameters would not have revealed the more fundamental structural issue, that the assumed classical demand model was inappropriate for the situation.

Roy and Oberkampf (2011) note that model form uncertainty is often the dominant source of uncertainty for the test cases that they have examined, but they base their quantification of this error on extrapolation from experimental data. However, an extrapolation would not capture a bifurcation in system behavior. (We do not believe that this assertion is inconsistent with Roy and Oberkampf's acknowledgment of the limitations of their method.) We will subsequently argue that we should expect bifurcations to be more common as we move away from purely physical systems and toward enterprise systems. The reasons are discussed in the subsequent sections.

Revisiting Complexity

In the RT-110 report, we discussed how complexity affects the modeling of enterprise system. To briefly review, the reason why we call such systems complex is because they are difficult to capture with compact, accurate models. (See Alderson and Doyle 2010). Consequently, we often need different models in different situations. This is in part due to the tendency of biological and social systems to adapt and change structure depending on the circumstances. The dynamically tolled road is a perfect example. Under some circumstances, the price serves as classical mechanism to regulate supply and demand. Under other circumstances, the price serves as a signal for the level of traffic congestion on alternative roadways.

To make matters worse, organisms, humans, and organizations all adapt. This means that they introduce new and different ontologies over time. For example, if we go far enough into the past, there was a time when things like currency and gross domestic products did not exist. These are effectively new ontologies that were introduced over time by human beings. Imagine that we had a particularly forward-thinking hunter-gatherer from pre-historic times that was interested in modeling and predicting the future progression of human beings. How would this hunter-gatherer know to include the impact of economics on future human behavior when it has not been invented yet?

To probe this point a little deeper, let us consider some work from biology and social science on this matter. Probably the most critical work that considers this issue is biologist Robert Rosen's “Fundamentals of Measurement and Representation of Natural Systems” (Rosen 1978). Rosen took a very abstract view of modeling systems. He framed the problem as one of systems inducing dynamics on meters and then investigated the implications of that setup. A key notion is that different meters are going capture different features of a system. If you do not know to build a particular meter, you would not measure the features of the system that it reveals. For example, you need a volt meter to measure voltage, but building a volt meter requires the concept of voltage. In a sense, we could argue that we must create new ontologies to create new useful meters (volt meter, gross domestic product). Two ideas that we can take from Rosen's work is that a complex system can rarely be captured by a single model and the notion of model bifurcation.

Bifurcation theory is a result of a very large body of work known as dynamical systems theory. An in depth discussion is not necessary for our purposes, and instead it is sufficient to note that bifurcation theory studies the circumstances when a small change in input results in a dramatic change in output for a dynamical system. In other words, there is a qualitative change in the system.

If we could fully represent a system with a single model, then that model would capture any bifurcation that occurred in a system. However, if we need to employ multiple, different models to capture a system, it suggests that we are missing some of the bifurcations. Thus, the model may bifurcate from the true system over certain spaces. If we cannot, as Rosen asserts, see the true system, then we can only see the models. Thus, what we are seeing is one model of the system bifurcating relative to another model of the system.

This viewpoint was expanded by Casti (1986) who considered the importance of bifurcations to levels of abstraction and system complexity. In particular, Casti asserts that levels of abstraction are the result of system bifurcations, and casts complexity as the number of non-equivalent descriptions required to represent a system for a given set of observables. The larger implication of Casti's work from our perspective is that if social systems are complex, there are many possible bifurcations; hence there are many different models and theories in social science. Each model or theory has an element of truth, but it is easy to over extrapolate, exceed the bifurcation point, and obtain bad predictions.

In work that that takes a slightly different approach but achieves complementary results, Harvey and Reed (1996) wrote “Social Science as the Study of Complex Systems.” They proposed a fourteen level hierarchy of ontological complexity. Paraphrasing, their hierarchy starts with physics and the bottom and ends with the evolution of social systems on the top. Harvey and Reed argue that the complexity of each layer is greater than the one below, which results in fundamental epistemological differences as we move up the hierarchy. In short, complexity limits what we can know about a system. As a direct consequence, different classes of models are appropriate for different levels of the hierarchy. What this means is that even though social systems are effectively physical systems at the bottom, as a practical matter, one cannot capture the behavior of a social system using a physical model. Multiple models at different levels of abstraction are a necessity for understanding enterprise systems.

What we can draw from this discussion is that a major challenge of modeling enterprise systems is shifting and incompatible models. In other words, there are likely to be more bifurcations when representing an enterprise system vice a purely technological system. When these bifurcations are not specifically known and quantified, they constitute a source of epistemic uncertainty. While uncertainty quantification methods aim to address model form uncertainty via extrapolation, it is not likely that such methods would address these bifurcations. A question naturally follows: How does one understand and manage the risk of bifurcations when modeling and simulating enterprise systems?

MODELS AND BIFURCATIONS

We initially discussed the concept of model bifurcation in RT-110. However, the work conducted as part of the RT-138 has allowed us to expand on the concept. The chief issues is when does the behavior of a model start to substantially diverge from the behavior of the system it is representing? Formally, Rosen defines a model bifurcation using ε and δ neighborhoods, but for our purposes it is sufficient to say that two models diverge.

In the RT-110 report we included, the example depicted in Figure 19. Suppose that the solid black line is the "true" system that we are trying to model. Suppose also that we have two models of the system that we can use, model 1 and model 2. Model 1 is represented by the dashed line and model 2 is represented by the dotted line. Note that on left portion of the diagram, model 1 is the better representation of the system while on the right side of the diagram, model 2 is the better representation. There is a region of overlap in the middle where neither model 1 nor model 2 is perfect, but together they bound the true system. The two vertical lines indicate the model bifurcation points. The left bifurcation point is where model 2 diverges from the true system while the right bifurcation point is where model 1 diverges.

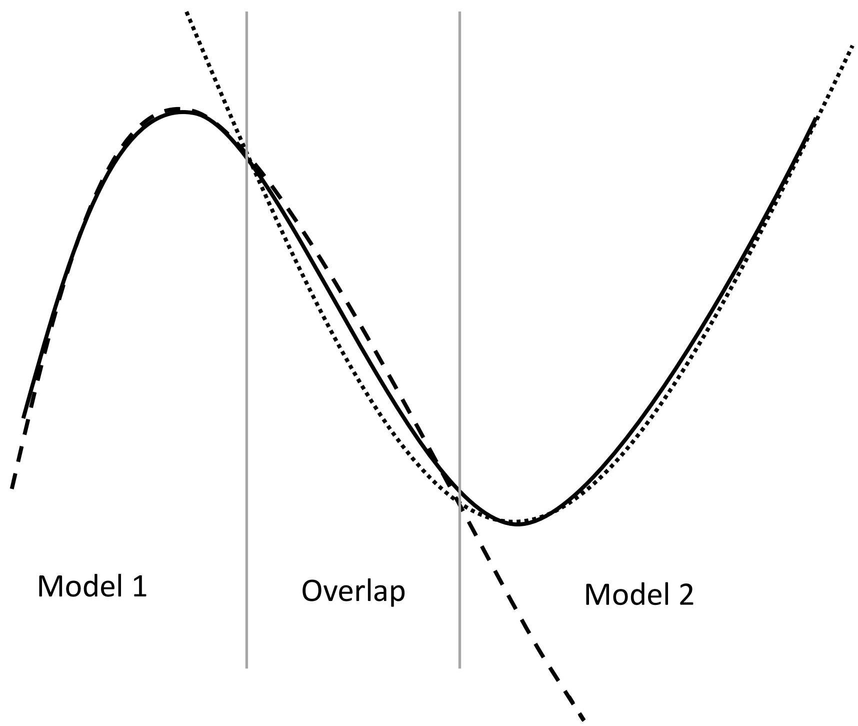


Figure 19 - Notional representation of models bifurcating from the true system

Perhaps the most important observation we can make about this example is that if we were to look at only model 1 alone or only model 2 alone, we would not detect the bifurcation points. Consequently, if tried to capture the model form error of model 1 (or model 2) via extrapolation from its area of fit, we would substantially underestimate it. If we were to look at both models 1 and 2 simultaneously we would see that the models bifurcate relative to each other, but we would not know which model is the better representation of the true system without additional information. This is the essence of the model uncertainty problem. We do not always know when our model is no longer valid since we cannot not find bifurcation points by simply running the model or performing a sensitivity analysis.

A more descriptive and intuitive term for a bifurcation is "phase shift." A bifurcation point is where the system of interest undergoes a phase shift, and the model we were using is no longer accurate. The canonical illustrative example of this principle seems to be the phases of water. For instance, we can model ice as a rigid solid...until the temperature gets above 0° C. At that temperature (assuming standard atmospheric pressure) water under goes a phase shift from solid to liquid. So our model of water as a rigid solid is no longer valid. Thus, the melting point of water is a bifurcation point. In fact, one could make the argument that model switching is how we define phases.

Solé (2011) discusses modeling the phases of water in great detail. Since we are using phase shifts as a metaphor, for our purposes, it will suffice to proceed with a highly simplified version. (We will not concern ourselves with supercritical fluids, etc.) We can extend our water example by continuing to increase temperature. Once we get to 100° C, water undergoes another phase shift from liquid to gas. This necessitates another model shift. Under a scenario of gradually increasing temperature we have to use three different, incompatible models of water: rigid solid, liquid, and gas. There are also two transition zones where the water is a mixture of two phases, and it is not clear what model we should use.

Empirically, we know the bifurcation points of water, but these would not necessarily be evident using only models for a solid, a liquid, and a gas. For example, a standard model of a rigid solid may not include temperature as a parameter at all. If we were to model water vapor using the ideal gas law, reducing the temperature would certainly not reveal the condensation point. Thus, the problem is that if these three models were all we had, we may not ever find the bifurcation points. Thus, we would not know when the models are wrong. And this does not even factor in the transition zones where none of the three models are accurate.

So how might we address this problem? The standard approach in science is reduction. We attempt to drill down a layer of abstraction and capture the bifurcation points. Of course, it almost goes without saying that the computational and data collection burdens increase rapidly as we move toward quantum mechanics. At that point, modeling a macroscopic system is computationally prohibitive. As a practical matter, all bifurcation points cannot be fully captured for most real world systems with a single model. The question then becomes whether or not one crosses a bifurcation point for the situation and model of interest.

For enterprise systems, the phase change issue is much more challenging than the water example. As we discussed in the previous sections, enterprise systems undergo multiple phase shifts at multiple layers of abstraction, and it is effectively impossible to represent them with a single model. Consequently, our only course of action is to try to identify the existence of the phase shifts or bifurcation points so that we can deal with them.

Returning to the water example above, knowing where the bifurcation points are allows us to know when to switch models. The problem is that we do not necessarily know the analog of melting point for a social system. In fact, social systems can be so complex that we have no idea which bifurcation points to look for when we pose a question of interest. So how might we go about identifying phase shifts or bifurcation points?

Expansion of Epistemic Uncertainty for Model Risk

Based on the discussion to this point, it seems that it is necessary to consider more sources of epistemic uncertainty than those typically considered by uncertainty quantification approaches when we attempt model enterprise systems. Roy and Oberkampf (2011) considered model form uncertainty, but for our purposes, we need to decompose this a bit more. To that end, Figure 20 organizes typical sources of uncertainty when modeling enterprise systems, where we assert that the lower rows become increasingly prominent as we move from modeling pure physical to enterprise systems.

For each source of uncertainty in the table, we provide a representative model structure, describe what varies, and a real life modeling example where the uncertainty is relevant. We should note that the representative model structures are deliberately simple to facilitate communication. We do not mean to suggest that these structures would be used in the examples.

Uncertainty	Model Structure	What Varies	Example
Deterministic	$y = ax + b$	N/A	Newton's Laws
Aleatory	$y = aX + b$ $X \sim \mathrm{N}(\mu, \sigma)$	X	Monte Carlo Financial Modeling
Epistemic - UQ	$y = aX + b$ $X \sim \mathrm{N}(\mu, \sigma)$	a, b, μ, σ	Robust Design Optimization
Epistemic – Phase shift	$y = \begin{cases} ax + b & t < t' \\ ax^2 + b & t \geq t' \end{cases}$	t'	Phase transition
Epistemic – Model Structure	$ax + b < y < ax^2 + b$	Model	Weather Modeling
Epistemic – Model Ontology	$y \approx ax + b$ $y \approx ct - du$	Ontology	Traffic Modeling

Figure 20 - Taxonomy of model uncertainty for enterprise systems

In the first row of the table, we start with deterministic modeling as a base case. The assumption, at least, is that there is no uncertainty in the system and a good example would be the application of Newton's laws. The next natural extension is to introduce random variables governed by probability distributions. As noted earlier, this is referred to as aleatory uncertainty. An example would be when businesses perform Monte Carlo simulations where cash flows are uncertain and governed by probability distributions.

The deterministic and aleatory cases are the two “classical” cases, and there are many optimization approaches to find or at least approximate the “best” solution when using such models. But what if we either do not have a probability distribution for an uncertain model parameter or we are not certain of that probability distribution? Then we move into the domain typically called uncertainty quantification in the literature. Methods in this domain vary model parameters and probability distributions. In the simplest form, this can be a basic sensitivity analysis. However, as noted in the literature review, more sophisticated techniques may be applied. Searching for optimal solutions relative to this level of uncertainty is certainly more challenging, but there is much ongoing work in the areas of robust design and MDO to address these challenges.

But what happens if we are uncertain regarding the structure of the model itself? This brings us into the area of bifurcation. The first case we will consider is where there are known models but the bifurcation point is uncertain as defined by a set of control parameters. For example, in our phase change example, we may be uncertain about a particular substance's melting point. An example of this in an enterprise system is when an organization operates in different modes (for example emergency response vs normal), but it is not entirely certain when a model shift will be triggered.

If we increase the level of epistemic uncertainty further, we may not be sure of the model structure at all. In the first case, we may more or less know the model ontology, but just not be sure about the relationships between the entities. However, we may be able to capture the spread of possibilities by comparing different possible model structures. An example of this would be weather modeling where it is common to run an ensemble of models. Each is based on the same basic physics of weather, but each emphasizes different aspects, employs different levels of resolution, etc. Finally, we may not even be sure what the best ontology is to apply to the question. An example of this would be predicting traffic jam density using either a flow based or an agent-based modeling approach. (Daganzo et al. 2011)

It is important to note that as we move from physical systems to enterprise systems we tend move down the table with increasing epistemic uncertainty. For example, the laws of physics are relatively stable, and for most practical applications we have no uncertainty about their applicability (though there are exceptions). However, if we were to build a model of a social organization, we would not be very confident in its structure over time. (Members may come and go. The organization may reorganize, etc.). Unfortunately, such uncertainties (phase, structural, ontological) are challenging to quantify or optimize over for enterprise systems. To illustrate this point, we will consider an example problem in the next section.

EXPLORATORY DECISION PROBLEM: CROP ALLOCATION

As discussed previously, the general MDO/uncertainty quantification paradigm attempts to capture the various sources of uncertainty within the model so that an optimization/search strategy may be employed to find an acceptable balance of performance and uncertainty. Consequently, if we were to apply this approach to an enterprise decision problem, we would have to identify all of the potentially relevant sources of uncertainty and capture them in the model (at minimum within the outer loop). In this section will consider a very simple decision faced by a farmer, and show how consideration of the epistemic uncertainties that result from the larger enterprise system in which the farmer is embedded quickly overwhelms one's ability to account for and manage that epistemic uncertainty within the model.

A key choice made by farmers across the agricultural landscape is to determine the mix of crops to produce in a particular growing season. There are a number of risk factors involved in the process that could have a detrimental effect on crop yield and revenue. For example, drought, production mishaps, demand fluctuation, and competitor supply levels can all affect the profit that may be made from a given crop. In essence, when the farmer makes an operational decision such as a crop mixture to plant, ideally, he or she should consider the impact of economic factors, market factors, and the physical environment just to name a few. (See Figure 21) To aid in determining the crop mixture for a given year, it would be helpful to develop a mathematical model of revenue under different crop mixtures.

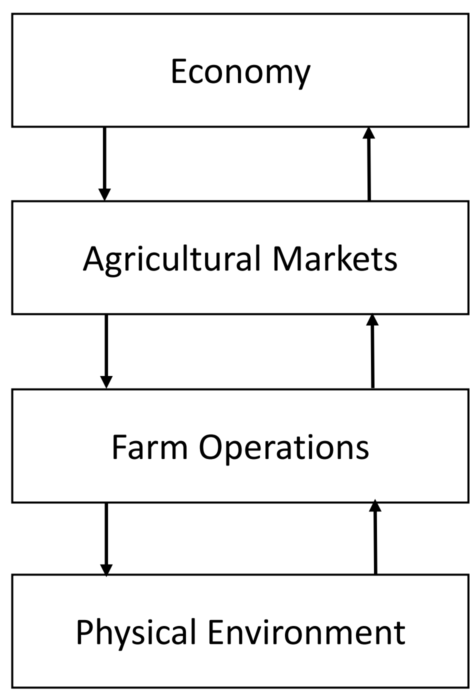


Figure 21 - Abstractions relevant to the farmer's decision problem

The obvious approach is to develop models that estimate the year's crop yield and crop price. These estimates can be used to determine revenues under different crop mixtures, and the farmer can compare these values to arrive at a preferred crop mixture. In this example we will start with the farmer's core operational decision then attempt to apply increasingly sophisticated methods to capture uncertainties that result from the agricultural market level and the economic level. We will find that we very quickly get overwhelmed before we even get to the uncertainties that result from the physical environment. In effect, following a reductionist approach in this example that increases model fidelity in an effort to both maintain consistency among the views while simultaneously capturing the sources of uncertainty can open up the possibility for different types of model uncertainty, showing the need for a balance between the realism of the model and the model uncertainty made possible by the model's exactness.

For simplicity, we assume that there are only two crops that can be planted, and that the proportion of one has no effect on the growth of the other. To balance revenue and risk, a simple mean-variance model of the form $E[X] - \lambda Var[X]$ is used as the objective function.

DETERMINISTIC/NAIVE MODEL

At the most basic level, we could ignore the many risk factors associated with the problem and model revenue simply as the product of point estimates of crop yield and crop price. One approach to estimating yield is to use the mean yield of that crop over the past several years. The price can be estimated by using the current price in the futures market (with settlement date closest to harvest time).

In particular, let x_i denote the price of crop i at harvest and y_i denote the yield of crop i. Let δ be the proportion of crop 1 planted, where $0 \leq \delta \leq 1$ and R_δ be the farm's revenue under proportioning δ, i.e., $R_\delta = \delta x_1 y_1 + (1-\delta) x_2 y_2$. Because the estimates for price and yield are non-stochastic, the variance is zero, and so the mean-variance formulation simplifies to a revenue maximization problem: $\max_\delta R_\delta = \delta x_1 y_1 + (1-\delta) x_2 y_2$.

In this simple case the solution is immediate: if $x_1 y_1 > x_2 y_2$, then $\delta = 1$, otherwise $\delta = 0$. The obvious drawback of this deterministic approach is that it makes no attempt at managing the potential variation in revenue. While this model is exceedingly simple to use and understand, attempts to optimize with this model are quite likely to produce an incorrect result, and decisions made using such a result will be problematic.

MODEL WITH ALEATORY UNCERTAINTY

To make the model closer to reality, we now assume that crop prices are random variables. To this end, suppose that the price of crop i is a random variable X_i with realizations x_i, and that the crop yield y_i is given a point estimate based on past yields, as above. The mean-variance model is then $MV(\delta) = E[R_\delta] - \lambda Var[R_\delta]$, where

$$\begin{aligned}
\mathrm{E}\ [R_\delta] &= \delta y_1 \mathrm{E}\ [X_1] + (1-\delta) y_2 \mathrm{E}\ [X_2] \\
\mathrm{Var}\ [R_\delta] &= \mathrm{Var}\ \ [\delta y_1 X_1 + (1-\delta) y_2 X_2] \\
&= \delta^2 y_1^2 \mathrm{Var}\ [X_1] + (1-\delta)^2 y_2^2 \mathrm{Var}\ [X_2] \\
&+ 2\delta(1-\delta) y_1 y_2 \mathrm{Cov}\ [X_1, X_2]
\end{aligned}$$

Differentiating with respect to δ and setting the derivative to 0, we obtain the first-order optimality condition

Equation 5

$$\delta = \frac{y_2^2 \text{Var}\ [X_2] - y_1 y_2 \text{Cov}\ [X_1, X_2]}{\text{Var}\ [y_1 X_1 - y_2 X_2]} + \frac{y_1 \text{E}\ X_1 - y_2 \text{E}\ [X_2]}{2\lambda\ \text{Var}\ [y_1 X_1 - y_2 X_2]}$$

Also, we observe that the second derivative, $\frac{\partial^2}{\partial^2 \delta} MV(\delta) = -2\lambda Var[y_1 X_1 - y_2 X_2] < 0$, and therefore (Equation 5) gives a global maximum.[1]

Although this approach to the problem seems to be more reasonable than the naive approach, we have introduced parametric uncertainty to the situation, as we are now faced both with choosing a distribution and parameterizing it.

MODEL WITH EPISTEMIC UNCERTAINTY

A more detailed model may make a more serious attempt at forecasting yield than merely relying on past averages. A common method of modeling crop yield utilizes statistical methods such as regression analysis or principle component analysis (Kantanantha et al. 2010). Such models typically operate on the same or similar ontologies, largely agreeing on the type of parameters used and their relationships.

We consider a simple method where linear regression analysis is used to parameterize the distribution of yield. For example, if the probability distribution calls for two parameters μ and σ, past data could be used to develop models $\sum_{i=1}^{n} \beta_i x_i$ and $\sum_{i=1}^{n} \gamma_i y_i$ which give estimates for μ and σ, respectively. We thus let both price and yield be random variables; price defined as above, and yield for crop i modeled by the random variable Y_i and parameterized by the above regressions.

As a further simplification, we assume pairwise independence among all variables. This allows the modeler to assign distributions to each phenomenon without concerns of describing how the variables relate to each other. The mean and variance simplify to

Equation 6

---

[1] Second order derivatives are not explicitly given in the remainder of this example, but we note that they are all of the form $c \cdot Var[Z]$, where $c < 0$ and Z is a random variable. Consequently, all first-order conditions describe maximal solutions.

$$
\begin{aligned}
\mathrm{E}\,[R_\delta] &= \delta \mathrm{E}\,[X_1 Y_1] + (1-\delta)\mathrm{E}\,[X_2 Y_2] \\
&= \delta \mathrm{E}\,[X_1]\mathrm{E}\,[Y_1] + (1-\delta)\mathrm{E}\,[X_2]\mathrm{E}\,[Y_2] \\
\mathrm{Var}\,[R_\delta] &= \mathrm{Var}\,\,[\delta X_1 Y_1 + (1-\delta) X_2 Y_2] \\
&= \delta^2 \mathrm{Var}\,[X_1 Y_1] + (1-\delta)^2 \mathrm{Var}\,[X_2 Y_2]
\end{aligned}
$$

Forming the mean-variance, differentiating, and setting equal to 0, we obtain the first-order condition

Equation 7

$$
\delta = \frac{\mathrm{E}\,[X_1]\mathrm{E}\,[Y_1] - \mathrm{E}\,[X_2]\mathrm{E}\,[Y_2]}{2\lambda\,\mathrm{Var}\,[X_1 Y_1 - X_2 Y_2]} + \frac{\mathrm{Var}\,[X_2 Y_2]}{\mathrm{Var}\,[X_1 Y_1 - X_2 Y_2]}
$$

Using these equations, the optimal δ can easily be calculated from Equation 7. For example, suppose that X_i and Y_i are lognormally distributed as follows: $X_1 \sim lnN(2.5, 0.5), X_2 \sim lnN(2, 0.25), Y_1 \sim lnN(1, 0.5)$, and $Y_2 \sim lnN(1.5, 1)$. Using Equation 7, we find the optimal mixture to be 83.67% crop 1 and 16.33% crop 2.

Although this framework appears to be superior to the previous models, there is still a great deal of uncertainty in the model, and much of this uncertainty has been induced by complicating the models. The use of a regression model introduces both quantification uncertainty and epistemic uncertainty of the model structure type, as the regression model faces its own parameter uncertainty, and additionally, it will never be clear that the chosen regression model (e.g., a simple linear regression model) is the “correct” model. Perhaps a nonlinear regression model would be more appropriate or perhaps a more abstract method such as principle components analysis would be preferable.

Additionally, we are now multiplying uncertainties. At the very least, at each decision point, there exists the potential for a binary phase transition, with small errors possibly propagating.

To move the model closer to reality, we might drop the assumption of independence between X_1 and X_2 and between Y_1 and Y_2, while maintaining the independence between crop price and yield. The expected revenue $E[R_\delta]$ is still given by Equation 6, but the variance is now

$$
\begin{aligned}
\mathrm{Var}\,[R_\delta] &= \mathrm{Var}\,\,[\delta X_1 Y_1 + (1-\delta) X_2 Y_2] \\
&= \delta^2 \mathrm{Var}\,[X_1 Y_1] + (1-\delta)^2 \mathrm{Var}\,[X_2 Y_2] \\
&\quad + 2\delta(1-\delta)\mathrm{Cov}\,[X_1 Y_1, X_2 Y_2].
\end{aligned}
$$

Forming the mean-variance, differentiating, and setting equal to 0, we obtain the first-order condition

$$\delta = \frac{\mathrm{E}\,[X_1]\mathrm{E}\,[Y_1] - \mathrm{E}\,[X_2]\mathrm{E}\,[Y_2]}{2\lambda\,\mathrm{Var}\,[X_1Y_1 - X_2Y_2]} + \frac{\mathrm{Var}\,[X_2Y_2] - \mathrm{Cov}\,[X_1Y_1, X_2Y_2]}{\mathrm{Var}\,[X_1Y_1 - X_2Y_2]}.$$

While this is similar to the previous δ, it is complicated by the covariance term (which also appears in the expansion of the variance term outside of the parentheses). The covariance is given by

Equation 8

$$\begin{aligned}\mathrm{Cov}\,[X_1Y_1, X_2Y_2] &= \mathrm{E}\,[X_1Y_1X_2Y_2] - \mathrm{E}\,[X_1Y_1]\mathrm{E}\,[X_2Y_2] \\ &= \mathrm{E}\,[X_1Y_1X_2Y_2] - \mathrm{E}\,[X_1]\mathrm{E}\,[Y_1]\mathrm{E}\,[X_2]\mathrm{E}\,[Y_2]\end{aligned}$$

The second term in Equation 8 is easy to compute, since the individual expectations of price and yield are assumed to be given, but the first term is difficult, and must be estimated. We see that, again, we have added uncertainty in the process of providing a more detailed model.

Continuing the numerical example from above, Table 8 gives the value of δ for different levels of covariance (which is uniquely determined by $E[X_1Y_1X_2Y_2]$).

Table 8 - Optimal δ for various values of $Cov[X_1Y_1, X_2Y_2]$

$\mathrm{Cov}\,(X_1Y_1, X_2Y_2)$	1000	800	600	400	200	0	-200	-400	-600	-800	-1000
δ	0.97	0.93	0.90	0.88	0.86	0.84	0.82	0.80	0.79	0.78	0.76

Depending on the magnitude of the covariance, the optimal crop mixture can be quite different from the mixture found by making the independence assumption. In particular, there is a point where the difference between the estimated mixture and the optimal mixture is significant, and this point defines a type of phase transition, where the simplified model breaks down and gives a poor prescription.

We now consider the implications of relaxing the assumption that crop price and yield are independent, presumably making the model even more accurate. The first order optimality condition becomes

$$\delta = \frac{\mathrm{E}\,[X_1Y_1] - \mathrm{E}\,[X_2Y_2]}{2\lambda\,\mathrm{Var}\,[X_1Y_1 - X_2Y_2]} + \frac{\mathrm{Var}\,[X_2Y_2] - \mathrm{Cov}\,[X_1Y_1, X_2Y_2]}{\mathrm{Var}\,[X_1Y_1 - X_2Y_2]}$$

Where

$$\begin{aligned}\mathrm{E}\,[X_iY_i] &= \mathrm{E}\,[X_i]\mathrm{E}\,[Y_1] + \mathrm{Cov}\,[X_i, Y_i] \\ \mathrm{Var}\,[X_iY_i] &= \mathrm{Cov}\,[X_i^2, Y_i^2] + \mathrm{Var}\,[X_i^2]\mathrm{Var}\,[Y_i^2] - (\mathrm{E}\,[X_iY_i])^2\end{aligned}$$

As a consequence of increasing the detail of the model, we encounter a significant increase in uncertainty. Indeed, we must now determine values for $E[X_i Y_i]$ and $Cov[X_i^2, Y_i^2]$ in addition to $E[X_1 Y_1 X_2 Y_2]$ from above. The potential for an incorrect estimate is thus quite high, and the likelihood of error is clearly greater in this case. This serves to amplify the possible impacts of the uncertainty types discussed above.

ALTERNATE METHODOLOGIES

The models from the previous sections generally shared a common ontology and were closely related. However, there are several alternate methodologies that might be used to model crop revenue.

- In the above methodology we considered the issue more or less in a vacuum, with the farmer in question having no impact on the market and the market having no impact on the farmer. For example, if every farmer in the area follows a similar strategy based on the same analysis, then the market may actually be different than anticipated, thus negating the individually focused models that were used in the first place.

- The above model did not consider the employment of risk management strategies (e.g., hedging and crop insurance). The use of such instrument could have a material effect on revenue, particularly when governmentally subsidized insurance is available.

- Instead of giving separate models for price and yield, a single model of revenue might be used which does not explicitly consider these factors.

The existence of these alternatives shows how epistemic uncertainty of the model ontology variety enters into the crop mixture problem.

At this point, the natural question to ask is what should be done in light of the uncertainty pointed out here. This simple analysis showed how quickly the crop mixture problem can become unwieldy if the modeler attempts to closely develop the intricacies of the problem. In reality, it is difficult to imagine that such an analysis is actually undertaken. We argue that there is good reason for this, as we have shown that these models would be fraught with assumptions and potential uncertainty, and optimizing over such an unstable space seems not to be worth the trouble. Adding models of weather and crop growth to the picture to account for that environmental impacts are unlikely to improve the situation.

As noted by Oberkampf, et al. (2002), “A model of limited, but known, applicability is often more useful than a more complete model. This dictum of engineering seems to be forgotten today with the advent of rapidly increasing computing power.” Consequently, the farmer may do well to consider simpler models and setting up adaption or hedging mechanisms to address the variability involved in the situation. Such mechanisms happen to be easily attainable in the agricultural market.

Epistemological Implications of the Crop Allocation Example

The takeaway from the farmer example is clear: the excessive amount of epistemic uncertainty makes the detailed decomposition and analysis of that uncertainty counterproductive. A real farmer would probably just use a simple decision rule and hedge the position in the futures market. While it is unlikely that his or her hedging position will be “optimal”, the approach seems reasonable considering the amount of effort required to find an optimal position. This approach effectively handles the uncertainty in bulk in a way that is likely to be good enough for practical purposes. This raises important epistemological questions regarding the use of multi-models to support decision making for enterprise systems versus physical systems.

Instinctively, we might think that a model that attempts to faithfully and fully implement every valid scientific theory relevant to the question of interest is more justifiable than one that makes substantial omissions and/or false but useful assumptions. Yet, the farmer example seems to suggest the opposite conclusion: we may be able to learn more from a coarse model where the modeler knowingly employs inaccurate yet useful assumptions than the theoretically comprehensive but unwieldy model. How could this be the case?

To understand this, we need to consider work on the epistemology of simulation. Particularly relevant to our problem are works by Andreas Tolk (2015) and Eric Winsberg (2010). Both consider how we can reach correct conclusions from simulations that we know are incorrect in some way. Tolk notes that if we follow the evolution of science, we see a similar problem. Over time theories became increasingly abstract and removed from everyday observations. In the end, a scientific theory is designed to answer a specific set of questions over specific set of conditions. There is no guarantee that it will be correct in all situations. In fact, we know it is not.

Also relevant to this discussion is Winsberg's examination of how multi-scale modeling is used to model crack propagation in a material. The issue is that depending on the scale, different models are required. For large scales, linear-elastic theory is used; for medium scales, molecular dynamics is used; and for small scales, quantum mechanics is used. The problem is that these three theories are inconsistent and incompatible. To make the simulation work, ``handshaking algorithms'' that require deliberate fictions must be introduced to translate parameter values back and forth among the three views.

For instance, fictitious “silogen” atoms are introduced on the boundary between the molecular dynamics view and the quantum mechanical view. Despite these fictions, the simulation can produce accurate results. Winsberg argues that the credibility of such approaches is established in very much the same way that credibility is established for a traditional experimental setup, through an evolutionary path that has been successful over time.

In short, consideration of some problems intrinsically requires multiple incompatible theories. These incompatibilities result from the fact that theories are themselves simplifications of reality. To accommodate these incompatibilities, we may be forced introduce fictitious entities to mediate between them, particularly when they are modeled in parallel. Unfortunately, development of “handshaking algorithms” would seem to require a great deal of trial and error. When we consider essentially physical systems such as an aircraft where the laws of physics are dominant, all else being equal, building a detailed multi-model and using uncertainty quantification coupled with robust optimization and related techniques seems to be a worthwhile approach. This is because the laws of physics are stable (i.e., they do not bifurcate over most questions of interest). For example, Newton's laws are still used for many practical engineering problems because they do not bifurcate over the space of interest. Colloquially, the rules of the game don't change in the middle of the game.

When we consider enterprise systems on the other hand, bifurcations are rampant. Human beings frequently switch modes of behavior depending on the circumstances and/or invent entirely new ontologies to address problems and adapt to new situations. In other words, they intentionally alter the rules of the game as the game proceeds. One could argue that this capability has contributed to our success as a species. On the other hand, it makes it incredibly difficult to model and predict our behavior. Consequently, constructing a detailed multi-model of an enterprise system and performing an expensive uncertainty analysis on it does not seem to pass the cost/benefit threshold. The more detailed the model, the more ways for it to be wrong.

So does this mean that one should not bother building multi-models for enterprise systems? Not at all. Rather, consideration of the issues discussed has implications for both how one constructs such and models and how one uses them. We need to capture are the bifurcation points so that we can manage the risk of crossing them. Of course, if the models in use do not contain the structure that would reveal the bifurcation, then increasing the fidelity of those models will not help. This shifts the emphasis of the multi-modeling effort from one of maximizing local predictive accuracy to one of detecting the existence of bifurcation points in the system. In other words, all else being equal, we would rather have an imprecise model that is viable over a large range of the decision space because it captures the bifurcations than an extremely precise model that has limited applicability because it does not capture the bifurcation points.

From this perspective, the point of multi-models of enterprise systems is to identify the existence of bifurcation points that the decision maker would not otherwise know about if he or she only focused on a single model. In essence, we looking for the existence of categories or modes of system behavior even if we cannot precisely and quantitatively demarcate the transitions between them.

Revisiting the Toll Road Example

With this perspective in mind, let us revisit the dynamically tolled road. How would the discussed concepts apply to this decision problem? In essence, we are asking: was there a way for engineers and decision makers to detect the potential for a bifurcation in driver behavior in the modeling and analysis phase prior to the expenditure of substantial funds? In that spirit, we want to perform the analysis using the same resources and expertise available to a typical engineer or analyst. This means that we need to be able to detect the bifurcation point without assuming that the analyst has a PhD in economics as well as every other relevant discipline. Following the approach advocated in this paper, we decided to develop a quick, low-fidelity model of the dynamically tolled road decision problem. Because we only need to establish the possibility for a bifurcation, not prove its existence, we can be a little loose in our modeling assumptions.

For this example, let us view it from the perspective of a hypothetical analyst who is faced with making a recommendation on whether or not to construct a dynamically tolled road. Simulating traffic flow itself is well established. Thus, the key issue here is the driver's decision whether or not to enter the dynamically tolled road. The whole point of the toll is to influence this decision. How should this be modeled? As previously discussed, the most natural approach would be to implement a classical demand function where an increase in the toll would result in a decrease in the number of drivers willing to enter the toll road. Potential drivers could be surveyed to assess their willingness to pay to save travel time, and the demand curve could be parameterized (This is essentially how it is done in real life, see (HNTB 2010)).

Of course, there would be uncertainty regarding this demand curve as the survey would not be completely reliable. So one could perform a sensitivity analysis on the demand curve and observe the impact on traffic flow in the simulation. If a more sophisticated analysis is desired, one could use one of uncertainty quantification techniques described previously. The results of this analysis would constitute a fairly conventional approach to support decision making. We would have a nominal understanding of the toll road's impact on traffic flow plus some understanding of the robustness of that result. But would it be correct? From a theoretical perspective, this analysis would seem to be sound.

As a first step to test this, we will intentionally alter the structure of decision problem (as opposed to just the parameters) to see how stable to the result is. Our first instinct might be to increase the fidelity of the driver decision model. One way to do this would be to employ the findings of behavioral economics and see if that changes driver response. Classical economics is based on decision makers being rational utility maximizers. However, real decision makers tend to be more focused on gains and losses relative to a reference point as opposed to the absolute value of a metric of interest. In particular, they tend to exhibit strong loss avoidance. We hypothesized that perhaps when drivers are surveyed, they view the use of the toll road as a gain in time, but when they are actually driving, they view use of the toll road as the avoidance of a loss. Thus, the toll rates set based on the initial survey might be too low to deter entry. This behavior is captured via prospect theory, which replaces a conventional utility function with a value function centered on a reference point (Kahneman and Tversky 1979).

To consider this hypothesis, we developed a Monte Carlo simulation of driver decision making. We generated a synthetic population of 1000 drivers, each with a different time value of money sampled from a lognormal distribution. To parameterize the distribution, we obtained national hourly wage data from the US Bureau of Labor Statistics (NBLS 2015) We assumed that each driver's time value of money was equivalent to his or her randomly sampled wage. We then generated a prospect theory style value function for money with an intentionally exaggerated penalty for lost money. For the sake of simplicity, we assumed that each driver had the same value function and relied on each driver's time value of money for differentiation.

To generate the demand curve, we used the value function to assess each driver's value for each option: enter the toll road or remain on the untolled road. We then computed these values over different toll levels and different anticipated delays on the untolled lanes. In short, if the value of taking the toll road exceeds the value of remaining on the untolled road, the driver will choose to enter the toll road. When these decisions are aggregated over the entire simulated population, we obtain a demand curve for the toll road that can be recomputed based on the anticipated difference in travel time between the tolled and untolled roads.

We generated demand curves for two different reference points: first, when drivers viewed switching to the toll road as a gain in time, and second when the viewed the switching to the toll road as avoiding a loss in time. Interestingly, the two scenarios resulted in identical demand curves, one of which is depicted in Figure 22. In retrospect, this should not have been surprising. While the two scenarios produced very different values for each driver, the order of preference did not change for any given driver. In essence, the increase in the fidelity did not have an impact on the outcome. Unfortunately, this means that this approach failed from the stand point of finding anomalous driver behavior. In other words, if this were all we had done, we would not have found the bifurcation point.

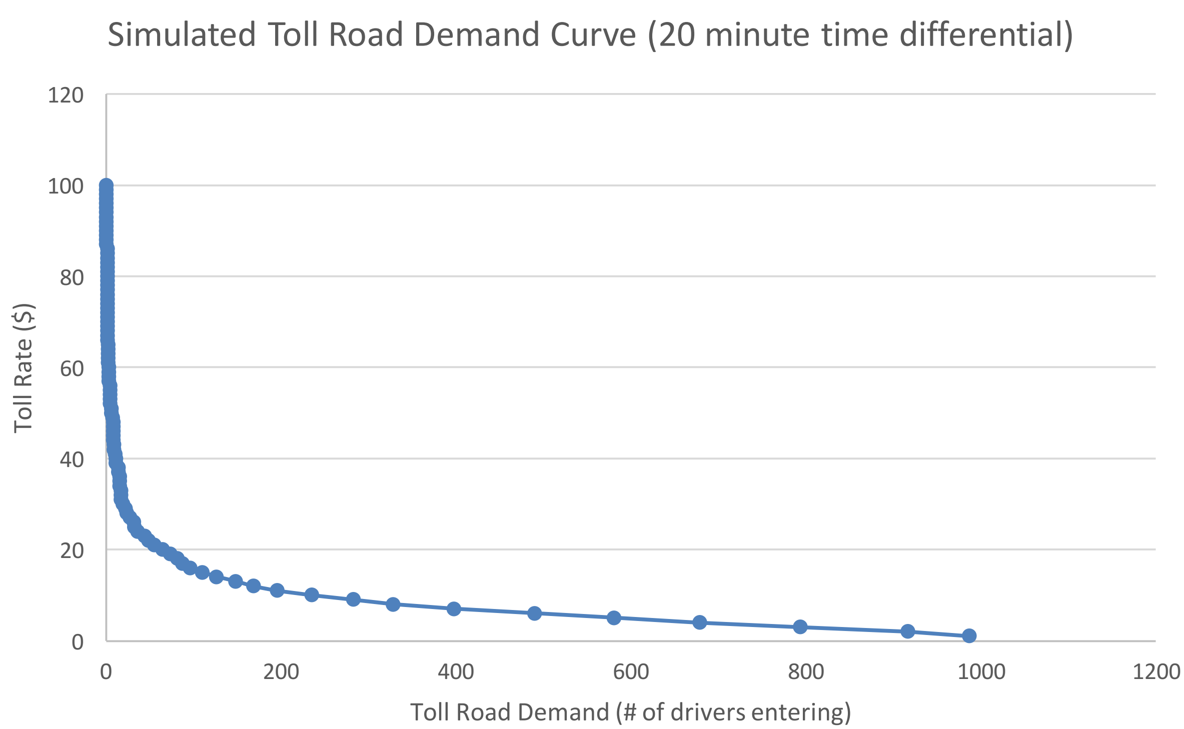


Figure 22 - Simulated demand curve generated using an assumed 20 minute difference in travel time between the tolled and untolled roads

Of course, the behavioral economic view is not the only way we could alter the classical demand model. Information economics is concerned with incomplete information, signaling, and their impact on markets. In the case of the dynamically tolled roads, as was previously suggested, the current toll could serve a signal to drivers. Drivers might assume that the higher the toll, the more congestion there is on the untolled road. This effectively creates a two-way coupling between the driver decision making model and the traffic model. Meaning for any given toll level, the driver would mentally translate that into an expected increase in travel time if he or she decides to remain on the untolled road. Before creating a very detailed model of this coupling, we want to determine is if it is even relevant (i.e., is there even the possibility of revealing a bifurcation point?). Consequently, we modified our simulation by replacing the delay parameter with a function that mapped the current toll level to the driver's perceived delay. Intuitively, we expected this relationship to be convex. So we postulated a function such that the driver expects the toll to increase exponentially as the delay on the untolled road increases. The result was the demand curved shown in Figure 23.

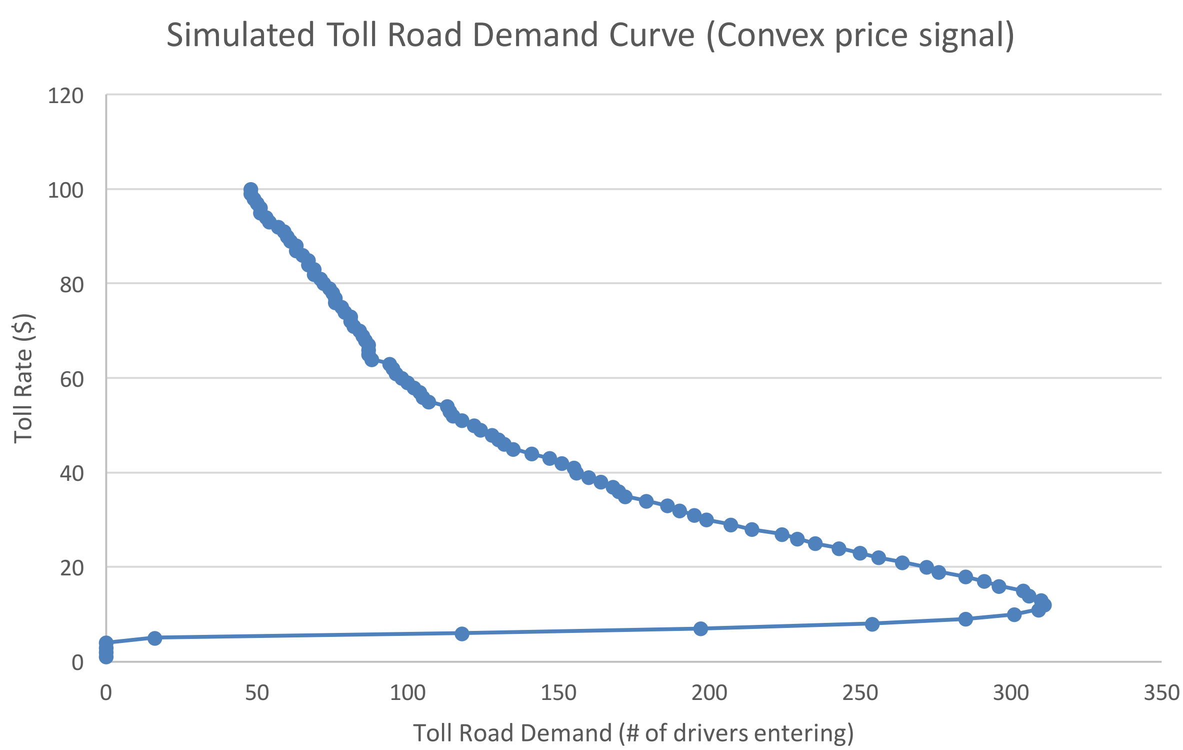


Figure 23 - Simulated demand curve generated using an assumed convex relationship between the expected delay and the toll level

Note that we immediately see the bifurcation that we were looking for. At the high toll levels, the toll serves as a deterrent to entry and we see a classical demand curve. However, at the lower toll levels, the demand curve reverses directions, and we see that more drivers enter the toll road as the toll increases. This non-classical demand response is qualitatively consistent with what was observed in the Janson and Levinson (2014) study in Minneapolis.

Actually, there is a second bifurcation point, though it is harder to see. At very low, but not zero, toll levels the demand drops to zero. In essence, drivers conclude from the low toll that congestion must be so low on the untolled road that no one thinks it is worth paying the toll.

From here, it is possible to run many additional excursions, for example changing the driver's belief in the relationship between the toll level and congestion to be say linear or concave. The former results in a vertical demand curve (i.e. fixed demand), and the latter results in the opposite response to the convex relationship. Both cases seem implausible in the real world. But really, such excursions are unnecessary for the purposes of this example. Our objective in this exercise was simply to detect potential bifurcations in the assumed driver behavior, and this was accomplished with an extremely rough, low-fidelity simulation.

The most obvious approach from here would be to fully couple the driver decision model with the traffic model, where the driver views the toll as a signal for price. Even it if were not entirely precise, it would give the analyst a feel for how the bifurcation points in driver behavior affect the traffic flow on the roads. This could allow experimentation with different toll schedules, more targeted surveys of drivers to capture the effect, etc. Of course, we very quickly end up back in our fidelity trap. There are additional possible feedback loops that could come into play with this simulation. Drivers may adjust their perception of the delay versus the toll over time, commuting patterns may change, etc. How could we account for all of the possibilities? The reality is that we cannot.

However, simply being aware of the existence of the bifurcation point opens up other strategies beyond just trying to find the optimal dynamic tolling approach. For instance, tolls could be set on a fixed schedule based on typical demand and only updated a quarterly basis. This defeats the use of the toll as a signal for current traffic conditions. This is what is done for a toll road in Singapore that has been shown to be effective (Olszewski and Xie 2005). Another approach might be to provide estimated travel time for the toll road and non-tolled alternatives along with the current toll information. Simulations could be used to explore these policies, too. There may be other possibilities as well, but none of these would have been considered if we had simply proceeded with our initial model.

Discussion

Establishing the existence of a bifurcation point provides a great deal of value to a decision maker. He or she can decide whether or not to attempt to reduce the uncertainty surrounding the bifurcation by actions such as identifying early warning indicators, collecting additional data, conducting pilot tests, creating contingency plans, purchasing options, etc. If multi-modeling can provide a means to identify relevant bifurcation points, then it provides value to the decision maker.

Pennock and Rouse (2014, 2016) argued that when real decision makers are faced with the sort of epistemic uncertainties we have discussed in this paper, they employ some combination of four basic strategies: optimize, adapt, hedge, and accept. When epistemic uncertainty is low, it is reasonable for one to search for an optimal solution. Note that this includes circumstances where uncertainties are aleatory and it is safe to add summary risk measures such as variances to the objective function to achieve the desired balance between uncertainty and performance. Additionally, epistemic uncertainty in the form of parametric uncertainty may be manageable using UQ and MDO techniques. Finally, known and well understood phase shifts can even be handled via controlled model switching or increasing fidelity.

As the level of epistemic uncertainty increases in the form of uncertainty in phase, structure, and ontology, optimization becomes problematic for all of the reasons discussed in this paper. If the decision maker does not know the optimal solution, then he or she would prefer either to plan to adapt to changing circumstances or hedge the decision (i.e., invest resources to provide some form of insurance). From a modeling standpoint, supporting an adapt or a hedge strategy under substantial epistemic uncertainty can result in a very different approach than that employed for supporting an optimization strategy.

In short, the objective shifts from finding the global optimal solution to finding a reasonable local solution coupled with knowledge of the bifurcation points at which this solution becomes invalid. In essence, the bifurcation points constitute mode shifts in the sense that the decision maker would need to shift to a different mode of operation if the bifurcation point were crossed. Under these circumstances, it is beneficial to the decision maker if one can at least determine the existence of bifurcation points even if one does not know exactly where they are.

If the decision maker's situation is such that he or she can rapidly change policies or solutions, then he or she may choose to watch for impending or actual bifurcations in the real system during operations and adapt as necessary. Alternatively, if the decision maker does not believe he or she can respond quickly enough, he or she may choose to hedge the decision. (This occurs in financial markets all of the time.) In either case, the decision maker needs to be at least aware of the existence of the bifurcation points in order to take the appropriate action.

CONCLUSIONS

In this section we considered the challenge of managing epistemic uncertainty when building multi-models of enterprise systems for decision support. We concluded that developing very detailed multi-models in order to find optimal solutions is not likely to be cost effective because such approaches are quickly overwhelmed by epistemic uncertainty in the form of phase, structural, and ontological uncertainty. These sources of epistemic uncertainty are more manageable for physics-based systems, but the complexity of social systems makes them problematic for enterprise systems. This is in part because human beings intentionally learn and alter their behavior in response to interventions.

We characterized these sources of uncertainty as bifurcations, and defined the objective of multi-modeling of enterprise systems as the identification of the existence of these bifurcations. Once they are identified, decision makers can develop strategies to adapt or hedge in response even if they cannot be precisely quantified. Of course, this is not a panacea, and sometimes the only choice is tried and true empirical methods (e.g., experimentation and testing).

This brings us back to our motivating example, the dynamically tolled road. How could this decision dilemma have been handled differently? Acknowledging substantial hindsight bias, a simple multi-model that captured the larger driver decision context might have revealed that other factors could overwhelm the price as a deterrent to entry. The example revealed that this does not require an in depth understanding of economics. A coarse implementation of the concept may be sufficient. Of course, this begs the question of which structural and ontological variations a modeler should try. An intriguing possibility is that scientific sub-disciplines themselves could serve as a guide, since each likely evolved in response to encountered bifurcations in the system of interest. Consequently, it may be possible to guide modelers to the appropriate structural variations based on the circumstances of their question. Further investigation of this potential approach will be the subject of future work.

So does where does this leave us with regard to analyzing enterprise systems? Figure 24 summarizes the situation with regard to increasing epistemic uncertainty. Obviously, if there are no sudden structural changes in the system, then conventional engineering modeling is sufficient. Even when there are behavioral shifts, these are sometimes well understood and predictable. These types of shifts can be captured in a single model, and this is the domain of dynamical systems theory. What happens when we are aware of the existence of behavioral shifts, but they are not explainable or predictable with a single model? This is the domain of multi-modeling. The fact that there are multiple inconsistent models of a system is an indicator of experienced behavioral shifts. That is why there are multiple models. In a sense, these behavioral shifts are captured in the scientific knowledge base. Given the complexity of enterprise systems, we expect that we will often find ourselves in this circumstance when analyzing enterprise systems.

Of course, we can't expect multi-models to capture everything. There are still going to be behavioral shifts that have been experienced but not captured in models or not experienced at all. This brings us to the right side of Figure 24. In the cases where we know of the existence of a behavioral shift but do not have a good model to explain its behavior, we may be interested in early warning signals of impending shift. This is the domain of the statistical warning signals discussed in Section 5. Finally, if an event is a true unknown unknown, there is no way to predict or anticipate it. In that case, we have no choice but to monitor results and learn.

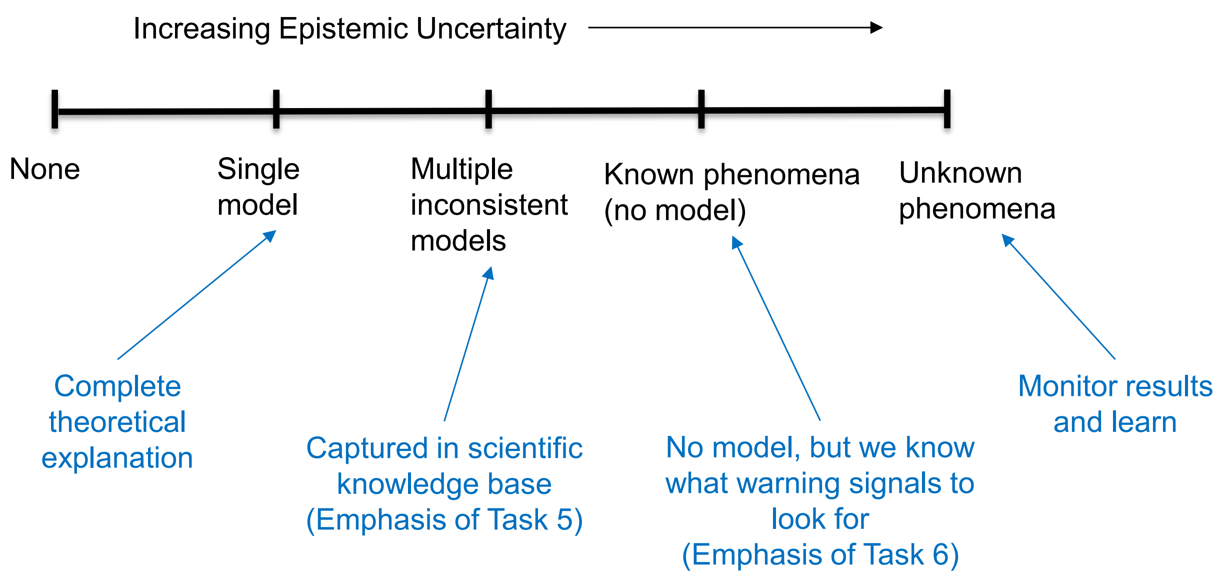


Figure 24 – Reponses to increasing epistemic uncertainty

Returning to the strategy framework (optimize, adapt, hedge, accept), we reach the conclusion that appropriate strategy is determined by the level of epistemic uncertainty. When the level of epistemic uncertainty is low (the first three rows of Figure 20), the best strategy is to optimize. The uncertainty can be managed with existing techniques for optimization under uncertainty. When the level of epistemic uncertainty is high (the second three rows of Figure 20), it is better to adapt or hedge. The uncertainty is driven by shifts in system structure that can invalidate a high-fidelity model and the resulting optimal solution. Consequently, under these circumstances it may be better to build a low-fidelity multi-model and search for the existence of bifurcations than a high-fidelity model to make "precise" predictions. In other words, it is better to be imprecisely right than precisely wrong.

7. Visualization Experiment

During RT-110, we considered how analysts and decision makers would use a model of an enterprise to diagnose situations and explore policy options. We discussed the use of visualization as mechanism to accomplish this. The concern we raised was whether the inherent complexity of enterprise problems would limit the ability draw inferences and whether it was possible to design interactive visual interfaces to mitigate this. We indicated the necessity of conducting an experiment to learn more about the ability of visual aiding to avoid the identification of spurious causal relationships under complex enterprise scenarios. This section describes the experiment we conducted under RT-138 to accomplish this.

BACKGROUND

Rouse, et al. (2016) conducted an enterprise diagnostic experiment in 2015 that applied Rasmussen's abstraction-aggregation hierarchy to historical cases from the US automotive industry. Each case dealt with the withdrawal of a brand from the market. These withdrawals were not attributable a single cause. Rather, they were the consequence of a complex interaction of factors. The intent of the experiment was to understand how subjects would use data and visualizations organized by level of abstraction to diagnose the factors that contributed to the car's withdrawal from the market. The experiment consisted of ten subjects drawn from Stevens students and faculty. Half of the subjects had a high-level of expertise with regard to the problem and the other half had a low-level of expertise. The subjects were given historical cases ranging from the 1930s to the 2000s and asked to identify all of the factors that contributed to each withdrawal. We will only briefly describe results of the experiment here.

A number of metrics were collected during the experiment and analyzed via a MANOVA analysis. The key finding ($p < 0.01$) was that the "experts" significantly outperformed the "non-experts" in only one era, the 2000s. This also happened to be era with the most complex causal relationships. Interestingly, the experts were not faster than the non-experts. Analysis of subject interaction with the interface suggests that the experts sought more information and made better use of the abstraction filter to facilitate sifting through the information.

One possible explanation is that non-experts had personal experience with the 2000s era. They may have even owned one of the cars in the experiment. Consequently, they may have felt that they knew enough about the era such that the need to search for more information via the interface was superfluous. Of course, this is just speculation, but it does lead to the question as to whether or not adding aiding that facilitates testing of hypotheses via the interactive visualization can close the gap between the experts and non-experts by counteracting the lack of information seeking behavior by the non-experts. This leads directly to the experiment performed under this task.

This experiment still involves diagnosing why the car failed but the data is presented and analyzed in a different way using a larger and more representative subject pool. First, the automotive cases are scored by the complexity of the situation (as determined by the number of contributing factors) so that it can be controlled as an experimental variable. Second, users are able to "tag" available evidence as either suggesting or contradicting potential causal factors. An aiding visualization then summarizes the results of this tagging for the subject. This enables them to consider explicitly the weight of the evidence both for and against any given factor. Third, engineers and managers in the automotive industry serve as the "expert" subjects. Stevens students continue to serve as the "non-experts." This setup allows us to consider three experimental variables: the complexity of the problem, the expertise of the subject, and the use of the aiding interface.

The critical question is whether or not the use of the aiding interface closes the performance gap between the experts and non-experts for the complex cases. A secondary question is whether the presence or absence of the aiding interface affects performance within the expert group itself.

AUTOMOBILE INDUSTRY APPLICATION

As discussed above, the cases presented to the subjects are historical examples from the automotive industry. A recently published study (Liu, et al., 2015) addressed the withdrawal of 12 car brands from the market during the 1930s, 1960s, and 2000s including the following cars:

- 1930s: Cord, Duesenberg, LaSalle, Pierce Arrow
- 1960s: DeSoto, Packard, Rambler, Studebaker
- 2000s: Mercury, Oldsmobile, Plymouth, Pontiac

The study focused on why these cars were removed. Explanations were derived at four levels: automobile, company, industry, and economy. Interestingly, only one of the twelve decisions was driven primarily by the nature of the car. Other forces usually dominated.

Data sources included quantitative data such as production levels for each car, market segment, and industry wide. Quantitative data also included financial information, e.g., revenues and profits, for companies and the industry as a whole. Data included text sources such as the ***New York Times*** archive, which contributed almost 100 articles published over the past 100 years on these vehicles. A variety of online sources were also accessed. There were also rich graphical components including, of course, picture of vehicles, but also pictures of executives, and graphical timelines.

HYPOTHESIZED USE CASES

The whole point of providing an interactive visualization for an enterprise problem is to support the user's foraging and sensemaking process. This process was discussed extensively in the RT-110 final report (Pennock, et al. 2015). To support the development of the experimental interface and aiding, we developed hypothetical use cases that allowed us to consider how a user might move through the available data to during the foraging and sensemaking process. We considered the available data from four levels of abstraction: Economy, Industry, Company and Car.

Problem statement: Brand X was removed from the market. Why did the company make this decision? Provide evidence to support your answer.

There are many ways to approach this question. Users with different level of expertise might have different approach in searching for data. An expert might think of looking at the Company level (industry publications or the Wall Street Journal) to find direct answers. Such articles might attribute to the withdrawal of brand X to decreasing sales.

Liu (Liu, et al., 2015) has shown that the causal chain can be traced back from the symptoms (withdrawal) to earlier decisions (investments, acquisitions, etc.). Thus, deeper answers are needed than "Brand X was withdrawn because it was not selling." The goal is to support users to identify the reasons it was not selling. They also should be able to determine the source(s) of the reasons. How did the company get into this situation?

To find deeper answers, one might start at the car level. How was brand X doing relative to competing brands? Were sales in this market increasing, flat, or decreasing across all companies? Causes of brand X sales decreasing in a decreasing market are likely very different from causes of decreasing sales in an increasing market. Is brand X losing and others winning, or is everybody losing?

If everybody is losing, one might then move from car to company to industry to economy, to determine why. If only brand X is losing, one would likely explore the company level to see whether brand X is really the problem rather than the rest of the company. To determine if brand X is the problem in itself, one might see how it competes with other brands in the market.

If brand X is not the source of its own problems, one would dig more deeply; one would look into company leadership and financial situations as potential reasons that brand X was sacrificed. It could be that product lines had to be trimmed and brand X was selected as the least painful alternative.

Another path would arise from discovering that everyone is losing, but other companies are not withdrawing brands. They may be better managed and have deeper pockets, or they may have a strategy that requires sustaining all of its brands.

Based on the results of the Rouse, et al. (2016) experiment, the concern is that a non-expert would only consider first order symptoms and jump to conclusions. Understanding the underlying relationships and structure of an enterprise requires the knowledge and/or genuine interest in that enterprise. Would an intuitive interface mitigate this issue? How difficult would it be to design an intuitive interface for a specific enterprise?

We associate meaning and functions to symbols that we see on daily basis. That's how we became familiar and avid users of innovations and new technologies like social networking. These innovations were designed for general purposes. Designing an interface with the same functionality in a technology niche area would be much more difficult that the general purpose one. First, these innovations and packages are accustomed for general purpose and don't fit technical cases. Second, it's difficult to design an intuitive interface for a technical purpose that doesn't require training. Experiments such this one discussed here are intended to investigate these issues.

EXPERIMENTAL DESIGN

The experiment is a three factor randomized design. The factors are the availability of the aiding interface, the expertise of the subject, and the complexity of the case. The response variable is the number of correctly identified causal factors.

For aiding, there are two levels: aiding and no aiding. Half of the subjects receive a version of the interface that has the capability to tag articles and collect these tags into single display as evidence that can be compared and contrasted. The other half subjects to not receive this capability.

For expertise, there are two levels: expert and non-expert. This will determination is made via the recruitment process. Subjects were recruited from the automotive industry to serve as experts. Non-experts were recruited from the Steven's student population.

Complexity is determined by the number of factors that contributed to the withdrawal in each case. Thus, a "simple" case has only a few contributing factors, while a "complex" case has several.

The experiment was web-based. Recruited subjects were invited to login to a website. The interface itself was developed using standard open source libraries. Data presented in the interface consisted of news articles as well as quantitative data such as production volumes, GDP, etc. presented in interactive graphs. User actions and responses were captured in a backend database. Each subject is randomly assigned to one era (1930, 1960, 2000) when registered. Aiding was provided to every other subject. Each subject is presented with four cars from the assigned era. They are asked to review the available evidence and identify the factors that contributed to the brand's withdrawal for each case. The cases are sequential and the subject must complete a case before moving to the next.

The interface itself has several components. There is an introduction for each case that provides an overview of the brand in question as well as representative photos (Figure 25)

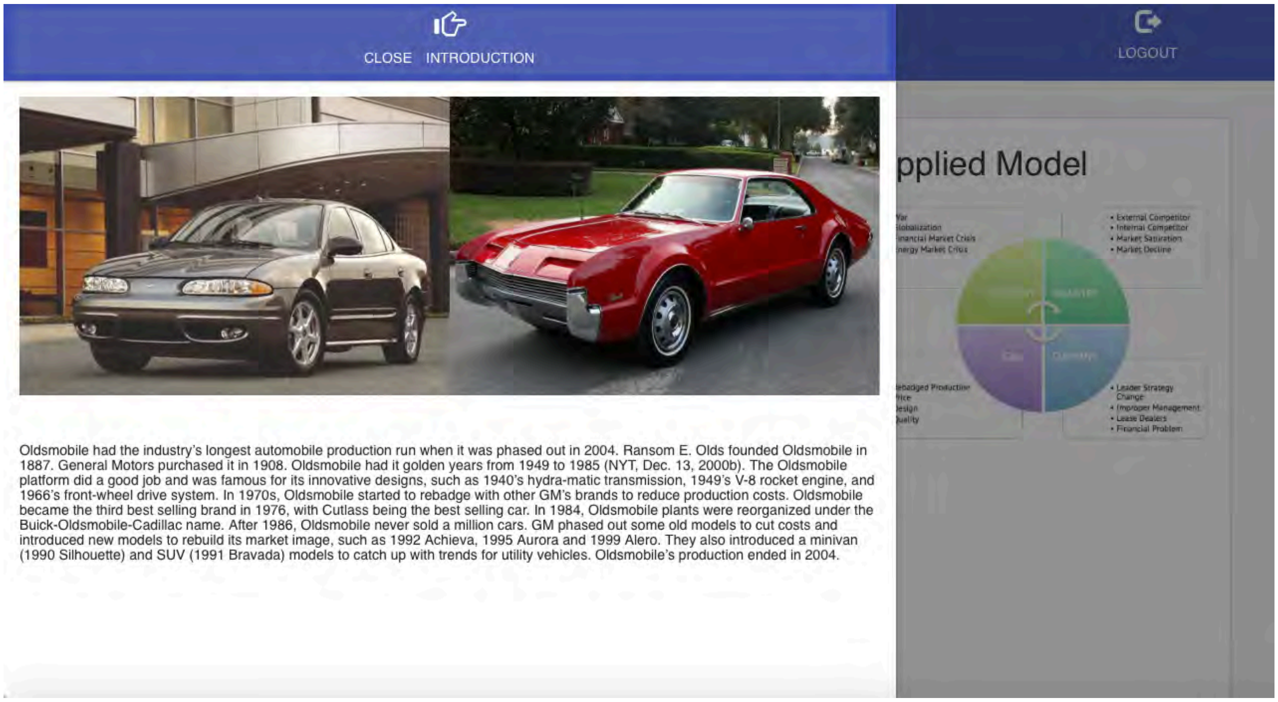


Figure 25 - Introduction view from side bar

Available data is organized into a tabbed interface. Each tab organizes data from a different level of abstraction (Economy, Industry, Company and Car). The subject may browse and view the available articles and graphs of quantitative data. (Figure 26) Subjects assigned to the aiding group can tag the articles with related words and mark charts for future reference.

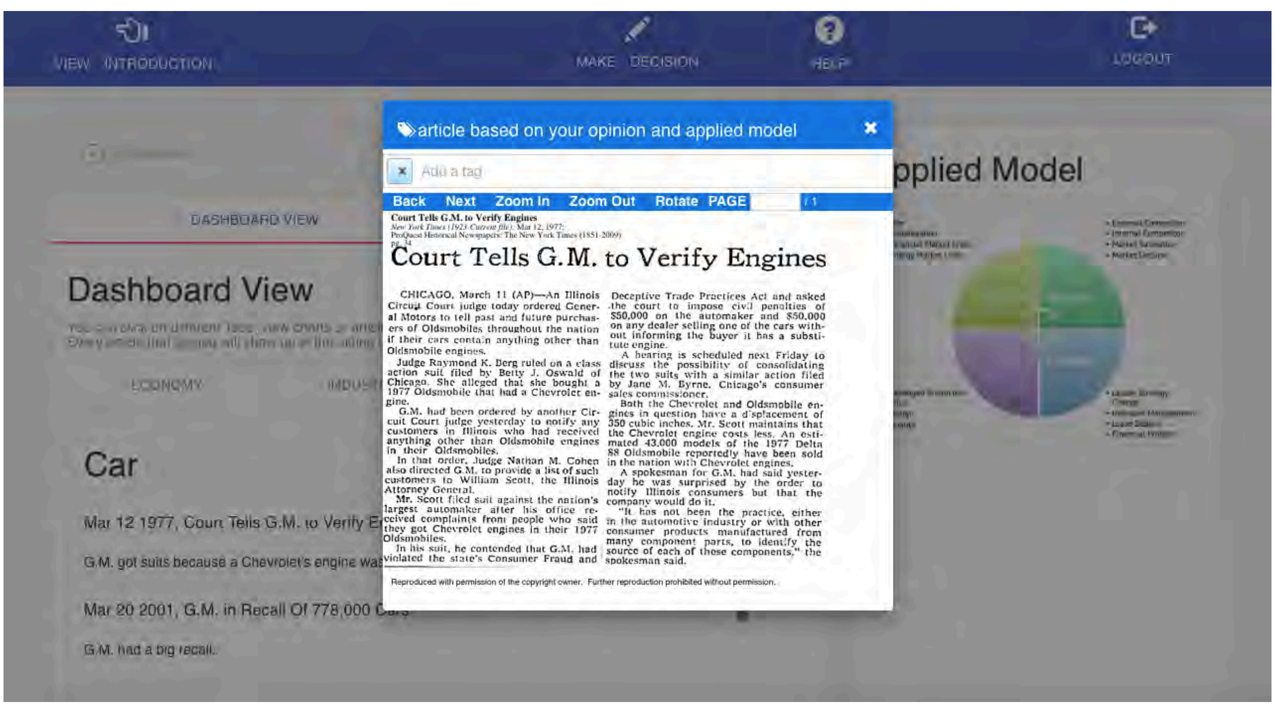


Figure 26 - Article dialog from dashboard view

Subjects assigned to the aiding group may also access the aiding interface. (Figure 27) The aiding interface consist of charts marked by the subject as relevant as well as a timeline that organizes the tags that the subject assigned to relevant articles. The subject is able to open associated the article and as well as edit the tags. The intent is to support subjects as they engage in the sensemaking process of posing and testing hypotheses.

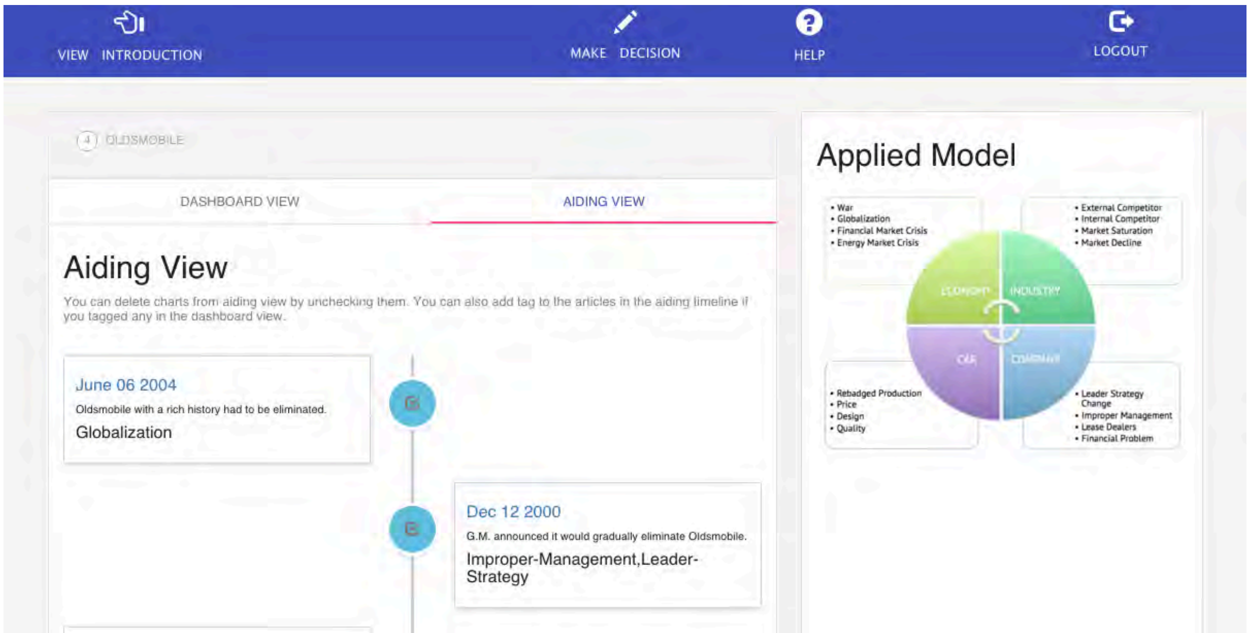
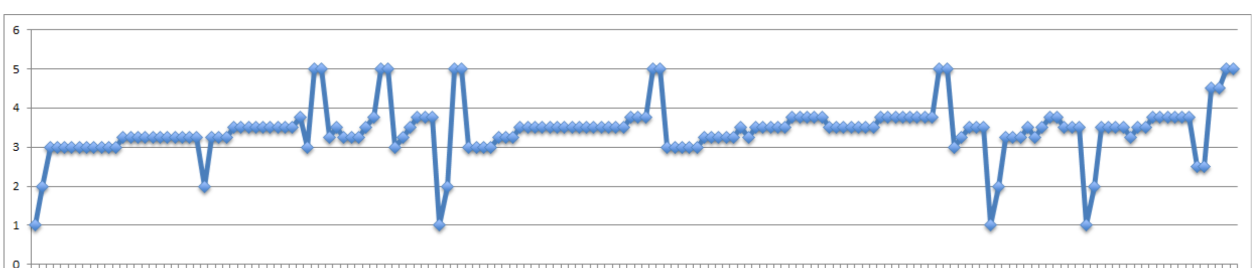


Figure 27 - Aiding view

Results and Analysis

At the time this report is was written, only a few subjects have completed the experiment. Consequently, we are unable to provide statistically significant results at this time. However, we can present analyses of the subject's usage of the interface features. Figure 28 shows the interface usage trajectories for three different subjects.



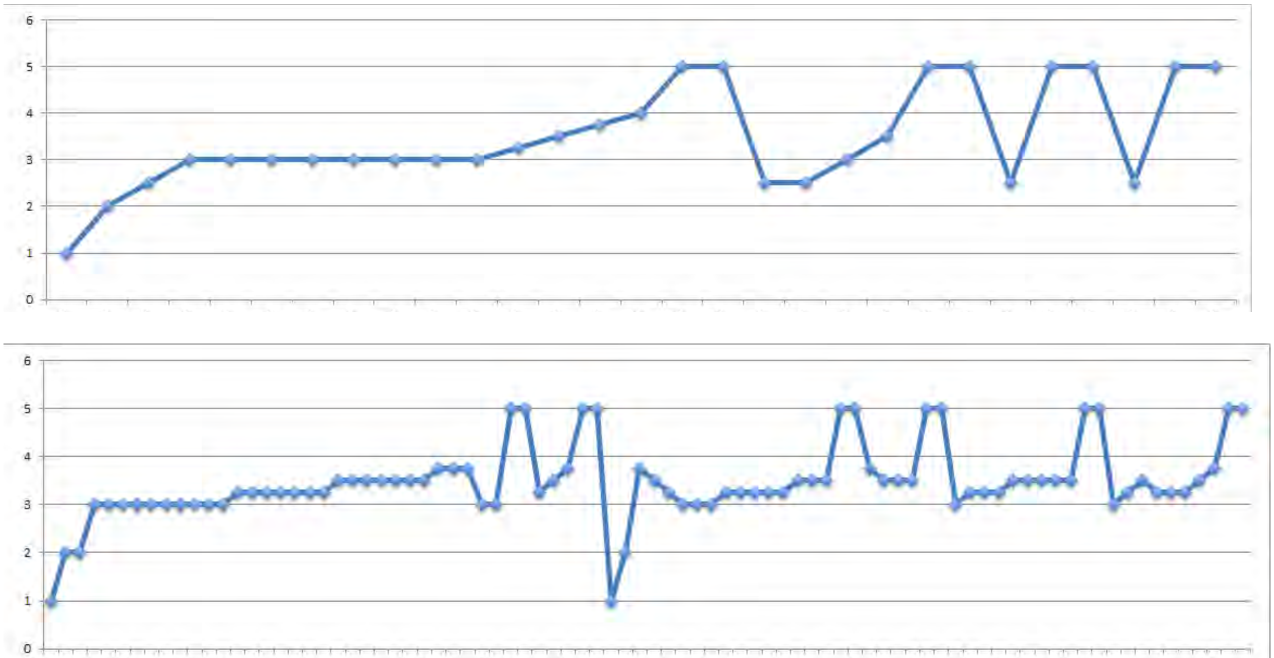


Figure 28 – Interface Usage Trajectories for Different Subjects

The usage trajectories can be interpreted in the following manner. Each marker indicates an action taken in the user interface. For each subject, the markers are ordered by the sequence in which they occurred. The position of each marker on the y-axis indicates the action taken. The relative distances between the markers are not meaningful. Table 9 provides the coding for the position of the marker on the y-axis.

Table 9 - Interface Action Coding

Code	Action
5	Decision Dialog
4.5	Help
4	Aiding
3.75	Dashboard-Car
3.5	Dashboard-Company
3.25	Dashboard-Industry
3	Dashboard-Economy
2.5	Introduction-sidebar
2	Description
1	login

Again, while there are not enough subjects yet to draw statistically significant conclusions, we can already see differences in usage behavior among the three subjects shown in Figure 28. The first subject made the most use of the available data by far, while the second user accessed some of the data for the first case, and then hardly any at all for the subsequent cases. The third subject fell somewhere in between. It is also worth noting that while the second subject had access to aiding, they only accessed it once. So, at least in the case of this subject, the presence of aiding did not seem to result in strong information seeking behavior. As additional subjects complete the experiment, we will be able to determine whether or not this is an anomaly.

8. REVISITING THE ENTERPRISE MODELING METHODOLOGY

As the enterprise modeling effort has progressed over several RTs (44, 110, 138), it is instructive to revisit the enterprise modeling methodology as it was originally proposed and consider how the subsequent research has impacted it. In short, our view has evolved. During RT-44, the objective was to build an accurate multi-layer model of the enterprise and then use it perform "what if" analyses of policy options to find the best alternative. At the time, we knew that model composition was problematic (see the RT-44a Phase 2 discussion of DMMF and IBM Splash (Rouse and Pennock 2013)), but we did know what to propose in its place.

Essentially, the approach we proposed is depicted in Figure 29. Relevant enterprise phenomena are organized into layers of abstraction. Identify the best in breed model for each layer. These models are coupled together to create a single integrated model. This integrated modes is then used to perform any necessary analysis.

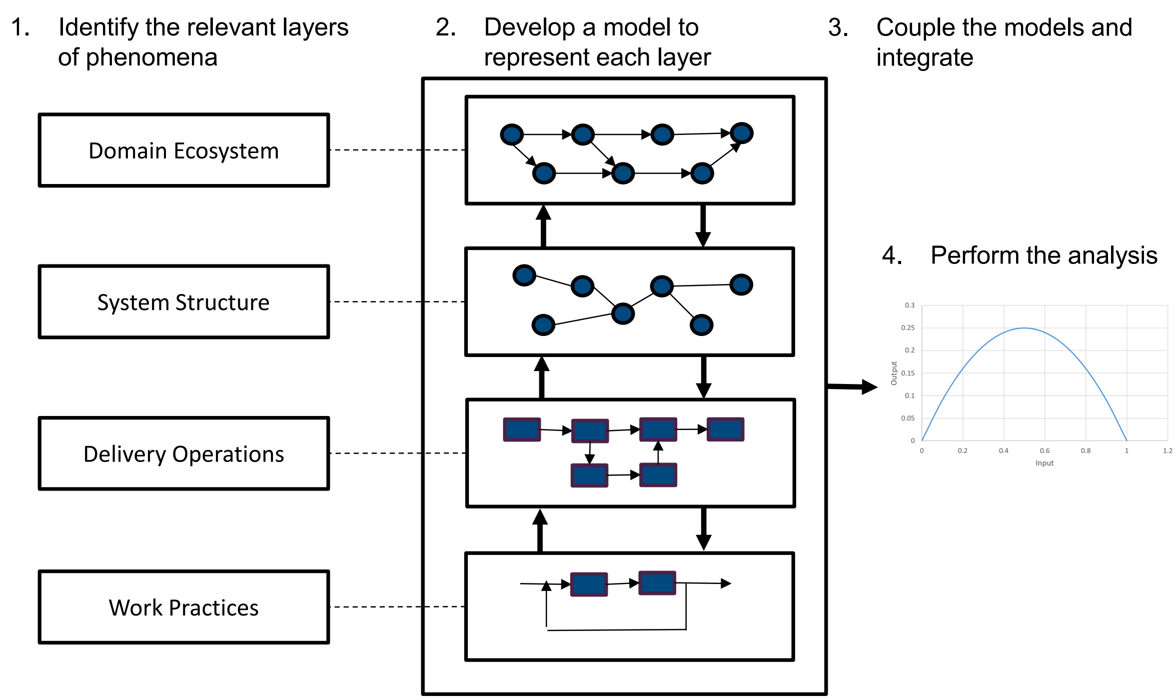


Figure 29- Preliminary Approach to Modeling Enterprise Systems

Since that time, the counterfeit parts simulation and other case studies allowed us to assess the feasibility of this approach. We identified three issues:

First, the nature of enterprise systems means that simply adding layers to the problem can cause a substantial increase in epistemic uncertainty (Section 6). Second, composition can be challenging due to ontological inconsistencies (Section 4). Finally, the "best" model depends on imprecisely known or unknown circumstances (Sections 3 and 6).

While composition of independent modeling layers at different scales has been accomplished for physics based systems (with difficulty), it is not clear that this is feasible for enterprise systems. Section 4 highlights the theoretical rationale behind this difficulty.

Consequently, modeling an enterprise system is more an exercise in managing uncertainty than maximizing predictive accuracy. "Good enough" relative to the problem at hand should be the goal, and this changes how we approach modeling enterprises. Here our experiences with developing the counterfeit parts simulation are instructive (Section 2).

At the start of the development process for the counterfeit parts simulation, it was conceptualized using multiple interacting layers of abstraction. However, it was not actually built this way. It turns out that building separate, independent layers for each model is impractical in many situations. As revealed by the analysis in Section 4, developing a separate model for each layer:

- Generates ontological and logical inconsistencies
- Likely increases epistemic uncertainty
- Requires the introduction of "fictions" to facilitate handshakes
- Fictions are determined via trial and error and are likely question specific

In short, this approach probably only makes sense for modeling a multi-scale physical system where the need to repeatedly examine the same class of questions justifies the effort to "calibrate" the handshakes.

Based on the lessons learned from the effort to develop the counterfeit parts simulation, a more practical approach is to construct a fully integrated core model that addresses one or two of the layers of abstraction. In the case of the counterfeit parts simulation, the supply chain/system model forms the core. The other layers of abstractions are addressed by developing peripheral models that are tailored to interact with the core. Examples in the counterfeit parts simulation include the broker behavior model, customs behavior model, and the recycling model.

The function of the peripheral models to "perturb" the core model to generate useful insights. For example, imposing supplier qualification policies causes "poor" programs to run out of suppliers more quickly. A major risk to implementing a policy option in an enterprise is crossing a tipping point that no one knew was there. The peripheral models can be used to trigger bifurcations in the behavior of the core model. Finding the bifurcations depends on exploring structural and ontological variations on the peripheral model, and scientific sub disciplines may serve as a source of such variations.

While the natural tendency in enterprise modeling seems to be to maximize predictive accuracy by maximizing the fidelity of the model (i.e., add as many relevant factors as possible), this approach has rapidly diminishing returns as this increases the degrees of freedom and risks over-fit with sparse data. Rather it is more productive to build a relatively simple core model and then selectively perturb it with structural variations in the peripheral models to see if this triggers any unexpected behaviors (i.e., bifurcations or tipping points). This idea is illustrated graphically in Figure 30.

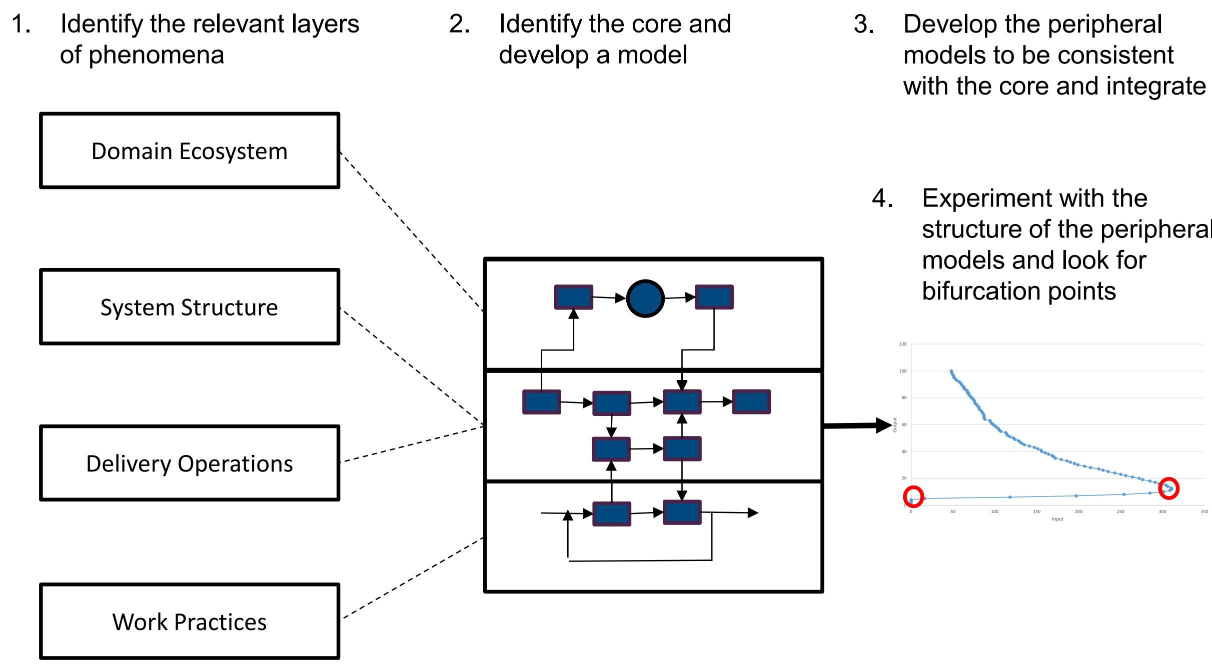


Figure 30 – Revised approach to modeling enterprise systems

This revised approach allows one to extract meaningful insights to support policy analysis while still keeping the level of epistemic uncertainty manageable. The conclusion that we reach is that at a high-level, the multi-level approach advocated in the enterprise modeling methodology is valuable for conceptualizing the model, but implementing the model requires a less literal interpretation of the layers. In short, in many circumstances, the enterprise simulation should not be a one-to-one mapping of layer to model.

9. Humanitarian Response Case Study

As a follow-on to the counterfeit parts case study, we sought another case study to demonstrate the effectiveness of the enterprise modeling methodology, as well as to help generate recommendations for its improvement. A variety of candidate DoD-related enterprise problems were considered, including humanitarian response to crises and disasters, planning and protecting critical infrastructure, and evolving an eco-system for design and development of modular systems to meet DoD needs in a cost-effective and time-critical manner. This section addresses a follow-on case study in humanitarian response.

BACKGROUND

For a variety of reasons, humanitarian crises have become increasingly common and important. Weather may be becoming more extreme. Increased populations have inhabited areas susceptible to disasters in greater numbers than before. Media is ubiquitous, including citizen journalists, and they broadcast the plights of imperiled populations to vast audiences, provoking pressure to respond.

In recent years, a body of research has emerged to study effective methods for humanitarian crisis response (Celik et al., 2014). Some research has focused on the nature of response problems. For instance, disaster response is impacted by complexity, which is in turn affected by population density and response lead time (Christenson & Young, 2013). Most research, though, has focused on specific methods to solve types of problems in the humanitarian response context.

Humanitarian response can be considered in three phases. In the planning phase, pre-positioning of supplies is a major concern. Duran et al. (2011) study methods for effectively pre-positioning supplies in this context. In the response phase, agencies deliver supplies and other goods and services to affected areas. Ekici et al. (2014) model effective food distribution strategies during pandemics. Finally, there is the post-disaster recovery phase. This phase is often neglected and can take long lead times before any semblance of normality is reached. Clearing debris after a disaster is one example. Celik et al. (to appear) develop efficient methods for clean-up of debris.

In any humanitarian crisis, a number of different organizations act and interact to resolve the situation. These range from government agencies, to non-governmental organizations, to private-sector firms. Case studies of disaster responses by these types of organizations are useful in terms of developing effective practices (Ergun et al., 2010). In addition, methods such as cooperative game theory can be applied to improve effectiveness of collaboration (Ergun et al., 2014).

The Department of Defense is a key player in humanitarian response in the U.S. and around the world. According to DoD Instruction 3000.05 (*Stability Operations*, September 16, 2009b), humanitarian relief is a "core U.S. military mission that the Department of Defense shall be prepared to conduct with proficiency equivalent to combat operations." In recent years, DoD has engaged in numerous humanitarian missions, ranging from Hurricane Katrina, to the 2010 Haiti earthquake, to 2010 monsoon floods in Pakistan. DoD has used extensive resources in such efforts. Dozens of ships, for instance, are used in a typical response to many crises (Apte et al., 2013).

Research has addressed a number of issues relating to DoD involvement in humanitarian response. Apte and Yoho (2013) study policy options for planning to determine preferred options for delivery of material under different conditions. Policy options include pre-positioning, proactive deployment, phased deployment and surge capacity. With a given portfolio of systems and platforms, it is important to determine the best set of platforms and systems to use in a response situation in terms of effectiveness and cost (Apte & Yoho, 2014). Finally, coordination has been studied between the military and NGOs, with at least one recommendation of better coordination between military and NGO efforts (Apte & Hudgens, 2015).

In addition, research has looked at various DoD humanitarian response efforts to make recommendations on future improvements. Cecchine et al. (2013) review the Army response to the 2010 Haiti earthquake and recommend updated policies, a national framework for responding to crises in foreign countries, improved familiarity of stakeholders with recommended practices, and the establishment of a standing organization for addressing crises. Moroney et al. (2013) review a number of crisis response efforts and make recommendations in terms of improving DoD response efficiency, enhancing interagency coordination, improving coordination with foreign government and NGOs, and building goodwill through relief efforts. Pirnie and Francisco (1998) review past relief and peacekeeping efforts, characterize them, and then make recommendations on effective response protocols that do not jeopardize the overall DoD capability for large-scale combat.

The existing research has certainly recognized the multi-stakeholder nature of the humanitarian response problem. Clearly, it is an enterprise problem in nature. Major humanitarian disasters involve a multitude of governmental jurisdictions. Those in the United States involve federal, state and local authorities. At each level, there are numerous agencies that coordinate response, including provisions, water treatment, accommodations for displaced persons, health services, rescue, and security, among others. In addition, non-profit NGOs and corporations typically are involved in provisions, accommodations and health services. Those in foreign countries similarly involve the relevant governmental agencies of the affected areas, the UN, U.S. federal government agencies and NGOs. Aside from the multiple organizations involved, the following features make this an enterprise problem in the context of the enterprise systems modeling research for this project.

- There is no locus of control.
    - Multiple agency/industry stakeholders are addressing the problem, ranging from DoD and its services, to NGOs, to private firms. In the U.S., state and local authorities are involved. In foreign countries, other governments and their agencies are involved, as is the Department of State.
    - Government agencies can promulgate policy, but must be careful of conflicts among different governments or levels of governments, as well as the potential for populations not to follow directives in times of crisis.
- There is significant adaptive behavior.

- Populations adapt to changing circumstances. Panic may ensue. Populations may overrun planned evacuation routes. Populations may horde supplies.
  - Looting and other criminal activity may occur.
  - Crises may occur in regions with conflicts, in which case combatants may adapt to policies and actions as part of their tactics and strategies.
  - Policy-makers must anticipate and adapt to potential behaviors of populations.
- There is significant complexity.
  - There is substantive socio-technical behavior (human behavior and social behavior interacting with technical system).
  - There are multiple organizations, processes and systems interacting with unpredictable effects.
- Each crisis has unique characteristics.
  - While there are useful classifications, each crisis has unique characteristics relating to cause, scope, location, effects on populations, population reactions. In addition, there is a unique "enterprise" created to address each crisis in terms of the set of responding organizations.

As noted above, each humanitarian crisis is unique. Thus, one need for enterprise modeling is to facilitate effective formations of "pick-up" enterprises to address crises. The issue here is to create a model generic enough that it will supply insights useful in a variety of situations, but realistic enough that those insights are practical and useful. The model should address a particular type of disaster (e.g., earthquake, tsunami, hurricane) and location type (e.g., major city, large rural or undeveloped area, foreign vs. U.S.). Data from previous similar disasters could then be used. Then customization would be built in so that the user can select such attributes as size of disaster, particular elements to include or not, and particular agencies involved.

This is in line with the approach of Pirnie and Francisco (1998), who develop a series of vignettes for different crisis types to support their recommendations. For purposes of this initial model, we select one of their vignettes to provide a concrete example for the model, while also discussing modeling features for humanitarian response in a more general context. This vignette involves military support to foreign authorities following a disaster such as a flood or hurricane on a littoral. Major activities include search-and-rescue, evacuation, and reconstruction afterward. There is a dissident population that can hinder the mission. The particular inspiration for this class of humanitarian responses is the Bangladesh cyclone of 1991, during which the Marines with support from other services provided relief and support to a devastated area of that country.

MODELING METHODOLOGY

Here, we use the modified version of the enterprise modeling methodology described previously. The next several sections guide development of the case study model through these steps.

CENTRAL QUESTIONS OF INTEREST

Here there are two primary questions of interest. The first is how to arrange effective "pick-up" enterprises to address crisis situations that may fall into broad classifications, but also have unique characteristics. The second is given a crisis situation, what are effective enterprise policies for managing it. Potential policies would address the following, keeping in mind trade-offs and anticipation of adaptive behaviors.

- Pre-positioning of supplies and assets
- Pre-positioning of trained personnel
- Selection of coordination mechanisms among agencies and NGOs (including foreign governments)
- Reinforcement/replacement/redesign of infrastructure
- Evacuation planning and inter-agency coordination
- Protocols for communicating information to population
- Budgeting for humanitarian efforts
- Addressing post-crisis recovery

KEY PHENOMENA

For this initial model, we organization key phenomena into the following generic categories.

- Crisis – the nature of the crisis itself and its ongoing effects are a key element of any model.
- Response assets – response assets include three major elements: personnel, platforms and systems, and supplies.
- Supply chain flows – supply chain flows model the movement of response assets through networks of facilities. In different locations, assets may be temporarily stored, may be consumed (in the case of supplies), or may provide services (in the case of personnel and platforms and systems).
- Infrastructure – infrastructure consists of systems that deliver critical service such as water or power, or that provide population-level functions such as transportation systems. It also includes permanent assets already in place to mitigate crisis effects, as

opposed to temporary response assets. Example would be back-ups for normal infrastructure, storm water barricades, etc.

- Enterprise actors – enterprise actors are organizations that may be part of a response pick-up enterprise, but do not have direct policy-making roles in the model. Examples may be NGOs, private firms or foreign government agencies. It should be noted that not all enterprise actors may be involved in a response, as part of the model's purpose is to determine effective methods to assemble a response enterprise from a set of potential participants.
- Policy actors – policy actors, on the other hand, have policy-making roles that are of interest in terms of modeling and assessing effectiveness. For purposes of this model, we limit such policy actors to DoD agencies and other U.S. government agencies.
- Exogenous effects – this category includes factors that are external to the direct enterprise, but that have significant effects on its performance and behavior. Funding levels for infrastructure, aggregate population behaviors, and cultural norms and effects are examples.

These are similar to the categories used for the counterfeit parts model, in part because both models have a substantive role for the supply chain. However, the nature of the crisis problem, as being caused by forces external to the enterprise, differs from the counterfeiting problem, which is caused by suppliers that essentially are part of the enterprise. Thus, the nature of the crisis is addressed as a separate category here. In addition, the supply chain differs in that it focuses on service and supply delivery rather than on manufacturing and assembly plus maintenance and repair. Finally, infrastructure is represented since it is an important characteristic of response effectiveness.

The model of the specific littoral response to a hurricane crisis provides the following phenomena to be represented, as shown in Table 10.

Table 10 - Key phenomena in littoral response situation

Category	Phenomena
Crisis	• Affected areas, flooding effects, destruction of population centers, homelessness and population displacement, ongoing, blocking of traffic routes with debris, effect on infrastructure
Response assets	• Military personnel – Marine Expeditionary unit, Army special forces, Army engineer company, helicopter support company • Amphibious ships, dock landing ships, helicopters, aircraft assets for airlift and evacuation

	• Potable water, medicine and medical supplies, rations • Temporary or constructed supporting infrastructure for response (airport, seaport, storage)
Supply chain	• Staging locations for response assets from initial deployment locations to locations of deployment during crisis and modes of transport
Infrastructure	• Existing infrastructure and repair to damage, rather than planning upgrades beforehand • Roads, harbors, sanitation systems, power systems
Enterprise actors	• Civil authorities, other U.S. government agencies, NGOs, dissidents
Policy actors	• DoD, Army, Marines, Air Force, Navy
Exogenous effects	• Cultural norms driving dissidents, population reactions to crisis and aftermath

Visualizations of Relationships among Phenomena

An initial visualization of the enterprise phenomena was created using the multi-level model approach (Rouse & Bodner, 2013). This is shown in Figure 31. In the future, visualizations of individual elements within this framework would be useful. For instance, the dynamics of population movements and evacuations, or the progress of rebuilding infrastructure on the availability of services would be of interest.

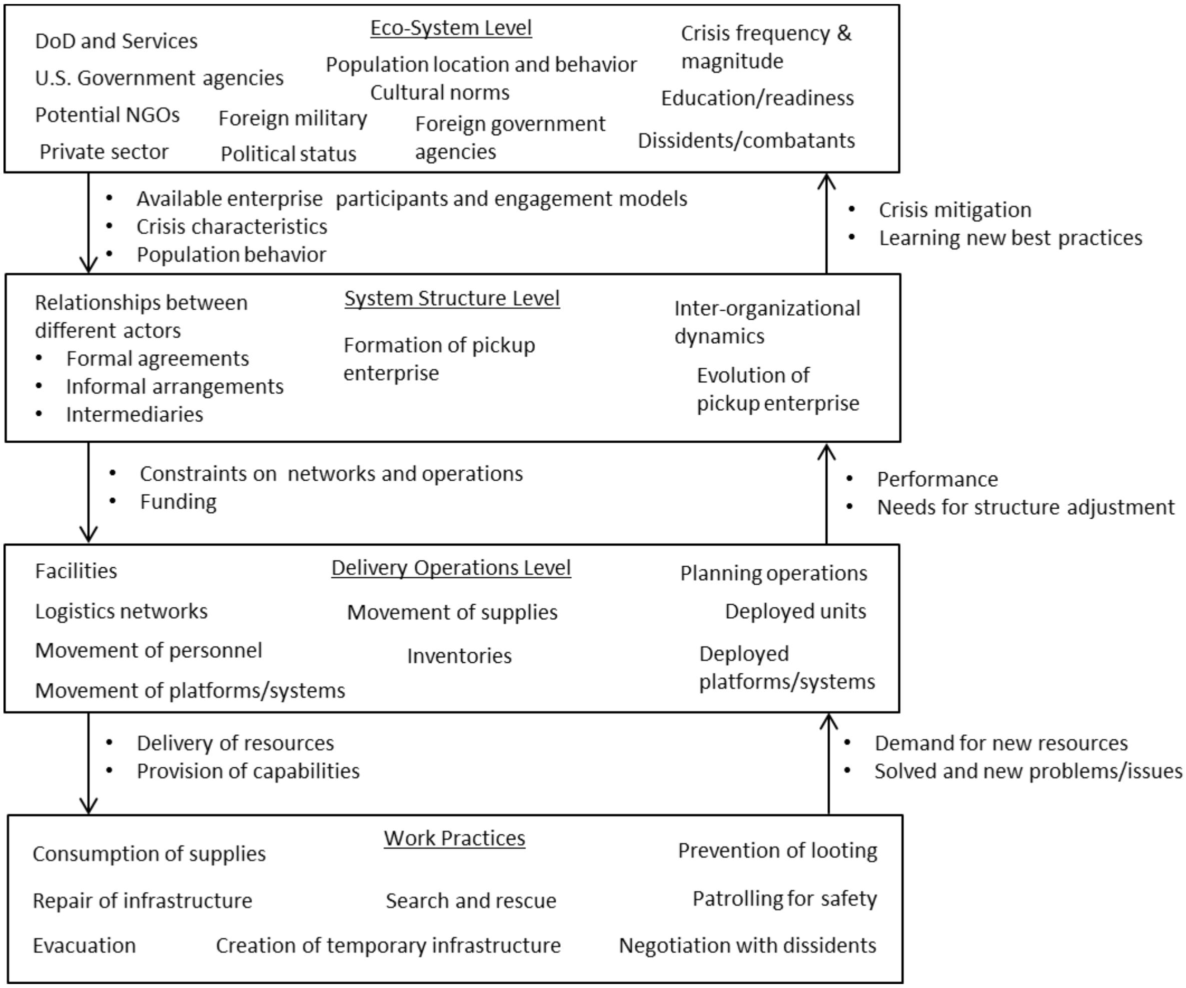


Figure 31 - Enterprise phenomena and relationships

KEY TRADEOFFS THAT APPEAR TO WARRANT DEEPER EXPLORATION

Here, we describe some key traded-offs for the model to address, specifically posed for the littoral response scenario.

- What is the trade-off between pre-positioning response assets versus just-in-time deployment in terms of cost and lead time response?
- What are trade-offs in various communication modes, frequencies and content timing in terms of desired population reaction (orderly evacuation) versus unintended effects (e.g., panic, overcrowding on evacuation routes, etc.).
- What is the trade-off between spending effort negotiating with dissidents versus other activities in terms of results in reconstruction of infrastructure and delivery of supplies?

- What are trade-offs in methods for establishing cooperation with local civil authorities in terms of lead times and getting access to resources/locations needed for supply delivery and reconstruction?
- What are trade-offs in alternate pick-up enterprise arrangements in terms of cost, lead times and response effectiveness?
- In what ways should the pick-up enterprise adapt under different circumstances to best manage trade-offs?

More generally, trade-offs could also focus on issues such as planned upgrades and strengthening of infrastructure in advance of crises versus developing agile infrastructure in terms of cost and construction/set-up lead times, and effectiveness. This particular trade-off would not apply in the littoral response scenario since the location is a foreign country that may not have resources to support such investments.

ALTERNATIVE REPRESENTATIONS OF THESE PHENOMENA

Simulation modeling provides three paradigms for representing phenomena. Discrete-event models view the world as a series of discrete state changes and are often used to model queueing systems where customers or other entities travel through a series of resources with time delays for service at the resources and delays travelling between them. System dynamics models view the world as a set of continuous flows, or continuous state changes, with feedback loops and lags modeled in the flow system. Agent based simulations operate using a discrete event worldview, but the focus is on modeling system elements individually with interactions provided via message-passing. Table 11 discusses representations for the various categories of

Table 11 - Representation alternatives

Category	Representations
Crisis	• System dynamics models represent the continuous impacts of the crisis in terms of storm progress and gradual degradation of power and water systems, and displacement of populations in different areas. Discrete effects are modeled by variables that are "triggered" by conditions in the system dynamics model.
Response assets	• Military units are represented as complex agents that perform specialized tasks. • Units have platforms and systems that are represented also as agents. These agents have location and state-based behaviors for availability. • Supplies (potable water, medicine and medical supplies, rations) are represented as variables indicating current

	levels at various locations. Shipments of supplies are represented as simple agents indicating the amount being shipped, expiration dates, etc. • Temporary or constructed supporting infrastructure is represented as agents indicating current capacity and state-based availability.
Supply chain	• Staging locations for response assets are represented as agents with collections of asset shipments. Assets are shipped via message-passing with delays for transport time. • Alternatively, the supply chain could be represented as a discrete event model using the process-interaction paradigm (i.e., network of queues).
Infrastructure	• Existing infrastructure is represented as represented as agents indicating current capacity and capability and state-based availability and degradation.
Enterprise actors	• The various enterprise actors are represented as complex agents that interact with one another.
Policy actors	• The various enterprise actors are represented as complex agents that interact with one another. These agents have policies that if enabled or set to certain levels, initiate changes in other model elements.
Exogenous effects	• Cultural norms are represented as a series of variables. Population dynamics and reactions to the crisis are represented by system dynamics models.

ABILITY TO CONNECT ALTERNATIVE REPRESENTATIONS

Similar to the counterfeit parts case study, we use AnyLogic® for model composition involving these three simulation paradigms. AnyLogic is a commercially available simulation software used to model complex systems. It also provides a Java™ API so that domain-specific customization can be used to enhance models. AnyLogic has proven quite capable for modeling the counterfeit parts case study and composing its different model elements.

CORE MODEL AND INTERACTING MODEL OVERVIEW

The core model consists of the response assets, the supply chain, the infrastructure, the enterprise actors and the policy actors insofar as their interaction with the rest of the core model. In terms of the littoral response scenario, the core model addresses the responding units and their systems/platforms, supplies, and temporary infrastructure items; plus the supply chain delivering supplies, systems and platforms, and personnel; and the services, NGOs, foreign government agencies and dissident groups. The interacting models consist of the exogenous effects model and policy decisions that a user can enable through the policy agents.

Similar to the counterfeit parts model, the humanitarian response model is primarily agent-based with some systems dynamics. Agent-based representations were selected over discrete-event representations due to the overall nature of the model as a set of interacting elements. In future versions of the model, though, discrete-event representations may prove useful in representing business processes important in the operation of most enterprises.

FUTURE WORK

This section has presented an initial model for humanitarian response. Future work will flesh out the initial model into a more fully functional model with the intent of identifying stakeholders, developing a focus for the model on a particular class of crisis response situations of interest, refining and enhancing the model (particularly with data from past crisis response efforts), and validating the model. In addition, customizations will be developed that will allow an analyst to experiment with different crisis "settings" within the particular class of crises for which the model is designed. For instance, such customizations may relate to the strength or path of a hurricane or the duration of flooding.

10. CONCLUSIONS AND FUTURE WORK

As part of the RT-44a task, we proposed a ten-step approach to modeling and understanding enterprises with the end goal of supporting policy makers. RT-110 and RT-138 applied and evaluated this methodology through a series of case studies, most notably the counterfeit parts case study. What we learned is that while the sequence of determining questions of interest, identifying the associated phenomena, and organizing them into levels of abstraction is useful for conceptualizing the model, the implementation requires a less than literal application of the layers.

Instead, we found that rather than trying to capture each layer in detail as a separate model, it is better to identify the layers that are most relevant to the questions of interest and develop a fully integrated but relatively simple "core" model of those layers. The other layers are then addressed by developing compatible peripheral models that are intended to perturb the behaviors of the core model.

A related conclusion, is that incorporating additional layers does not necessarily improve the "accuracy" of the model in terms of quantitative prediction. Rather, adding the additional layers increases the degrees of freedom and, consequently, increases the risk over-fit when limited data is available (which is the typical case). Rather the motivation behind adding the layers is to uncover the possibility of unexpected or counterintuitive behaviors. This insight was reinforced through a series of expert reviews of the counterfeit parts simulation.

What we conclude from the results of this research task is that what is needed to support policy makers and analysts concerned with enterprise systems is not a model that produces accurate forecasts. This is likely impossible. The situation is this: The Government and many business enterprises are fairly adept at developing first order models of policy impacts. This is effectively an exercise in extrapolating the trend. As long as the trend continues, these forecasts are sufficient. The challenge is that sometimes higher order effects can overwhelm the first order effect and/or trigger unintended, negative side effects. It is these higher order effects that policy makers would like to identify.

Given the underlying complexity of enterprise systems, it is highly unlikely that one will be able to forecast these effects with any precision. However, it sometimes possible to at least establish their possibility. This is where enterprise modeling approach developed through the course of this research effort can have impact. Consequently, in future work, we intend to develop a systematic approach to identify counter-intuitive results and policy tipping points while simultaneously considering the enterprise at multiple scales of resolution and multiple perspectives. We believe the that core-peripheral view coupled with the mathematical analysis of model composition presented in this report provide a starting point to develop a more directed approach to finding these higher order effects that will be more effective than relying on serendipity.

11. References

ABA (2012). A White Paper Regarding Department of Defense Implementation of Section 818 of the National Defense Authorization Act for Fiscal Year 2012. Report issued by Task Force on Counterfeit Parts of the Committee on Acquisition Reform and Emerging Issues of the American Bar Association Section of Public Contract Law, October 5, 2012.

Agarwal, H., Renaud, J. E., Preston, E. L., & Padmanabhan, D. (2004). Uncertainty quantification using evidence theory in multidisciplinary design optimization. *Reliability Engineering & System Safety*, *85*(1), 281-294.

AIA (2011). Counterfeit Parts: Increasing Awareness and Developing Countermeasures. Arlington, VA: Author.

Alderson, D. L. and J. C. Doyle, (2010), Contrasting Views of Complexity and Their Implications for Network-Centric Infrastructures, *IEEE Transactions on Systems, Man, and Cybernetics-Part A: Systems and Humans*, 40(2): 839-852.

Apte, A., & Hudgens, B. (2015). Analysis of Capabilities of Organizations in the Areas of Responsibility for U.S. Commands of the United States Navy to Humanitarian Assistance and

Disaster Relief. Report NPS-LM-15-125. Monterrey, CA: Graduate School of Business & Public Policy, Naval Postgraduate School.

Apte, A., & Yoho, K. (2014). Optimizing Resources of United States Navy for Humanitarian Operations. Report NPS-LM-14-180. Monterrey, CA: Graduate School of Business & Public Policy, Naval Postgraduate School.

Apte, A., & Yoho, K. (2013). Strategies for Logistics in Case of a Natural Disaster. Report NPS-LM-11-188. Monterrey, CA: Graduate School of Business & Public Policy, Naval Postgraduate School.

Apte, A., Yoho, K., Greenfield, C., & Ingram, C. (2013). Selecting maritime disaster response capabilities. Journal of Supply Chain Management, 6(2), 40–58.

Barberis, N. C. (2012). *Thirty years of prospect theory in economics: A review and assessment* (No. w18621). National Bureau of Economic Research.

Barberis, N., P. J. Haas, C. Kieliszewski, Y. Li, P. Maglio, P. Phoungphol, P. Selinger, Y. Sismanis, W. Tan, I. Terrizzano, (2012), *Splash: A computational platform for collaborating to solve complex real-world problems*, IBM Research - Almaden, http://researcher.watson.ibm.com/researcher/view_project.php?id=3931, accessed 10/28/13.

Barnett, L., Bossomaier, T. (2012) Transfer Entropy as a Log-Likelihood Ratio. Physical Review Letters 109(13):138105.

Bartholomew, David. Knott, Martin. Moustaki, Irini. (2011) Latent Variable Models and Factor Analysis: A Unified Approach, 3rd Edition. Wiley.

Bodner, D. A. (2014). Enterprise modeling framework for counterfeit parts in defense systems. In *Proceedings of the 2014 Complex Adaptive Systems Conference*.

Bodner, D. A. (2015). Mitigating counterfeit part intrusions with enterprise simulation. In *Proceedings of the 2015 Complex Adaptive Systems Conference*.

Bossomaier, T. Barnett, L. Harre, M. (2013) Information and Phase Transitions in Socio-economic Systems. Complex Adaptive Systems Modeling, v1, 9.

Brownstone, D., Ghosh, A., Golob, T. F., Kazimi, C., & Van Amelsfort, D. (2003). Drivers' willingness-to-pay to reduce travel time: evidence from the San Diego I-15 congestion pricing project. *Transportation Research Part A: Policy and Practice*, *37*(4), 373-387.

Bureau of Labor Statistics, (2015) "May 2014 national occupational employment and wage estimates," http://www.bls.gov/oes/current/oes nat.htm (accessed 10/12/15), United States Department of Labor.

Business Insider (2012). Counterfeit Chinese microchips are getting so good they can't be identified. http://www.businessinsider.com/counterfeit-parts-from-china-raise-grave-concerns-for-both-us-companies-and-national-security-2012-6.

Camerer, C. & Loewenstein, G. (2004). Behavioral economics: Past, present, future, in C. Camerer, C., Loewenstein, G. & Rabin, M. *Advances in Behavioral Economics*. New York, NY: Princeton University Press.

Casti, J. L. (1986). On system complexity: Identification, measurement, and management. In *Complexity, language, and life: Mathematical approaches* (pp. 146-173). Springer Berlin Heidelberg.

Cecchine, G., Morgan, F. E., Wermuth, M. A., Jackson, T., Gereben Schaefer, A., & Stafford, M. (2013). The U.S. Military Response to the 2010 Haiti Earthquake Considerations for Army Leaders. Santa Monica, CA: RAND Corporation. http://www.rand.org/pubs/research_reports/RR304.html

Celik, M., Ergun, O., Johnson, B., Keskinocak, P., Lorca, A., Pekgün, P., & Swann, J. (2014). Humanitarian logistics. In *New Directions in Informatics, Optimization, Logistics, and Production, TutORials in Operations Research*, (pp. 18-49), P. Mirchandani (ed.) and C. Smith (series ed.), Hanover: INFORMS.

Celik, M., Ergun, O., Keskinocak, P. (to appear). The post-disaster debris clearance problem under incomplete information. *Operations Research*.

Christensen, J. K., & Young, J. K. (2013). Drivers of Complexity in Humanitarian Operations. Thesis. Monterrey, CA: Graduate School of Business & Public Policy, Naval Postgraduate School.

Cohen, B. S., & Lee, K. (2014). On the limits of test in establishing products assurance.

Congressional Research Service (CRS). (2005). Hurricane Katrina: DoD Disaster Response Report RL33095. Washington, DC: Author.

Daganzo, C. F., Gayah, V. V., & Gonzales, E. J. (2011). Macroscopic relations of urban traffic variables: Bifurcations, multivaluedness and instability. *Transportation Research Part B: Methodological*, *45*(1), 278-288.

DAU (2013). Anti-counterfeiting. Defense Acquisition Guidebook. https://acc.dau.mil/CommunityBrowser.aspx?id=638350&lang=en-US.

DeYoung, B. Barange, M. Beaugrand, G. et al (2008) Regime Shifts in Marine Ecosystems: Detection, Prediction and Management. Trends in Ecology and Evolution, v23, 7.

Ditlevsen, P. Johnsen, S. (2010) Tipping Points: Early Warning and Wishful Thinking. Geophysical Research Letters, v37, L19703.

DoC (2012). Defense Industrial Base Assessment: Counterfeit Electronics. Report. Washington, DC: Author (Dept. of Commerce).

DoD (2014). Defense Federal Acquisition Regulation Supplement: Detection and avoidance of counterfeit electronic parts (DFARS Case 2012-D055). Federal Register, 79, 26091-26108.

DoD (2013). DoD Counterfeit Prevention Policy. DoD Instruction 4140.67. Washington, DC: Author (Dept. of Defense).

DoD (2011). DoD Supply Chain Materiel Management Policy. DoD Instruction 4140.01. Washington, DC: Author (Dept. of Defense).

DoD (2012). Protection of Mission Critical Functions to Achieve Trusted Systems and Networks (TSN). DoD Instruction 5200.44. Washington, DC: Author (Dept. of Defense).

Duran, S., Gutierrez, M., & Keskinocak P. (2011). Pre-positioning of emergency items worldwide for CARE International. Interfaces, 41(3), 223-237.

Economist (2012). Huawei: the company that spooked the world. Author, August 4, 2012.

Ekici, A., Keskinocak, P., & Swann, J. (2014). Modeling influenza pandemic and planning food distribution. *Manufacturing and Service Operations Management*, 16(1), 11–27.

Eldred, M. S., Swiler, L. P., & Tang, G. (2011). Mixed aleatory-epistemic uncertainty quantification with stochastic expansions and optimization-based interval estimation. *Reliability Engineering & System Safety*, *96*(9), 1092-1113.

Ergun, O., Guyi, L., Heier-Stamm, J. L., Keskinocak, P., & Swann, J. (2014). Improving humanitarian operations through collaboration. *Production and Operations Management*, 23(6), 1002-1014.

Ergun, O., Heier-Stamm, J. L., Keskinocak, P., & Swann, J. (2010). Waffle House restaurants hurricane response: a case study. *International Journal of Production Economics*, 126(1), 111-120.

Fry, J. (2012) Exogneous and engdogenous market crashes as phase transitions in complex financial systems. The European Physical Journal B, v85(405).

GAO (2012a). Additional Efforts Needed by National Security-Related Agencies to Address Risks. Report GAO-12-579T. Washington, DC: Author (Government Accountability Office).

GAO (2010). DoD should leverage ongoing initiatives in developing its program to mitigate risk of counterfeit parts. Report GAO-10-389. Washington, DC: Author (Government Accountability Office).

GAO (2011). Periodic Assessment Needed to Correct Parts Quality Problems in Major Programs. GAO-11-404. Washington, DC: Author (Government Accountability Office).

GAO (2012b). Suspect counterfeit electronic parts can be found on internet purchasing platforms. Report GAO-12-375. Washington, DC: Author (Government Accountability Office).

Geels, FW. (2005) The Dynamics of Transition in Socio-technical Systems: A Multi-Level Analysis of the Transition Pathway from Horse-drawn Carriages ot Automobiles 1860-1930. Technology Analysis & Strategic Management, v17, 4, 445-476.

Guin, U., DiMase, D., & Tehranipoor, M. (2014). Counterfeit integrated circuits: detection, avoidance, and the challenges ahead. *Journal of Electronic Test*.

Guo, S. Seth, A. Kendrick, K. Zhou, C. Feng, J. (2008) Partial Granger causality - Eliminating Exogenous Inputs and Latent Variables. Journal of Neuroscience Methods, v172, 79-93.

Haimes, Y. Y. (1981). Hierarchical holographic modeling. *Systems, Man and Cybernetics, IEEE Transactions on*, *11*(9), 606-617.

Hall, A. D. (1989). *Metasystems methodology: a new synthesis and unification*. Pergamon.

Harvey, D.L. and M. Reed, (1996), Social Science as the Study of Complex Systems, in L.D. Kiel and E. Elliot (Eds.), *Chaos Theory in the Social Sciences: Foundations and Applications*, Ann Arbor: The University of Michigan Press.

Hlavackova-Schindler, K. Palus, M., Vejmelka, M. Bhattacharya, J. (2007) Causality Detection Based on Information-Theoretic Approaches in Time Series Analysis. Physics Report, v441, 1-46.

HNTB Corporation, (2010). “Atlanta regional managed lane system plan final report,” http://www.dot.ga.gov/BuildSmart/Studies/ManagedLanesDocuments/FINALREPORT.pdf (accessed 10/16/15).

Ives, A.R. & Dakos, V. (2012) Detecting dynamical changes in nonlinear time series using locally linear state-space models. Ecosphere, 3, art 58.

Janson, M., & Levinson, D. (2014). HOT or not: Driver elasticity to price on the MnPASS HOT lanes. *Research in Transportation Economics*, *44*, 21-32.

Kahneman, D., & Tversky, A. (1979). Prospect theory: An analysis of decision under risk. *Econometrica: Journal of the Econometric Society*, 263-291.

Kantanantha, N., Serban, N., & Griffin, P. (2010). Yield and price forecasting for stochastic crop decision planning. *Journal of agricultural, biological, and environmental statistics*, *15*(3), 362-380.

Kéfi S, et al. (2014) Early Warning Signals of Ecological Transitions: Methods for Spatial Patterns. PLoS ONE 9(3)

Kraskov, A. Stogbauer, H. Grassberger, P. (2004) Estimating Mutual Information. Physical Review E, v69, 0006138-1.

Liu, C., Rouse, W. B., & Yu, Z. (2015). When Transformation Fails: Twelve Case Studies in the American Automobile Industry. *Journal of Enterprise Transformation*, *5*(2), 71-112.

Livingston, H. (2014). Avoiding counterfeit electronic components – myths and unreliable facts. Presentation. BAE Systems.

Livingston H. (2007a). Avoiding counterfeit electronic components. IEEE Transactions on Component Packaging Technology, 30, 187-189.

Livingston, H. (2007b). Avoiding counterfeit electronic components – Part 2: Observations from recent counterfeit detection experiences. Report. BAE Systems.

Maydeu-Olivares, Albert. Joe, Harry (2005) Limited- and Full-Information Estimation and Goodness-of-Fit Testing in 2n Contingency Tables: A Unified Framework. Journal of the American Statistical Association, v100, 471.

Maydeu-Olivares, Alberto & Joe, Harry (2008) An Overview of Limited Information Goodness-of-Fit Testing in Multidimensional Contingency Tables. International Meeting of Psychometrics Society, 253-262.

McFadden, F. E., & Arnold, R. D. (2010). Supply chain risk mitigation for IT electronics. *Proceedings of the IEEE International Conference on Technologies for Homeland Security*, 49-55.

Moroney, J. D. P., Pezard, S., Miller, L. E., Engstrom, J., & Doll, A. (2013). Lessons from Department of Defense Disaster Relief Efforts in the Asia-Pacific Region. Santa Monica, CA: RAND Corporation. http://www.rand.org/pubs/research_reports/RR146.html

Mullen, F. E. (2013). Dynamic Multilevel Modeling Framework Phase 1 -- Feasibility. Washington, DC: Department of Defense, Modeling and Simulation Coordination, Office.

Oberkampf, W. L., DeLand, S. M., Rutherford, B. M., Diegert, K. V., & Alvin, K. F. (2002). Error and uncertainty in modeling and simulation. *Reliability Engineering & System Safety*, *75*(3), 333-357.

Olszewski, P., & Xie, L. (2005). Modelling the effects of road pricing on traffic in Singapore. *Transportation Research Part A: Policy and Practice*, *39*(7), 755-772.

Pan, X., & Zuo, Z. (2014). A Stochastic User Equilibrium Model and Optimal Congestion Pricing with Prospect Theory. *Procedia-Social and Behavioral Sciences*, *138*, 127-136.

Park, H., T. Clear, W. B. Rouse, R. C. Basole, M. L. Braunstein, K. L. Brigham, and L. Cunningham, (2012), Multilevel Simulations of Health Delivery Systems: A Prospective Tool for Policy, Strategy, Planning, and Management, *Service Science*, 4(3): 253-268.

Pecht, M., & Tiku, S. (2006). Bogus: electronic manufacturing and consumers confront a rising tide of counterfeit electronics. *IEEE Spectrum*, 43, 37-46.

Pennock, M. J. & Gaffney, C. (2016). Managing Epistemic Uncertainty for Multi-models of Socio-technical Systems for Decision Support, under review, *IEEE Systems Journal*.

Pennock, M. J., & Rouse, W. B. (2016), The epistemology of enterprises, in press *Systems Engineering*.

Pennock, M. J., & Rouse, W. B. (2014). The Challenges of Modeling Enterprise Systems. In *Proceedings of the 2014 International Engineering Systems Symposium*.

Pennock, M. J., & Rouse, W. B. (2014). Why connecting theories together may not work: How to address complex paradigm-spanning questions. In *Systems, Man and Cybernetics (SMC), 2014 IEEE International Conference on* (pp. 373-378). IEEE.

Pennock, M. J., Rouse, W. B. Bodner, D. A. Gaffney, C., Oghbaie, M., & Prasad, P. (2015) Enterprise Systems Analysis Technical Report SERC-2015-TR-020-4, Systems Engineering Research Center, Stevens Institute of Technology, Hoboken, NJ.

Perretti, C. Munch, S. (2012) Regime shift indicators fail under noise levels commonly observed in ecological systems. Ecological Application, v22(6), 1772-1779.

Pirnie, B. R., & Francisco, C. M. (1998). Requirements for Peacekeeping, Humanitarian Assistance, and Disaster Relief. Santa Monica, CA: RAND Corporation.

Rasiel, E. B., Weinfurt, K. P., & Schulman, K. A. (2005). Can prospect theory explain risk-seeking behavior by terminally ill patients?. *Medical Decision Making*, *25*(6), 609-613.

Reiser, M (1996) Analysis of Residuals for the Multinomial Item Response Model. Psychometrika, v61, 3, 509-528.

Reiser, M (2008) Goodness-of-fit Testing using Components Based on Marginal Frequencies of Multinomial Data. British Journal of Mathematical and Statistical Psychology, 61, 331-360.

Rhodes, D. H., & Ross, A. M. (2015). Interactive Model-Centric Systems Engineering (IMCSE): Pathfinder Workshop Report. Technical report, Systems Engineering Research Center, Stevens Institute of Technology, Hoboken, NJ.

Rodrigues, S. Barton, D. Marten, F. (2010) A method for detecting false bifurcations in dynamical systems: application to neural-field models. Biological Cybernetics, v102, 145-154.

Rosen, R. (1978). *Fundamentals of measurement and representation of natural systems* (p. 221). New York: North-Holland.

Rouse. W.B. (2012). *State of Modeling and Simulation in Commercial Industries*. Atlanta, GA: Tennenbaum Institute, Georgia Institute of Technology.

Rouse, W.B. (2015). *Modeling and Visualization of Complex Systems and Enterprises: Explorations of Physical, Human, Economic, and Social Phenomena.* Hoboken, NJ: Wiley.

Rouse, W. B., & Bodner, D. A. (2013). Multi-Level Modeling of Complex Socio-Technical Systems. Technical Report CCSE-2013-01. Hoboken, NJ: Center for Complex Systems and Enterprises, Stevens Institute of Technology.

Rouse, W.B. & Pennock, M.J. (2013). Multi-Level Modeling of Complex Socio-Technical Systems. Technical Report CCSE-2013-02. Hoboken, NJ: Center for Complex Systems and Enterprises, Sevens Institute of Technology.

Rouse, W.B., M.J. Pennock., M. Oghbaie, and C. Liu. (2016) Interactive Visualizations For Decision Support: Application Of Rasmussen's Abstraction-Aggregation Hierarchy, Under Review, *Journal of Applied Ergonomics.*

Roy, C. J., & Oberkampf, W. L. (2011). A comprehensive framework for verification, validation, and uncertainty quantification in scientific computing. *Computer Methods in Applied Mechanics and Engineering*, *200*(25), 2131-2144.

SAE (2014). Compliance verification criterion standard for SAE AS6081, fraudulent/counterfeit electronic parts: avoidance, detection, mitigation, and disposition – distributors. http://standards.sae.org/wip/as6301/. Retrieved 6/4/2014.

Sandia National Laboratory (2015), Dakota, https://dakota.sandia.gov/ (last accessed 3/1/15).

SASC (2012). Inquiry into counterfeit electronic parts in the Department of Defense supply chain. Washington, DC: Author (Senate Armed Services Committee).

Scheffer, M., et al (2009) Early-warning signals for critical transitions. Nature, 461, 3.

Schreiber, T. (2000) Measuring Information Transfer. Physical Review Letters, v85, 2.

Seekell DA, Carpenter SR, Pace ML (2011) Conditional heteroscedasticity as a leading indicator of ecological regime shifts. American Naturalist 178: 442-451.

Solé, R.V. (2011). *Phase Transitions*. Princeton University Press.

Sornette, D. (1998) Discrete Scale Invariance and Complex Dimension. Physics Reports 297, 239-270.

Stradley, J., & Karraker, D. (2006). The electronic part supply chain and risks of counterfeit parts in defense applications. *IEEE Transactions on Components Packaging Technology*, 29, 703-705.

Tolk, A. (2003). The levels of conceptual interoperability model. *Proceedings of the Fall Simulation Interoperability Workshop*. Orlando, Florida, September.

Tolk, A. (Ed.).(2013). *Ontology, Epistemology, and Teleology for Modeling and Simulation*. Berlin Heidelberg: Springer-Verlag.

Tolk, A. (2015). Learning Something Right from Models That Are Wrong: Epistemology of Simulation. In *Concepts and Methodologies for Modeling and Simulation* (pp. 87-106). Springer International Publishing.

Tversky, A., & Kahneman, D. (1992). Advances in prospect theory: Cumulative representation of uncertainty. *Journal of Risk and uncertainty*, *5*(4), 297-323.

Vallacher, R. & Nowak, A. (1997) The Emergence of Dynamical Social Psychology. Psychological Inquiry, v8, 2, 73-99.

Villasenor, J., & Tehranipoor, M. (2013). Chop-shop electronics. *IEEE Spectrum*, 50, 40-45.

Walters, C.J. (1997) Challenges in adaptive management of riparian and coastal ecosystems. Conserv Ecol. 1, 1.

Wang, R. (2012) Flickering gives early warning signals of a critical transition to a eutrophic lake state. Nature, v492, 419-422.

Wilson, J.A. (2002) Matching social and ecological system in complex ocean fisheries. Ecological Society, 11, 9.

Winsberg, E. (2010). *Science in the age of computer simulation*. University of Chicago Press.

Xu, H., Lou, Y., Yin, Y., & Zhou, J. (2011). A prospect-based user equilibrium model with endogenous reference points and its application in congestion pricing. *Transportation Research Part B: Methodological*, *45*(2), 311-328.

Yan, W. Woodard, R. Sornette, D. (2010) Diagnosis and Prediction of Tipping Points in Financial Markets: Crashes and Rebounds. Physics Procedia v3, 1641-1657.

Yao, W., Chen, X., Luo, W., van Tooren, M., & Guo, J. (2011). Review of uncertainty-based multidisciplinary design optimization methods for aerospace vehicles. *Progress in Aerospace Sciences*, *47*(6), 450-479.