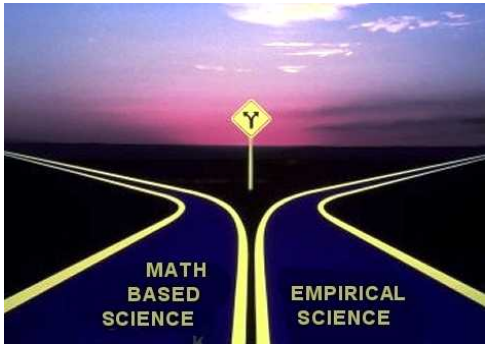
SAND2016-5906C

Exceptional service in the national interest

Sandia National Laboratories



Model/Use Characteristics that Can Challenge Current VVUQ : Building Credible Models

June 22, 2016

Stephen J. Verzi and Asmeret Naugle

U.S. DEPARTMENT OF ENERGY
National Nuclear Security Administration

Sandia National Laboratories is a multi-program laboratory managed and operated by Sandia Corporation, a wholly owned subsidiary of Lockheed Martin Corporation, for the U.S. Department of Energy's National Nuclear Security Administration under contract DE-AC04-94AL85000. SAND NO. 2011-XXXXP

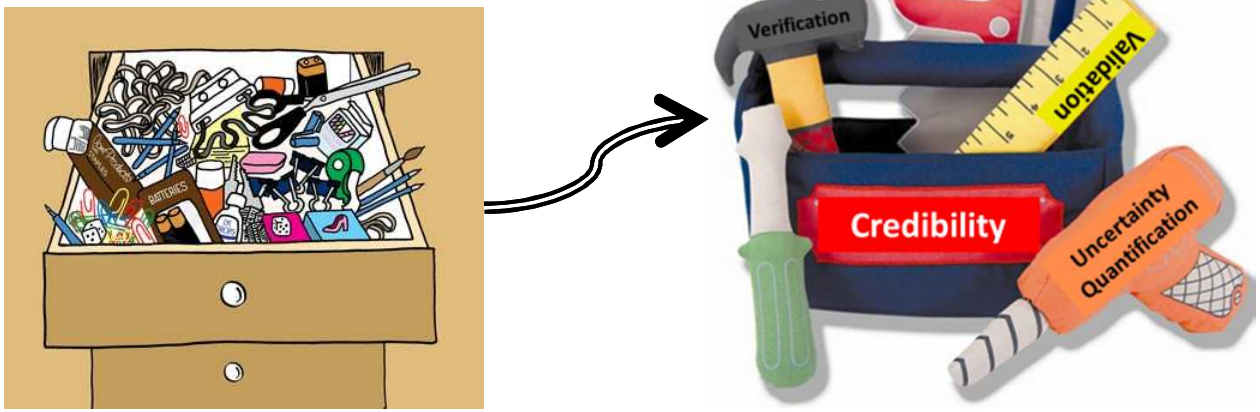
Our technical team

Sandia National Laboratories

- Jeffrey Tsao
- Asmeret Naugle
- Mark Smith
- Tatiana Flanagan
- Kasimir Gabert
- Matthew Lave
- Eric Vugrin
- Sasha Outkin
- Curtis Johnson
- Timothy Trucano
- Laura Swiler
- Laura McNamara
- George Backus
- Patrick Finley
- Dean Dobranich
- Timothy Wildey
- Brian Adams
- William Rider

Outline

- Modeling process
- Characteristics of complex systems
- Purpose of VVUQ in modeling complex systems
- Model uses for complex systems
- When can modeling challenge current VVUQ?
- Opportunities for further research
- Summary



Modeling process

Model use as well as validation will likely lead to refinement or even new generation of problem questions

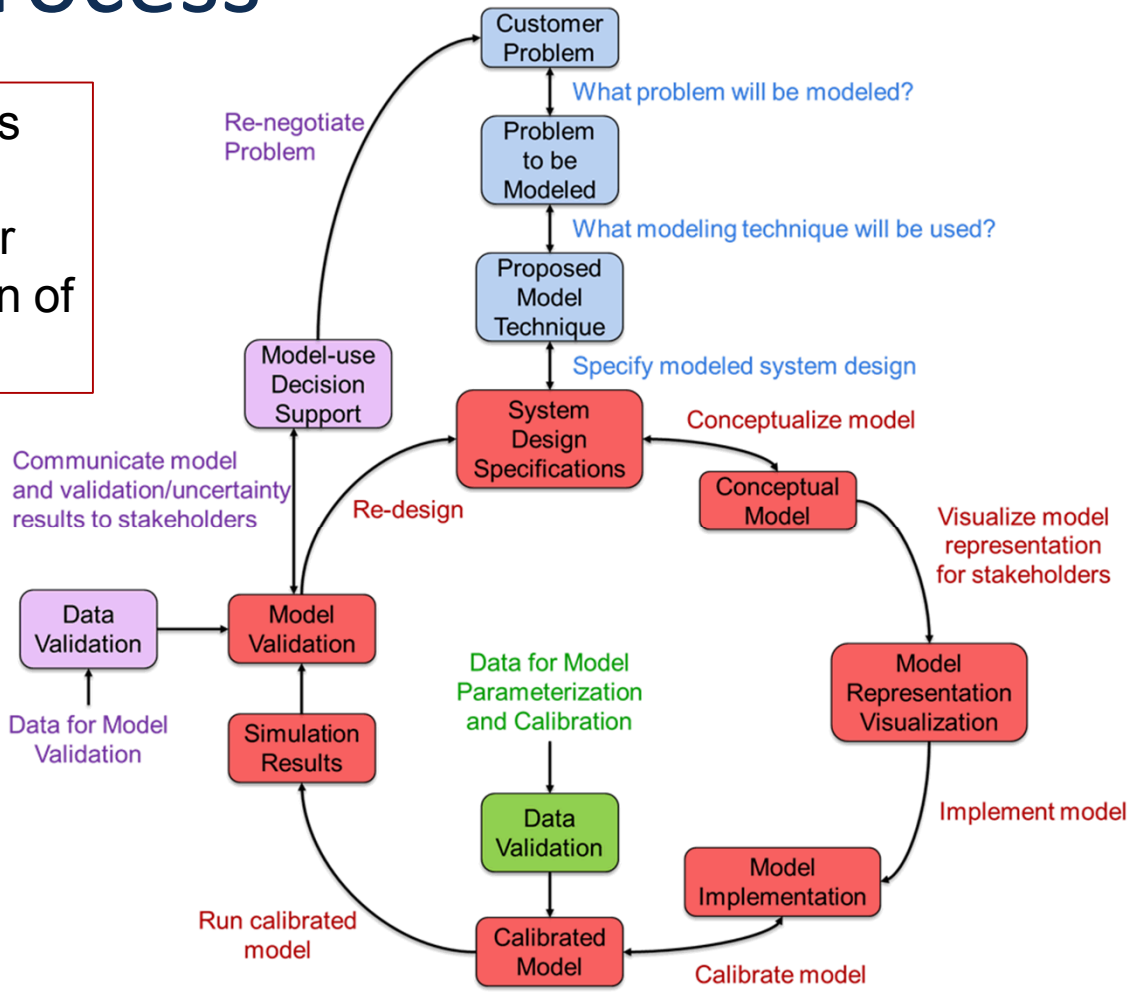


*Diagram is a modification of Balci modelling process diagram (Balci, 1998).

Modeling is an iterative process to build confidence and understanding.

Model uses

Sandia National Laboratories

According to Shannon (Shannon, 1975)

- Evaluation of system behavior
- Forecasting
- Comparison of different operating policies
- Optimization
- Sensitivity analysis
- Determination of functional relationships
- Training

Formal model validation

$$E = S - D$$
$$E = (T + \delta_S) - (T + \delta_D) = \delta_S - \delta_D$$
$$E = \delta_{\text{model}} + \delta_{\text{numerical}} + \delta_{\text{input}} - \delta_D$$
$$\delta_{\text{model}} = E - \left(\delta_{\text{numerical}} + \delta_{\text{input}} - \delta_D\right)$$
$$u_{\text{validation}} = \sqrt{\delta_{\text{numerical}}^2 + \delta_{\text{input}}^2 + \delta_D^2}$$
$$\text{if } |E| \leq u_{\text{validation}} \text{ then "valid"}$$

Formal validation: compare simulation results with experimental or observational data

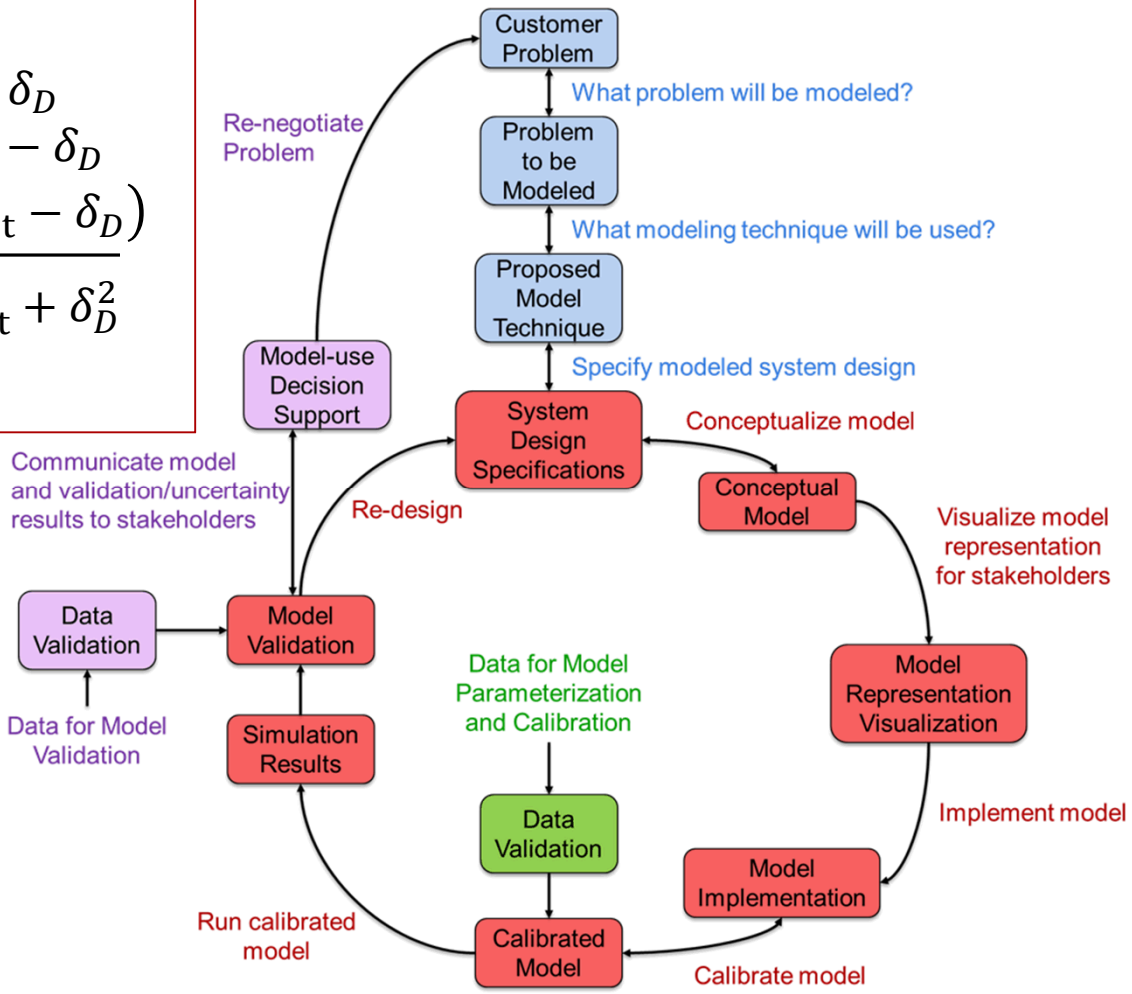


*Diagram is a modification of Balci modelling process diagram (Balci, 1998).

Model error within noise level imposed by numerical, parameter input and experimental data uncertainties → "valid".

Characteristics of complex systems

Sandia National Laboratories

- adaptive or self-organized behavior
- high throughput
- heterogeneity of subcomponents
- multi-scale interaction
- bifurcations and phase change
- cascading and/or emergent behavior
- synergistic components
- feedback loops
- non-linearity
- humans in the loop
- lack of established theory and/or unknown basic physical laws
- feedback from model to system
- inability to conduct experiments, lack of data and/or low signal-to-noise
- reliance upon soft quantity data
- out of equilibrium
- results and assessment focus on dynamics or dynamical behavior
- reductionist approach is inappropriate (irreducible)
- open system
- imbalanced information exchange
- exhibits power laws
- multi-objective behavior
- non-locality
- complementary quantities of interest
  - Heisenberg uncertainty principle
- inability to specify closed-form description but can be simulated
- potentially unpredictable
- social dynamics

List is incomplete, non-orthogonal and not definitive or universally agreed upon.

Purpose of VVUQ in modeling complex system

- Continue to use formal validation where possible
- Build confidence in modeling process where formal validation cannot be applied

Includes stakeholders and potential stakeholders

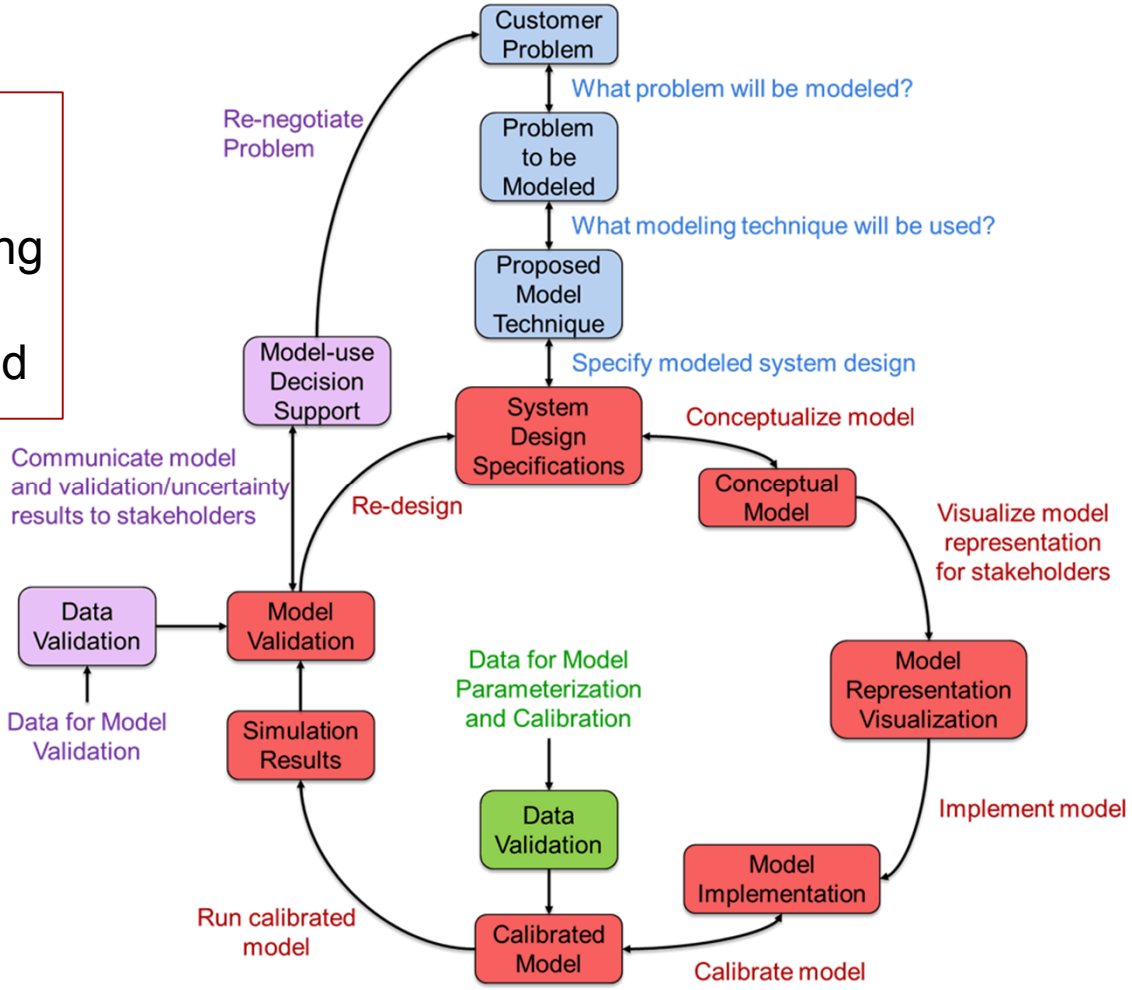


*Diagram is a modification of Balci modelling process diagram (Balci, 1998).

Understand and quantify credibility in the use of complex systems modeling approaches with respect to a particular use.

Model uses for complex systems

Sandia National Laboratories

Key model/use factors to be addressed

- What is the question of interest (knowledge of interest)?
- Which model is appropriate for answering this question?

Potential model uses

- Prediction
- Policy exploration
- Risk analysis
- Empirical demonstration of system theory
- Scientific exploration of alternate system theories
- Real-time operation (feedback & control)

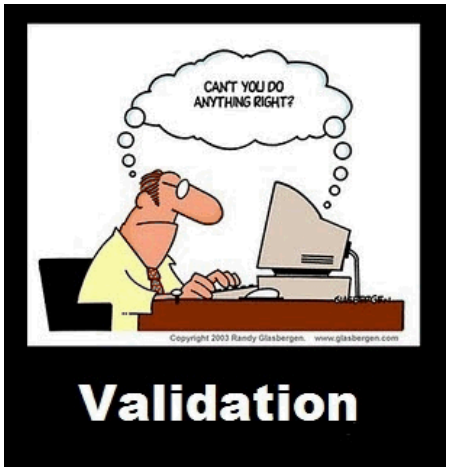
Model/use credibility

- Do we have the right model for our intended use?
  - Engineered systems – "Assurance that a product, service, or system meets the needs of customers and other identified stakeholders"
  - Software – "Process of evaluating software design and implementation to determine whether it satisfies specified requirements"

ISO/IEC 15288

IEEE SA 1012

- Are we asking the right question of our model?
  - Customers do not always know what they want
  - Customers need Y but ask for X
  - Have model Z that we want to apply
  - Want to research/build model W



Quote from Swiler, 2016.

"Perform due diligence and communicate frankly about assumptions, approximations, and limitations affecting simulation credibility."

Prediction is hard:

- *Limited* physical data (observational or experimental)
- *Limited* simulations (high computational demands…)
- *Imperfect* computational models (missing physics, etc.)
- *Under-resolved* approximations or numerics
- *Unknown* model parameters and boundary conditions
- *Imperfect* humans
- We want to *extrapolate* to conditions beyond validation regime…

Slide from Swiler, 2016.

When can modeling challenge current VVUQ?

Sandia National Laboratories

- Limited (observational or experimental) data
  - Modeled system is unpredictable
- Additional data challenges
- Imperfect models
  - Computation and/or physics
  - Complex behavior dynamics
- Limited model simulations (time & cost)

Current approaches to assess model credibility

- Limited data
- Additional data challenges
- Imperfect models
- Extrapolating beyond validated (validate-able) regime

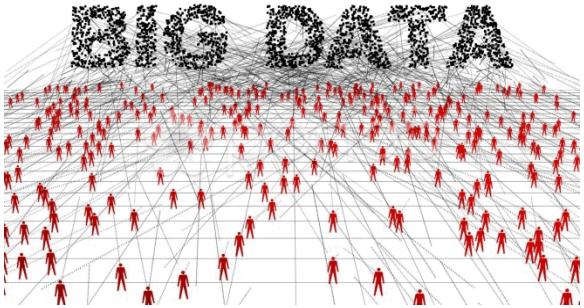
Limited data

- inability to conduct experiments
  - detonate “dirty bomb” in NYC
  - natural disaster
- imbalanced information
  - stealth
  - one-sided learning
- Modeled system is unpredictable

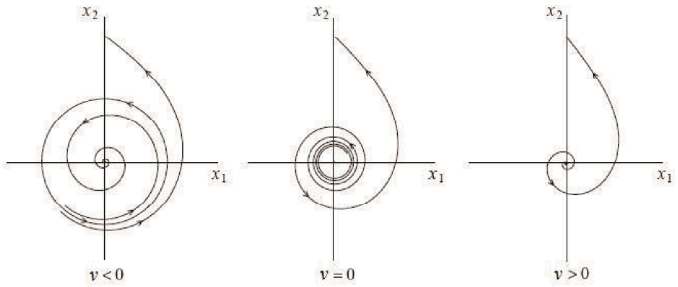
Need to understand potential effects we cannot create ourselves.

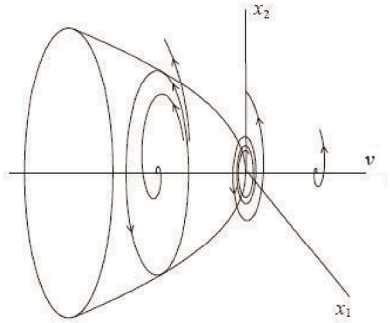
Additional data challenges

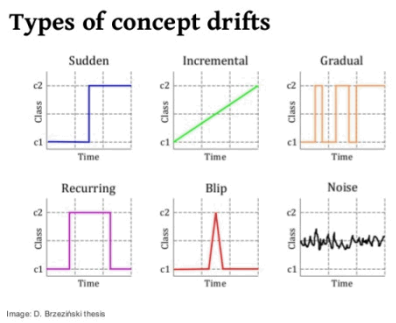
- Data quality
- Dimensionality
  - heterogeneity of subcomponents
- Too much data
  - multi-scale interactions
  - low signal-to-noise
- Continuous dynamic data
  - bifurcations and phase change
  - out of equilibrium



- Volume
- Velocity
- Variety
- Value
- Visibility







Imperfect models

Computation and/or physics

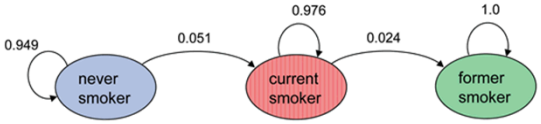
- non-linearity
- synergistic components
- open system
- reductionist approach is inappropriate (irreducible)
- lack of established theory and/or unknown basic physical laws

Exploratory modeling

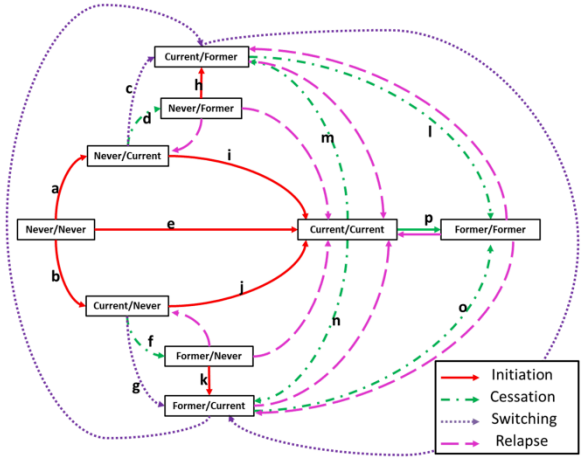
Complex behavior dynamics

- adaptive or self-organized behavior
- cascading and/or emergent behavior
- multi-objective behavior
- social dynamics
- humans in the loop

Re-negotiate problem



Verzi et al., 2012.



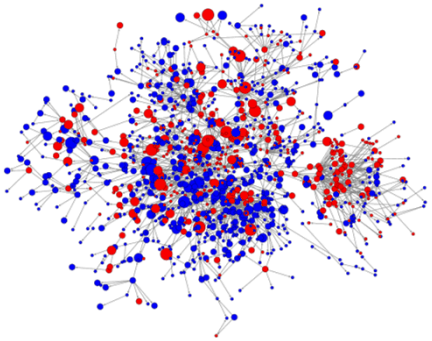
Vugrin et al., 2015.

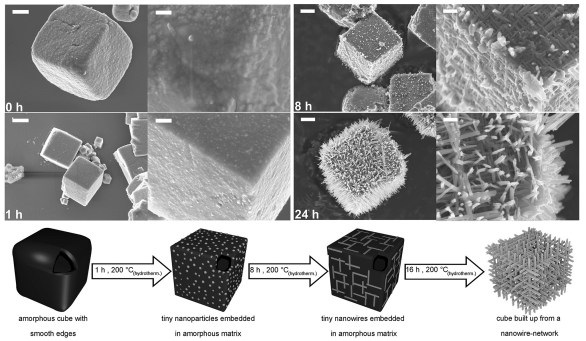
Limited model simulations

- high throughput
  - turbulent flow

- complex behavior dynamics

  {cascading, emergent, self-organized, multi-objective, adaptive}
  - self-organizing materials
  - social dynamics



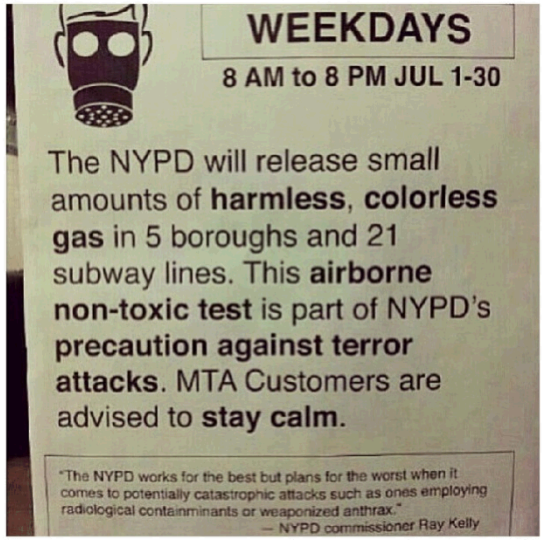


Need to understand potential effects that are too expensive to simulate completely (a large number of times).

Current approaches to assess model credibility

Limited data

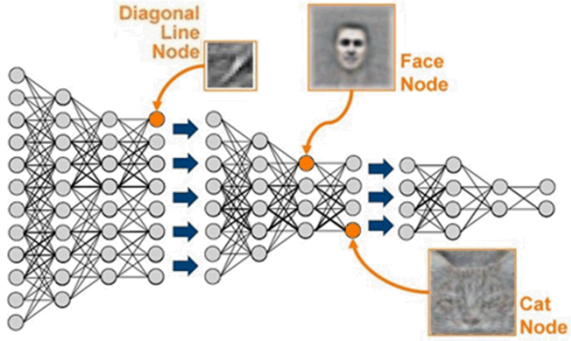
- Find surrogate source
- Simulate it
  - Use known data to design generative model
  - Use micro scale model to generate macro scale data
- Conduct our own experiment
  - Small sample size – biased?
- Survey subject matter experts (SMEs)
- Model without it
  - Use “best guess”

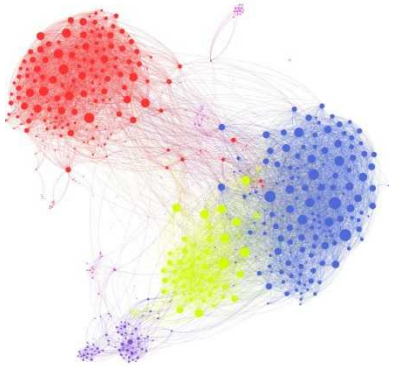


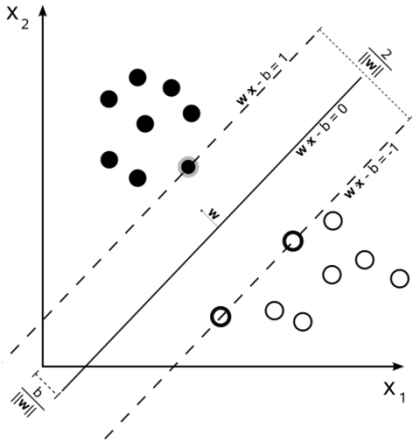
Current approaches to assess model credibility

Additional data issues

- Data quality V&V
- Dimension reduction
  - Feature extraction – deep learning
  - Principle components analysis
  - Compression
- Data clustering
- Machine learning







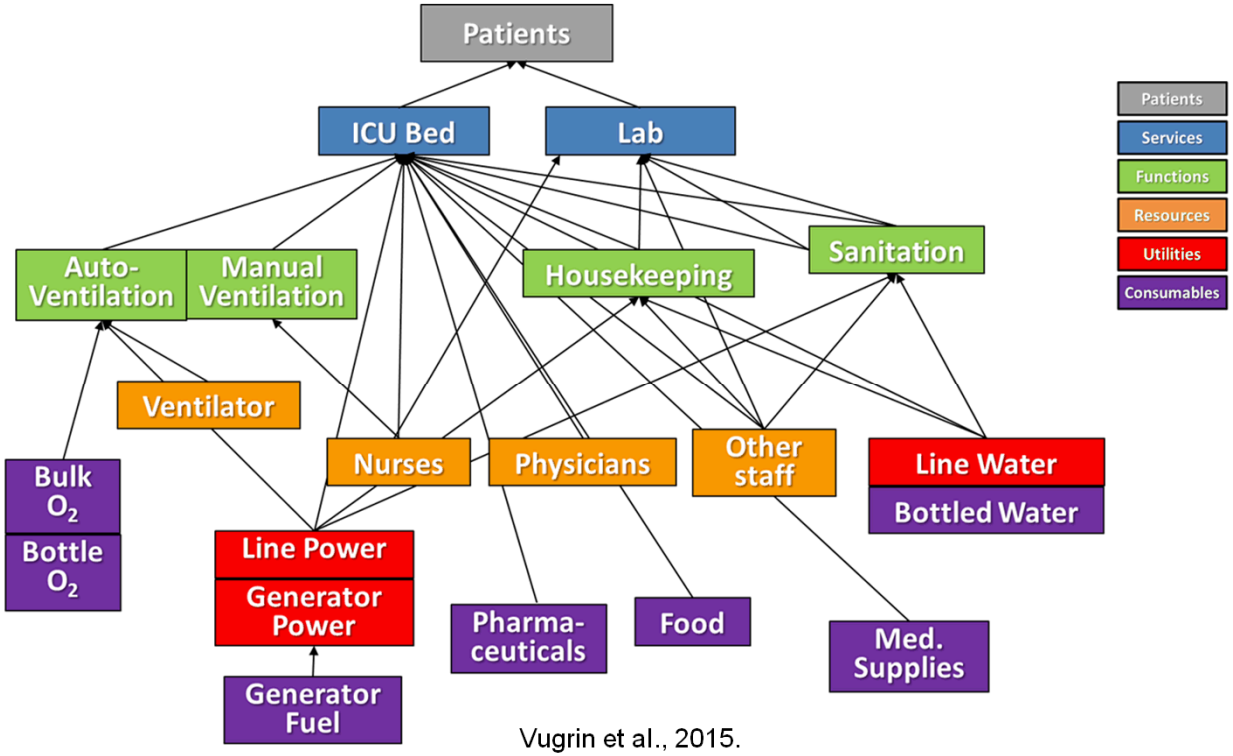
Current approaches to assess model credibility

Sandia National Laboratories

Model representation visualization

Imperfect models

- Adaptive resilience in hospital ICU



Vugrin et al., 2015.

Work with customers to derive insights from model for adaptive substitution and to build credibility by recreation of historical event.

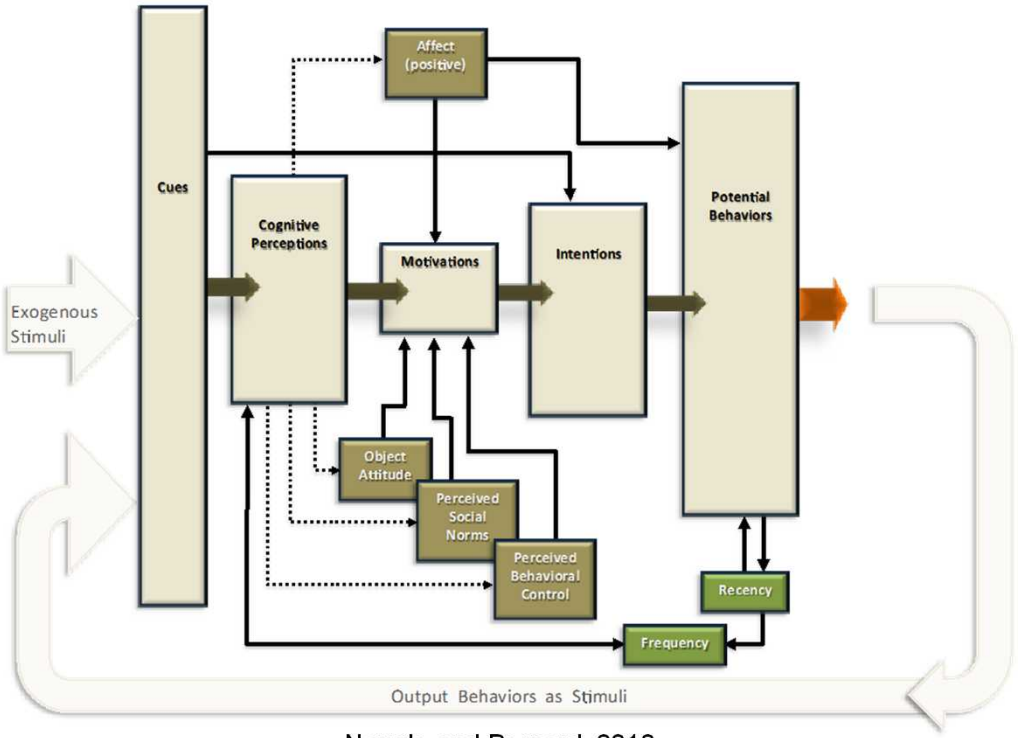
Current approaches to assess model credibility

Sandia National Laboratories

Model representation visualization

Imperfect models

- Complex socio-cognitive behavior



Naugle and Bernard, 2010.

Work with customers/stakeholders to derive insights from modeling process and to understand modeled behavior equations.

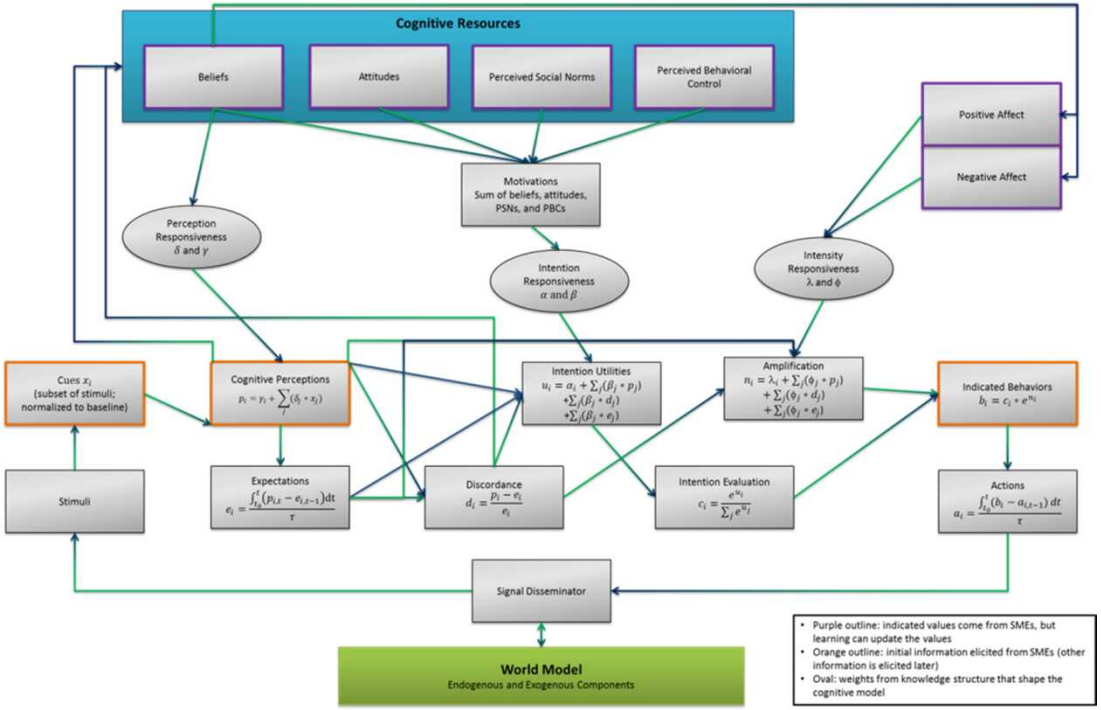
Current approaches to assess model credibility

Sandia National Laboratories

Model representation visualization

Imperfect models

- Complex socio-cognitive behavior



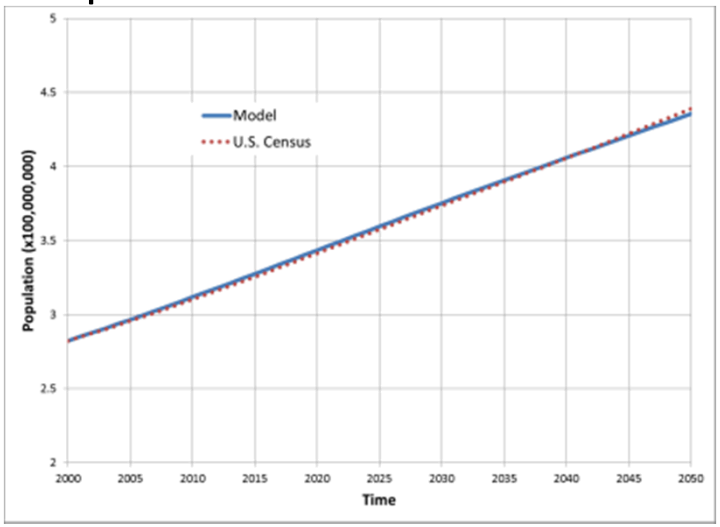
Naugle and Bernard, 2010.

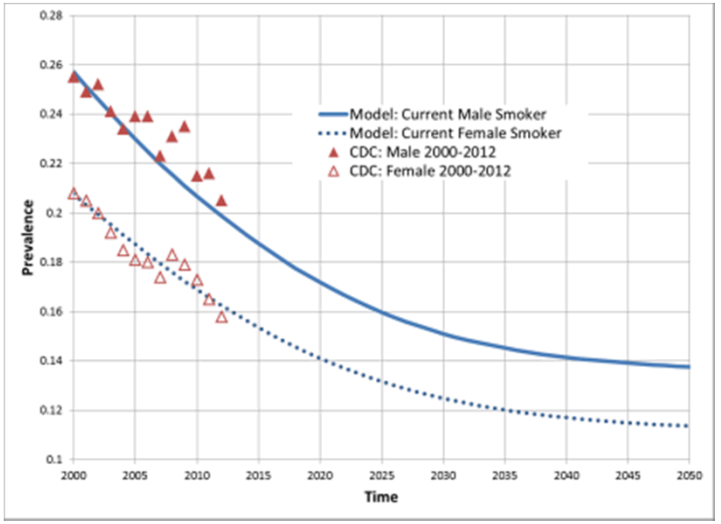
Work with customers/stakeholders to derive insights from modeling process and to understand modeled behavior equations.

Current approaches to assess model credibility

Extrapolating beyond validated (validate-able) regime)

- Model-to-model comparison
  - docking
  - co-validation (model verification)
- Independent review
  - subject matter experts
  - peers

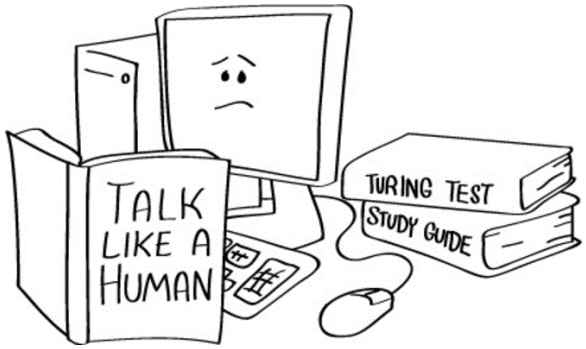




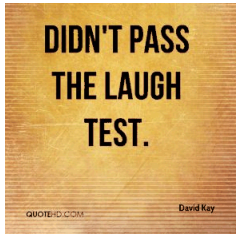
Vugrin et al., 2015.

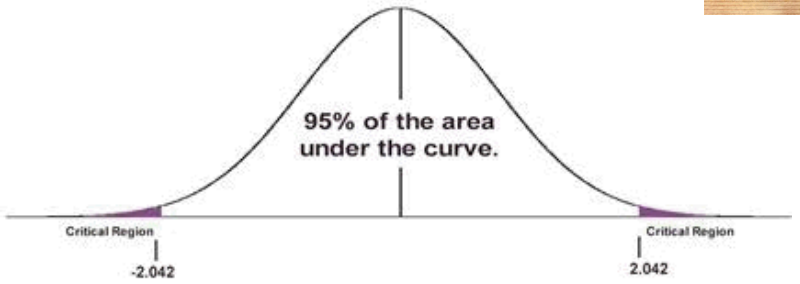
Opportunities for further research

- Lack of data
  - Use analogous, synthetic or expert-elicited data
  - Reproduction of
    - system theory
    - phenomena
    - behavior



- Extending model credibility
  - Face validation
  - Turing test
  - Extreme-value testing





Summary

- Modeling process
- Complex systems modeling characteristics
- Goal: understanding and quantifying model credibility
- VVUQ gaps
- Current approaches
- Opportunities

References

- Backus, G., Bernard, M., Verzi, S., Bier, A. and Glickman, M. 2010. "Foundations to the Unified Psycho-Cognitive Engine." *Sandia National Laboratories Technical Report*, SAND2010-6974.
- Balci, O. 1998. "Verification, Validation and Testing." (Chapter 10) In *Handbook of Simulation: Principles, Methodology, Advances, Applications, and Practice*, Banks, J. (Ed.), (Wiley Online Library) John Wiley & Sons, Inc., Hoboken, NJ. doi: 10.1002/9780470172445.ch10.
- Johnson, C.M., Backus, G.A., Brown, T.J., Colbaugh, R., Jones, K.A. and Tsao, J.Y. 2012. "A Case for Sandia Investment in Complex Adaptive Systems Science and Technology." *Sandia National Laboratories Technical Report*, SAND2012-3320.
- Naugle, A.B. and Bernard, M.L. 2014. "DYMATICA Modeling and Assessment Dynamic Multi-Scale Assessment Tool for Integrated Cognitive Behavioral Actions." *Sandia National Laboratories Technical Presentation*, SAND2014-15338PE.
- Rider, W.J., Kamm, J.R. and Weirs, V.G. 2010. "Verification, Validation and Uncertainty Quantification Workflow in CASL." *Sandia National Laboratories Technical Report*, SAND2010-234P.
- Sargent, R.G. 2011. "Verification and Validation of Simulation Models." *Proceedings of the 2011 Winter Simulation Conference*, Jain, S., Creasey, R.R., Himmelspach, J., White, K.P. and Fu, M. (Eds.): 183-198. http://dl.acm.org/citation.cfm?id=2431538.
- Shannon, R.E. 1975. *Systems Simulation: the Art and Science*, Prentice-Hall, Englewood Cliffs, NJ. doi: 10.1109/TSMC.1976.4309432.
- Sornette, D., Davis, A.B., Vixie, K.R., Pisarenko, V. and Kamm, J.R. 2007. "Algorithm for model validation: Theory and applications." PNAS, 104(16): 6562-6567. doi: 10.1073/pnas.0611677104.
- Verzi, S.J., Apelberg, B., Rostron, B. Brodsky, N., Brown, T.J., Husten, C. and Glass, R.J. 2012. "An Agent-Based Approach for Modeling Population Behavior and Health with Application to Tobacco Use." *Sandia National Laboratories Technical Report*, SAND2012-6898.
- Vugrin, E.D., Rostron, B.L., Verzi, S.J., Brodsky, N.S., Brown, T.J., Choiniere, C.J., Coleman, B.N., Paredes, A. and Apelberg, B.J. 2015. "Modeling the Potential Effects of New Tobacco Products and Policies: A Dynamic Population Model for Multiple Product Use and Harm." PLOS ONE, 10(3): e0121008. doi:10.1371/journal.pone.0121008.
- Vugrin E, Verzi S, Finley P, Turnquist M, Griffin A, Ricci K, Wyte-Lake T. 2015. "Modeling hospitals' adaptive capacity during infrastructure disruptions." *Journal of Healthcare Engineering, 6*(1): 85-120. PM ID: 25708379. doi: 10.1260/2040-2295.6.1.85.

Additional References

- Allen, E.H., Tallant, G.S. and Elliot, M.A. 2014. “Computer systems and methods for quantum verification and validation.” US Patent 8832165, *Google Patents*. US 8832165.
- Koulouriotis, D.E., Diakoulakis, I.E., Emiris, D.M. and Zopounidis, C.D. 2005. “Development of dynamic cognitive networks as complex systems approximators: validation in financial time series.” *Applied Soft Computing, 5*: 157-179.
- McElhaney, J. 2011. “No Laughing Matter: Failing the Giggle Test Might Leave You Crying.” ABA Journal. http://www.abajournal.com/magazine/article/the_giggle_test.

References

Sandia National Laboratories

Chronology of V & V for Complex Systems and Complex Adaptive Systems

- Wise, J.A., Hopkin, V.D. and Stager, P. (Eds.). 1993. *Verification and Validation of Complex Systems: Human Factors Issues, NATO ASI Series*, Springer, Heidelberg, Germany.
- Van Dijkum, C., De Tombe, D. and Van Kuijk, E. 1999. *Validation of simulation models*. Siswo Publication, Amsterdam.
- Kelton, W.D. and Law, A.M. 2000. *Simulation and modeling analysis*. McGraw Hill, Boston, MA.
- Stevenson, D.E. 2002. "Verification and Validation of Complex Systems." *Intelligent Engineering Systems through Artificial Neural Networks (2002)*, http://citeseerx.ist.psu.edu/viewdoc/download?doi=10.1.1.6.8571&rep=rep1&type=pdf.
- Küppers, G. and Lenhard, J. 2005. "Validation of simulation: Patterns in the social and natural sciences." *Journal of Artificial Societies and Social Simulation, 8*(4): 3. http://jasss.soc.surrey.ac.uk/8/4/3.html.
- Macal, C.M. and North, M.J. 2005. "Validation of an Agent-based Model of Deregulated Electric Power Markets." *Proceedings of the North American Computation Social Organization Science (NAACSOS) 2005 Conference*. doi: 10.1.1.72.6692.
- North, M.J., Hower, T.R., Collier, N.T. and Vos, J.R. 2007. "A declarative model assembly infrastructure for verification and validation." *Advancing social simulation: the first world congress*, Springer, Japan. 129-140. doi: 10.1007/978-4-431-73167-2_13.
- Windrum, P., Fagiolo, G. and Moneta, A. 2007. "Empirical validation of agent-based models: Alternatives and prospects." *Journal of Artificial Societies and Social Simulations, 10*(2): 8. http://jasss.soc.surrey.ac.uk/10/2/8.html.
- Omerod, P. and Rosewell, B. 2009. "Validation and Verification of Agent-Based Models in the Social Sciences." In *Epistemological Aspects of Computer Simulation in the Social Sciences: Second International Workshop, EPOS 2006*, Springer, Heidelberg, Germany: 130-140. doi:10.1007/978-3-642-01109-2_10.
- Bruch, E. and Atwell, J. 2013. "Agent-based models in empirical social research." *Sociological Methods & Research, 44*(2): 186-221. doi: 10.1177/0049124113506405.
- Luna, S, Lopes, A., Tao, H.Y.S, Zapata, F. and Pineda, R. 2013. "Integration, Verification, Validation, Test, and Evaluation (IVVT&E) Framework for Systems of Systems (SoS)." *Procedia Computer Science, 20*: 298-305. doi:10.1016/j.procs.2013.09.276.
- Bert, F.E., Rovere, S.L., Macal, C.M., North, M.J. and Podestá, G.P. 2014. "Lessons from a comprehensive validation of an agent-based model: The experience of the Pampas Model of Argentinean agricultural systems." Ecological modeling, 273: 284-298. doi: 10.1016/j.ecolmodel.2013.11.024.

Thank you

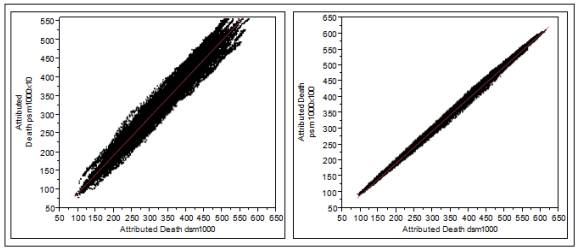
- Questions?

Backups

Sandia National Laboratories

Dealing with computationally intensive model simulations

- Employ more tractable mathematical/model representation (use an emulator)
  - continuous ←—→ discrete
  - mean behavior ←—→ individual behavior
  - use (very) complicated (not complex) modeling approach

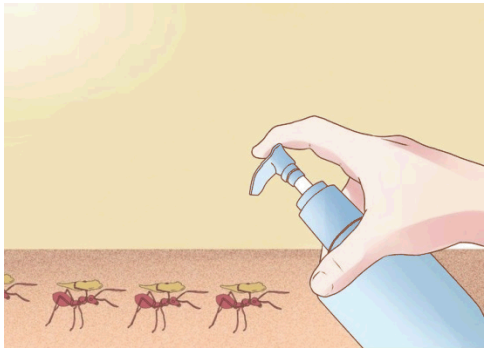


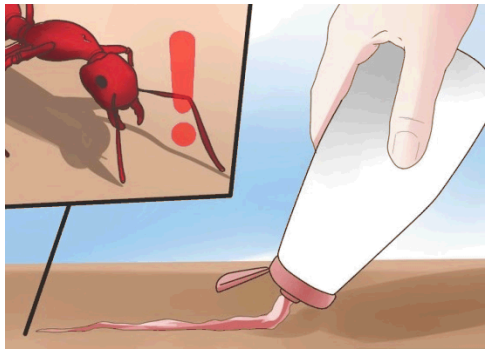
- Buy a bigger/faster computer

Modeling Interventions

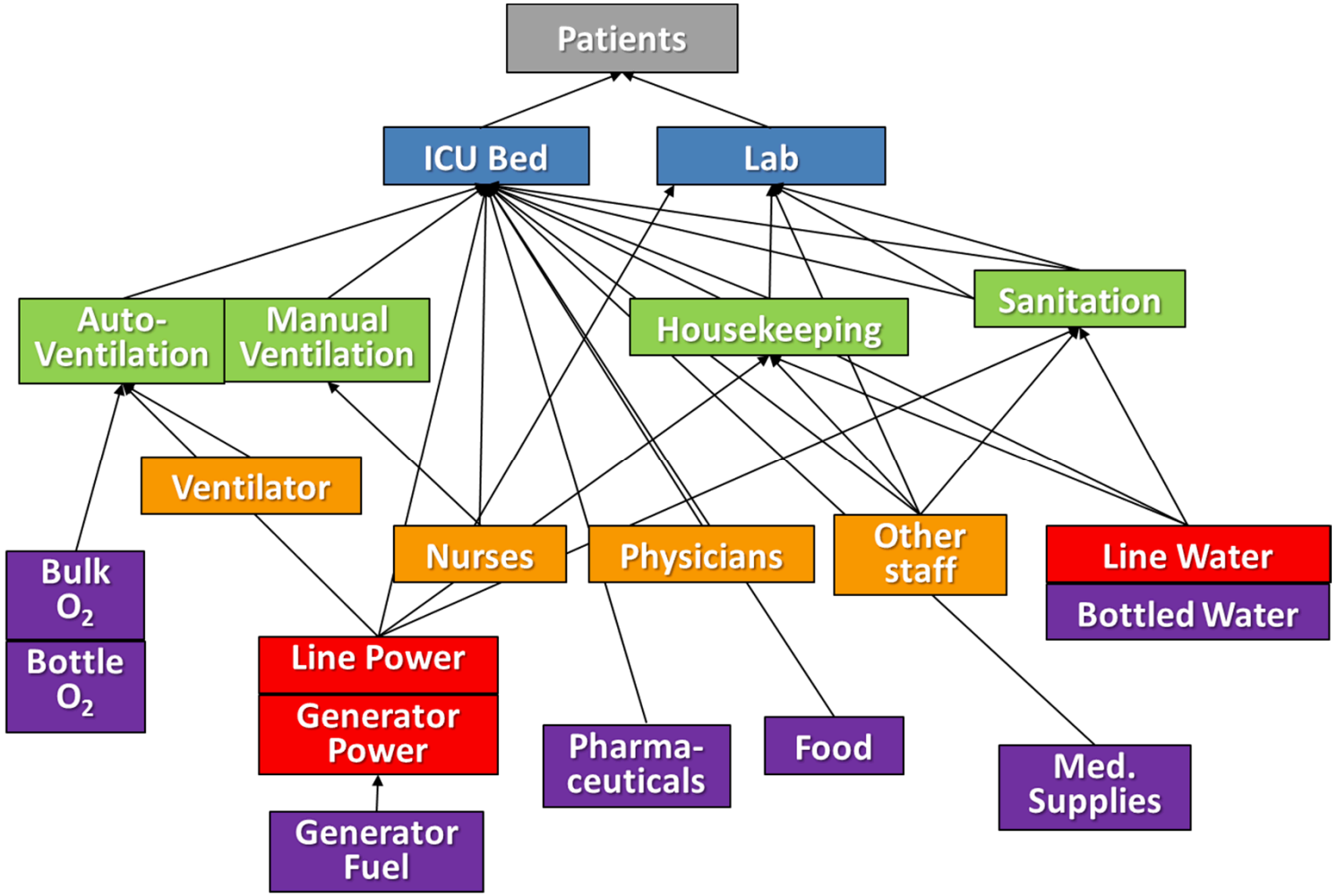
Example:

- Want to keep ants outside
- Possible interventions
  - Attempt to kill them
  - Build a barrier
  - Entice them to stay outside
  - …
  - Combinations
- How detailed does the ant model need to be?
- Can it be validated?





Hospital ICU resilience




Verification
Validation
Credibility
Uncertainty Quantification


Verification
Validation
Credibility
Uncertainty Quantification
Distinguisher