On Validating Multi-Agent System Applications

Charles M. Macal, PhD, PE, macal@anl.gov

Director, Center for Complex Adaptive Agent Systems Simulation (CAS^2),

Argonne National Laboratory

Senior Fellow, The Computation Institute, University of Chicago

Computation Institute

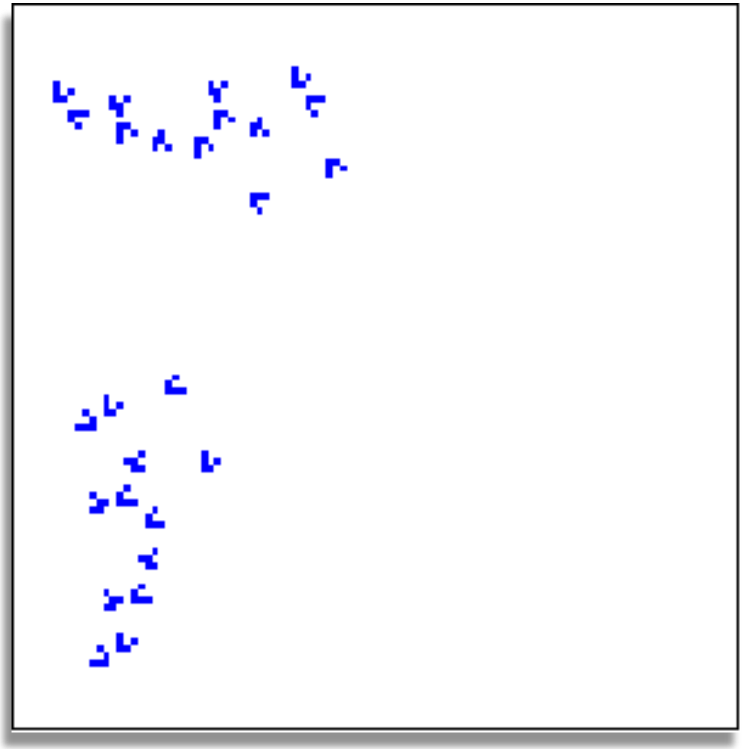
Keynote Address MABS 2013

14th International Workshop on Multi-Agent-Based Simulation

May 7, 2013

Multi-Agent Model Validation?

1. How should we ***think about*** model validation?
2. How should we ***do*** model validation?
3. What should we ***do about*** model validation?



3 Simple Rules.
Two sets of 13 **Life** *Gliders* create *Glider Guns* each of which creates a *Glider Stream*, which eventually destroy the *Glider Guns.*

A Polling Question

- What do **YOU** think that the main challenge of **MAS model validation** is (choose one):
  a. A research question worthy of **scientific research** and formal investigation?
  b. A matter of *doing* **good practice** that are already known, but perhaps seldom applied?
  c. Primarily about building **awareness and consensus** in the modeling community?

The Agent-Based "Name Game"

- **Agent**
  - "an autonomous, software entity with adaptive behaviors" or "someone who works for the FBI?"
- Agent modeling is known by many names:
  - **ABM**: "Agent-Based Modeling" or "Activity-Based Modeling"
  - **IBM**: "Individual-Based Modeling" or "International Business Machines?"
  - **MAS**: "Multi-Agent Systems"

A more formal view of agent-based modeling

- An Agent-based Model:

  ABS = {A, I, E, T}

  1. A set of agents
  2. I agent interaction space
  3. E environment independent of the agents
  4. T time advance mechanism

- An Agent:

  a = {B, S, D, N, M} $a \in A$

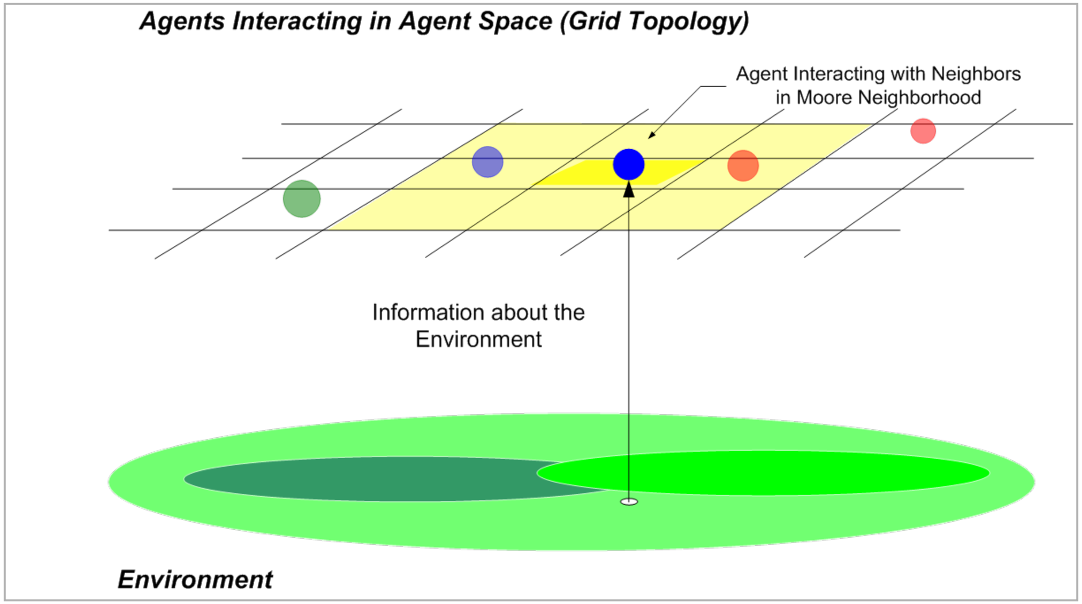
  B set of agent behaviors

  S set of static attributes for the agent

  D set of dynamic attributes for agent

  N agent's neighborhood

  M mechanisms for updating agent state



Validation in Modeling & Simulation

My Large-Scale Agent-based Modeling Applications

- *EMCAS:* A large-scale agent-based model of deregulated electric power markets
- *Virtual Learning Lab:* An ABM of the markets for consumer packaged goods and supporting supply chains
- *AgentCell:* A large-scale multi-level simulator for the motile behavior of *E. Coli*, in which agents are molecular and agents are bacteria cells
- *Mobilizing for Change*: A model for simulating political movements in armed conflicts, applied to Central Asia
- *PAH Model:* A multi-scale, 3D model of mammalian arterial blood vessels focusing on the pathobiology of Pulmonary Arterial Hypertension
- *SimIndia CVD:* A model that forecasts cardio-vascular disease for the population of a state in India
- *Pampas Model*: An ABM to explore the dynamics of structural and land use changes in agricultural systems of the Argentine Pampas
- *MRSA ABM:* A large-scale model of Community Associated MRSA transmission for Chicago
- *… and there's more …*

The Need for Validation

- Model validation is an essential part of the model *development* process -- *if* models are to be *accepted:*
  - *Used* to support policy- and decision- making, or
  - As *credible*, scientific representations of real-world phenomenon
- One of the very *first questions* that a person who is promoting a model or its results will be asked is:

"Has your model been validated?"

If the answer to this critical question is ***No***…

- Experience has shown that the model is unlikely to be adopted or even tried out in a real-world setting
- Often the model is sent "back to the drawing board"
  - Developers lose funding and are fired
  - It's a black eye for modeling in general, and we all are lose
- The challenge then becomes one of being able to say "yes" to the critical question:

"Yes! My model has been validated."

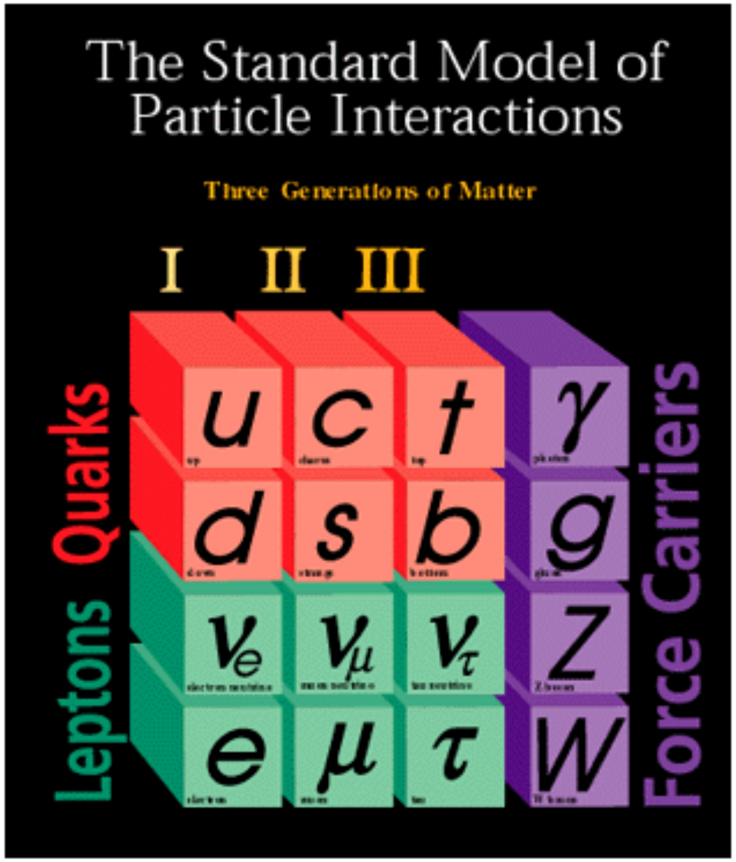
.. and let me tell you how…

This is what people usually are thinking of when they say "validation" …

The Model Validation "Gold Standard"

- In Physics, "The Standard Model" is the name given to the current theory of fundamental particles and their interactions

- The Standard Model is a good theory because it has been **validated**
  - Its predictions have matched experimental data, decimal place for decimal place, with amazing precision (17 sig. figs.)
  - All the particles predicted by this theory have been found
- Can such a theoretical model be found for social systems?



A Validation Test for a Model

- It would be great if we could write a universally applicable function, call it **CRRBRWS** for short:

  Correctly_Represents_and_Reproduces_the_Behaviors_of_the_Real_World_System[…]

- … that we could apply **CRRBRWS** to any **model** (or even any model in a restricted domain of application) to test whether the model is valid:

  ```
  ValidationQ = CRRBRWS[model] == True //??
  ```

- ... but we don't have one of those right now -- that I am aware of

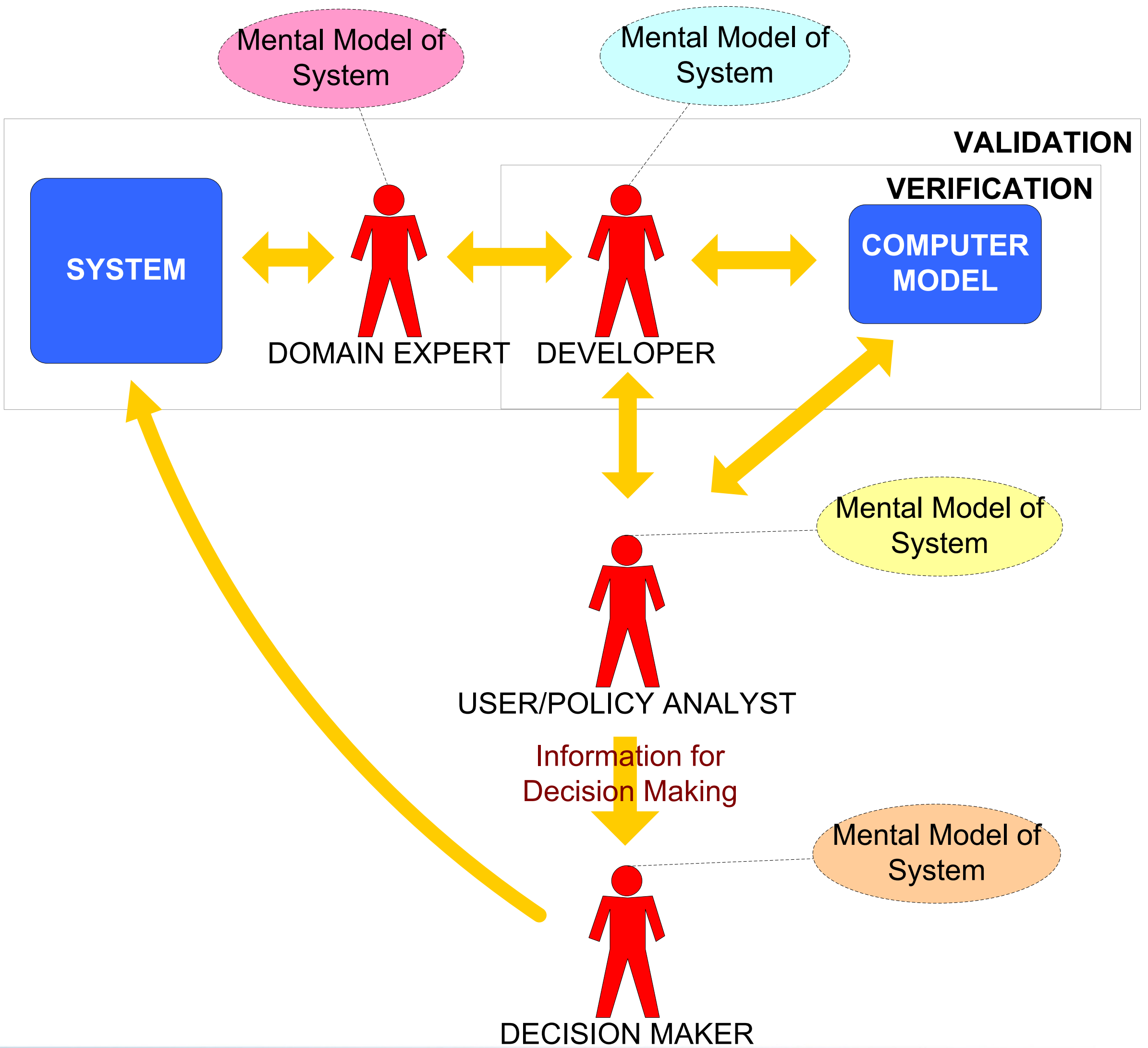
So We Have a Validation Process for a Model

- Validation ensures that the model meets its intended requirements and *purpose* in terms of the methods employed and the results obtained
- The goal of model *validation* is to make the model *useful* in the sense that the model addresses the right problem, provides accurate information about the system being modeled, and to helps to makes the model is actually *used*

Enter Subjectivity: Why We Model

- We model for insights, not numbers
  - As an exercise in "thought space" to gain insights into key variables and their causes and effects
  - To construct reasonable arguments as to why events can or cannot occur based on the model
- We model to make quantitative or *qualitative* predictions or forecasts about the future

An Information Focused View of Model V&V



Questions About Validation

How can a model be validated if…

- Controlled experiments cannot be performed on the system, or if only a single historical data set exists?
- The real-world system being modeled does not exist, yet?
- The model is inherently stochastic?

Further Questions About MAS Validation

- How can agent-based models be validated with their added complexities?:
  - Agent behaviors and interaction mechanisms
  - "Butterfly Effect" - possible extreme sensitivity to initial conditions
  - Adaptive agent or agent population behaviors
  - Self-organization and emergence of patterns and structures

"Does anybody here have <u>any</u> answers?"

ABM Validation Challenges

- Difficulties in measuring subjective choices, complex psychology, and irrational behavior
- Simulating detailed behaviors of agents can be computationally intensive
- Identifying and understanding emergent behavior

Swarming "Boids" agents with four simple rules and:

•No leader

•No purpose

•Purely reactive behaviors

Validating Theory

- Often, multi-agent models are interdisciplinary in nature involving various social, behavioral, and systems sciences
- "Validating Theory" refers to how the technical details of a model relate to the relevant disciplines, experts, and underlying theories or hypothesized about the dynamic mechanisms upon which the model is based
- Theory?
  - What theories are used in the model?
  - How are the theories used together or combined in the model?
  - Where are the theory gaps in the model?

Practical Validation

Establishing Credibility

- Unlike physical systems, for which there are well-established procedures for model validation, no such guidelines exist for social modeling
- In the case of models that contain elements of human decision making, validation becomes a matter of establishing *credibility* in the model
- Verification and validation work together by removing barriers and objections to model use
- The task is to establish an argument that the model produces sound insights and sound data based on a wide range of tests and criteria that "stand in" for comparing model results to data from the real system
- The process is akin to developing a legal case in which a preponderance of evidence is compiled about why the model is a valid one for its purported use

Pathways to Validation

- Cases
  - Exploration of critical cases
  - Exhaustive exploration of cases
- Using models as exploratory e-laboratories
  - Rapid prototyping
- Multiple models
- Maximally diverse model ensembles
- Using subject matter experts
  - Evaluation
  - Role playing, participatory simulation
- Computational simulations as a special cases of analytical modeling

ABM Validation

ABM requires two additional validation considerations, compared to validation issues found in traditional modeling:

- Agent behaviors
- Agent interaction mechanisms
- Emergent structures and properties produced by the ABM

Putting Things in Perspective

There are at least two kinds of MAS models, purposes, and intentions which are perhaps polar extremes

- Abstract, elegant agent models that provide insights about the real world
- Large-scale agent models that are serious attempts to faithfully represent the dynamic processes of a system in the computer and reproduce observed behaviors

ABM is Being Applied in Many Areas Including…

- Transportation
  - Traffic dynamics
  - Regional transportation
  - Land-use planning
- Economics
  - Emergent market structures
  - Financial markets
- Medicine
  - Epidemics
  - Healthcare
- Energy
  - Electric power markets
  - Technology diffusion
- Business
  - Manufacturing
  - Supply chains
  - Consumer markets
- Defense
  - Autonomous vehicle networks
  - Force-on-force engagement
  - Terrorist networks
- Pedestrian movement
  - Evacuation modeling
  - Architectural planning
- Biology
  - Ecological networks
  - Cellular behavior
- Social processes
  - Crime reduction
  - Policy assessment
- Archaeology
  - Human settlement
  - Societal collapse

A Note on the on Modeling Literature Pertaining to Validation

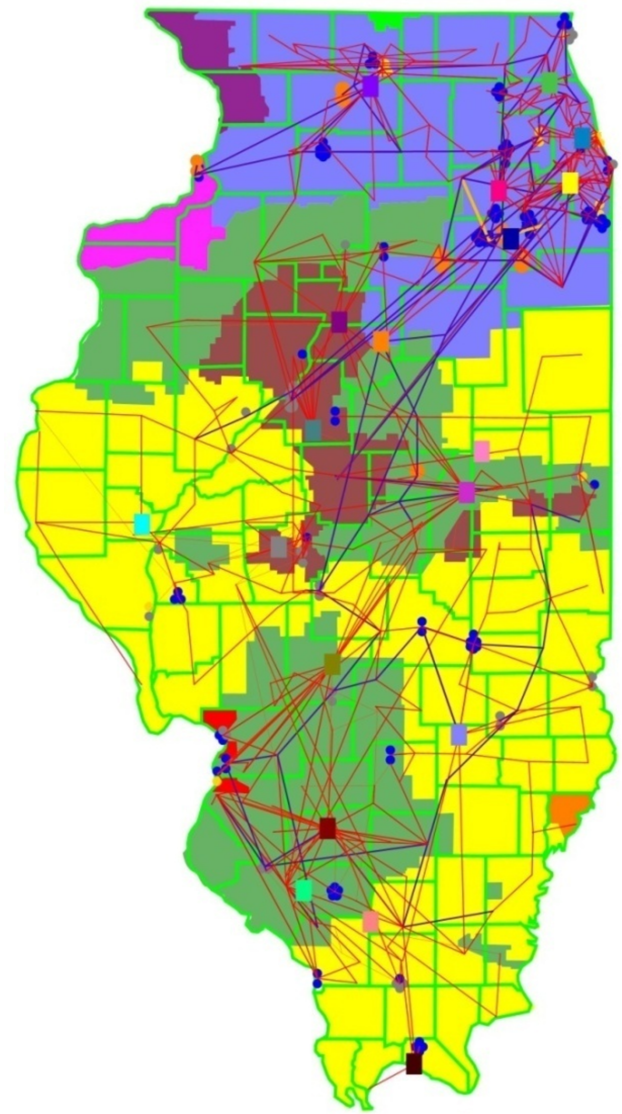
- Few applied papers do a convincing job of describing how the authors /developers went about validating their models
- Perhaps a full paper on how a model was validated would be appropriate and helpful in addition to a paper that describes the model itself!
- We could ask the question "What does a good application validation paper look like – and what would be included?"

Characteristics of Real-World Large-Scale ABM Applications

- Faithful representations of the real world
- Real data
  - Populations of "synthetic" agents and characteristics
  - Real agent behaviors and salient interactions
- Geographic landscapes and geo-spatial attributes
- Having an Impact
  - Providing key information for making decisions and setting policies
  - Changing business processes in organizations
  - Adding to scientific knowledge through *in-silico* experiments
- "Validated"

Electricity Power Deregulation

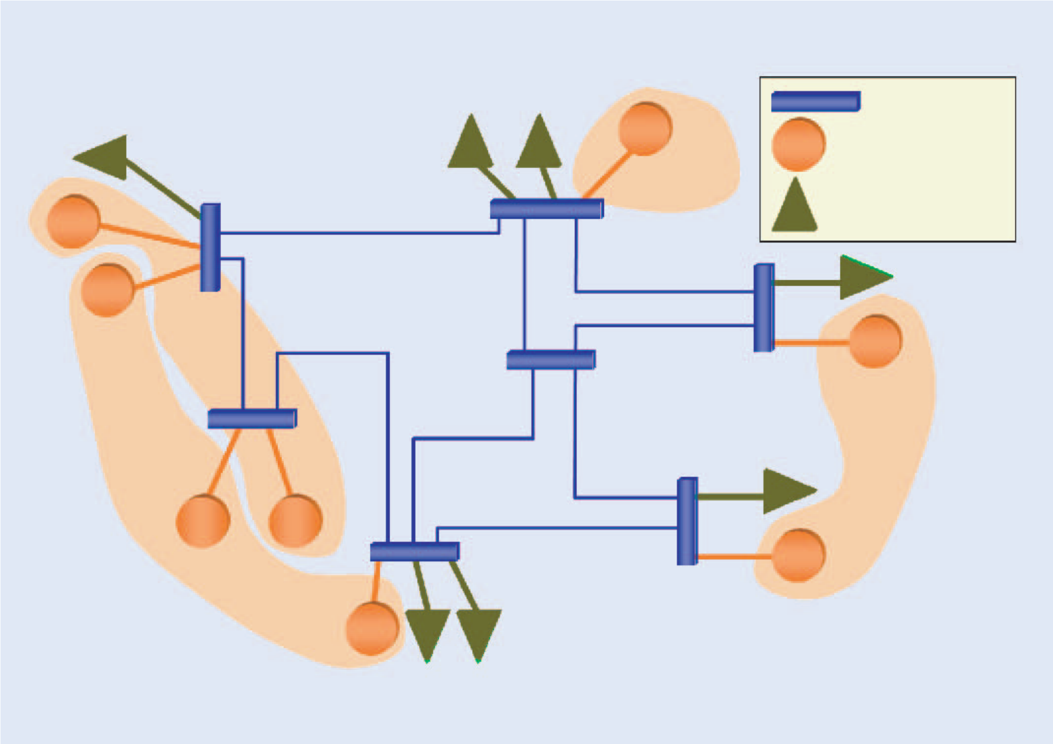
- Centralization - *Before*
  - Single electricity price for whole state
  - Rate of return regulated by Illinois Commerce Commission (ICC)
- Decentralization – *As of January 1, 2007*
  - Companies free to price their production by bidding into power pools
  - Independent System Operator (ISO) matches supply and demand and clears the market
  - People make their own decisions on consumption
- New ways to calculate electricity prices
  - Locational marginal pricing (LMP)
  - 30 separate pricing zones on the grid
- These issues can only be addressed through agent-based modeling



Illinois electric power transmission grid and service areas

Why Agents?

- The rules of business and social interaction are at least as important as the rules of physics when it comes to the generation, sale, and delivery of electrical power
- Agents can be used to model decentralized competitive decision-making
- Learning and adaptation of agent behavior can be modeled
- Alternative market rules can be tested



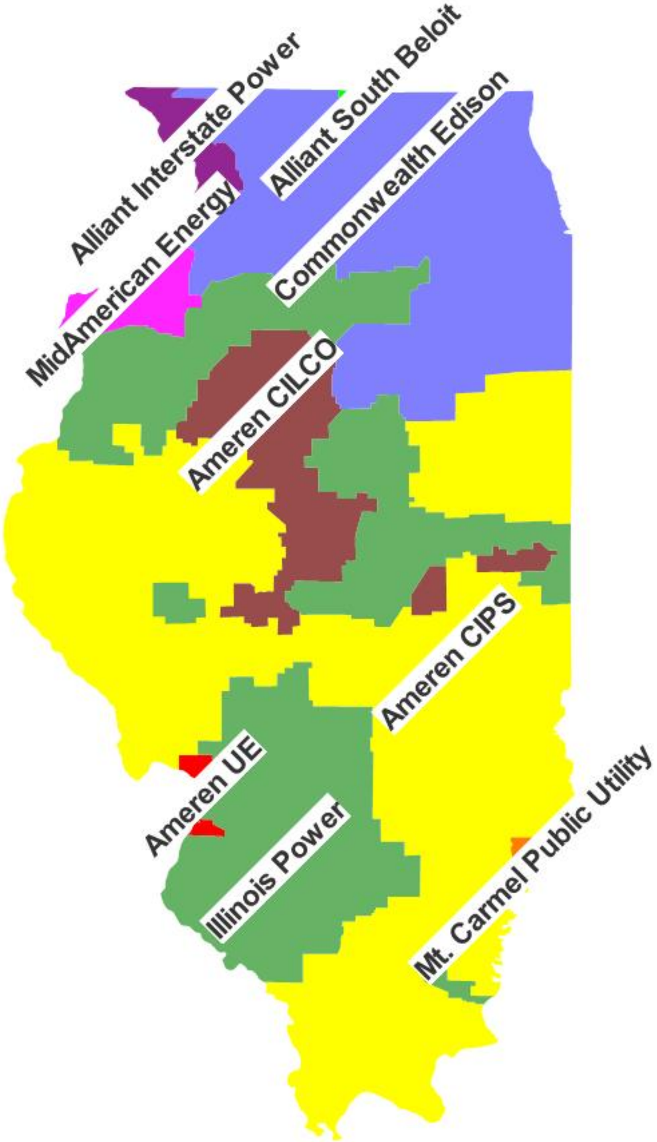
Validation of EMCAS

- EMCAS deals with an important public policy issue in which virtually all segments of society have important economic interests
- The modeling results must be able to withstand scrutiny and criticism from many parties having vested interests in what the model results may imply
- Burden for model validation is severe by most modeling standards

EMCAS Model

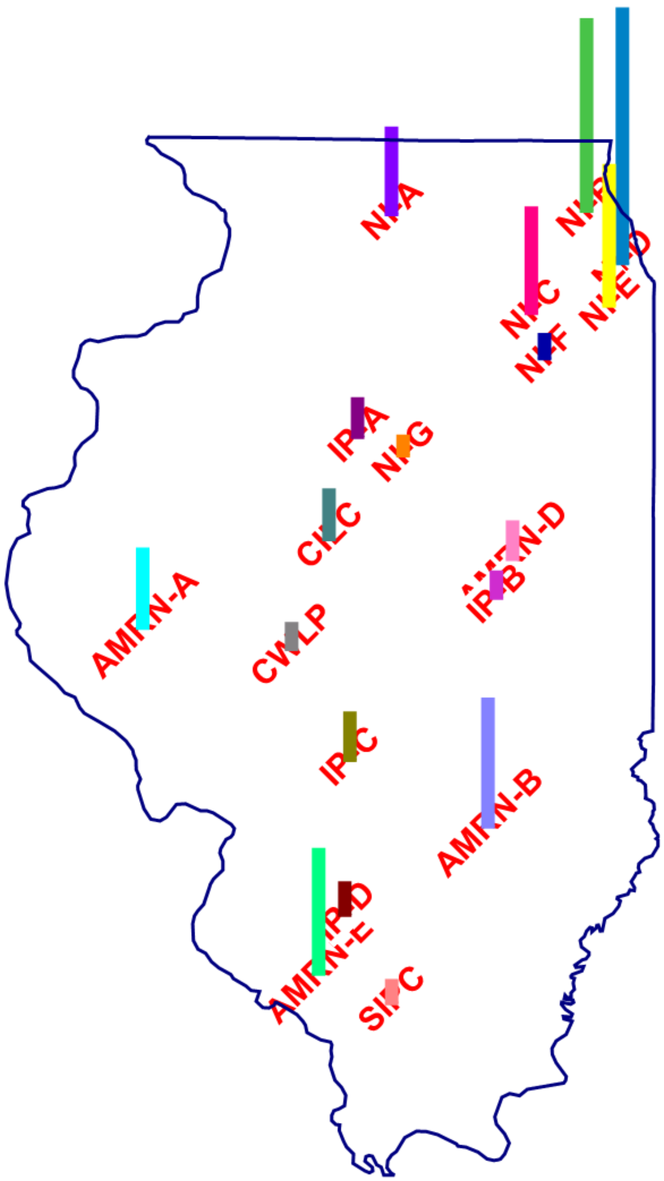
- Agents represent participants in the electricity market
  - Generation companies (GenCos)
  - Demand companies (DemCos)
  - Transmission companies (TransCos)
  - Distribution companies (DistCos)
  - Independent system operators (ISOs)
  - Consumers
  - Regulators
- Agents are specialized, performing diverse tasks using individual rules of behavior
  - Maximize profits or more generally utility
  - Learn about the market based on submitted bids and market outcomes
  - Explore new strategies in attempt to adapt to dynamic supply and demand forces in the marketplace

Service Areas



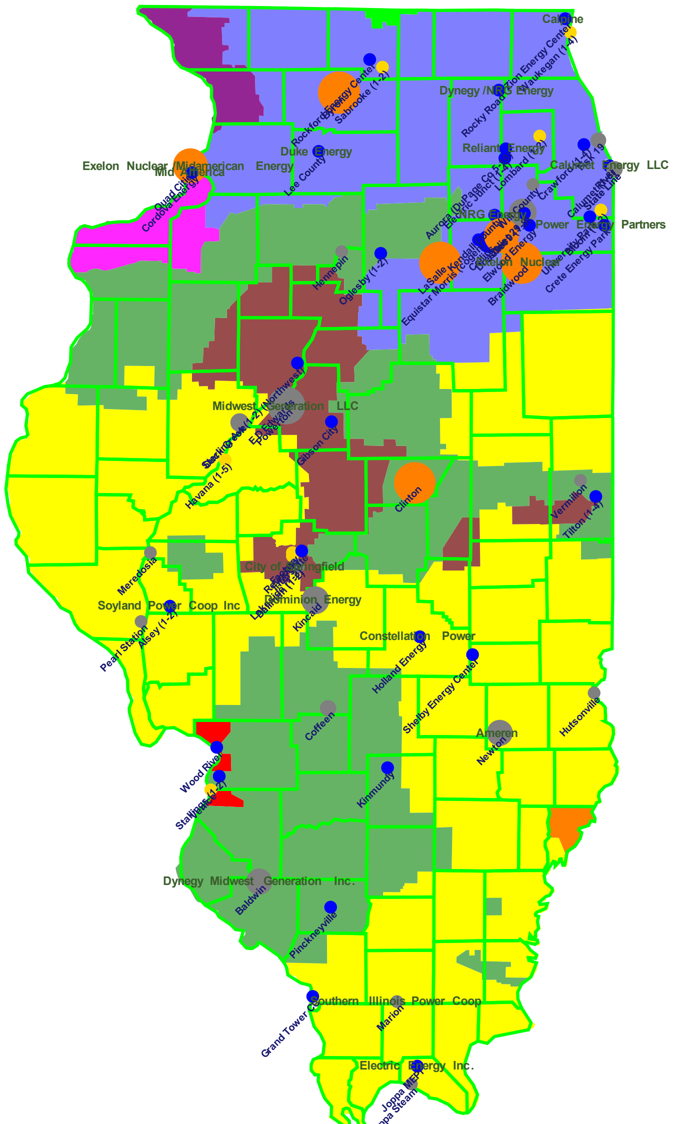
- 9 service areas

Load Zones

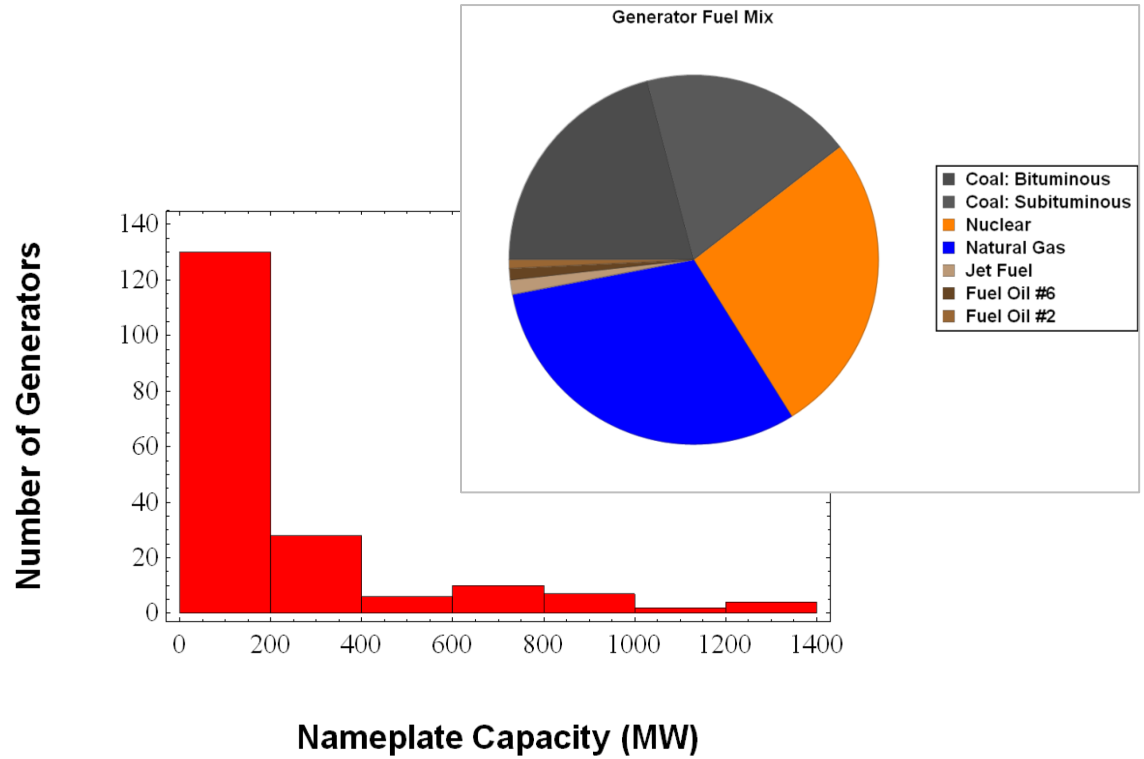


- 852 buses with loads
- 8760 hours (1 year)
- Collected into 18 load zones

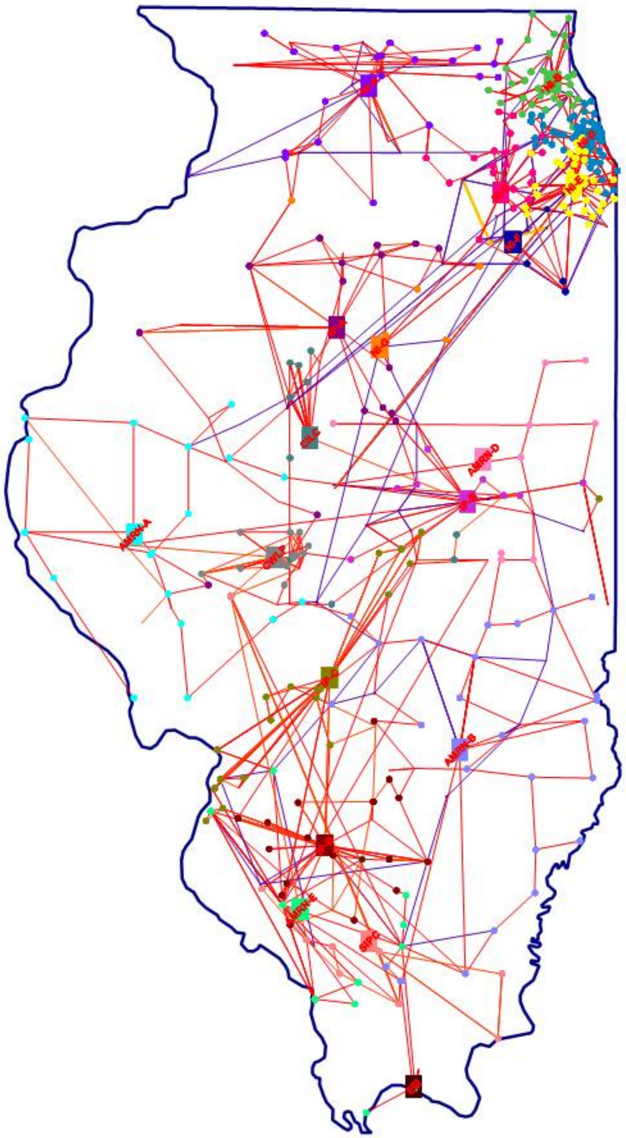
Generating Plants



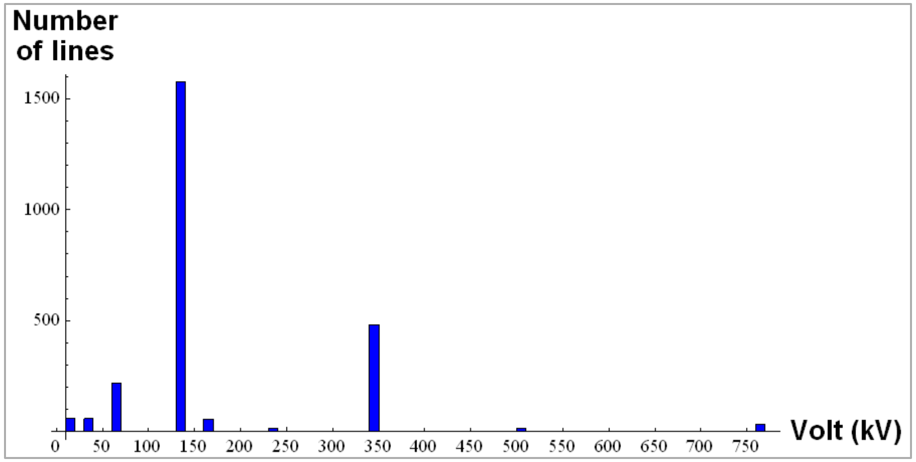
- 66 plants
- 237 generating units
- 638 total generator blocks
- 20 GenCos



Transmission Network



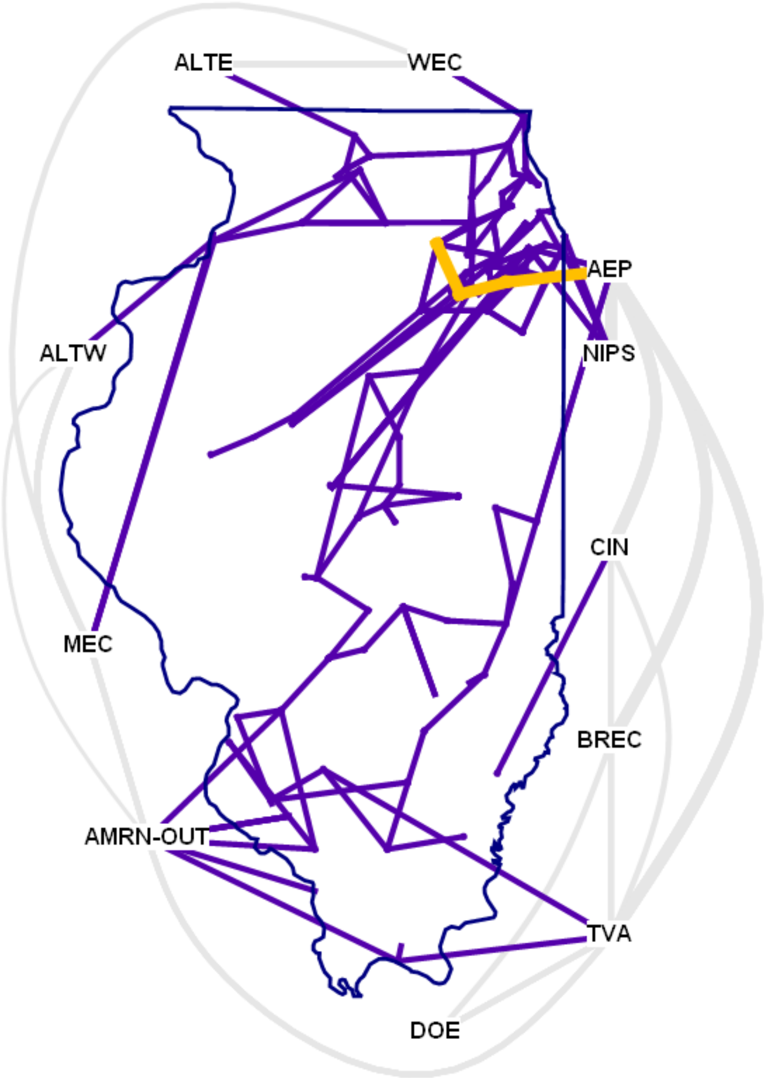
- 2522 transmission line segments
- 1908 buses
- 852 buses with loads for 8760 hours (1 year)



There are 2522 unique rows and 20 data fields.

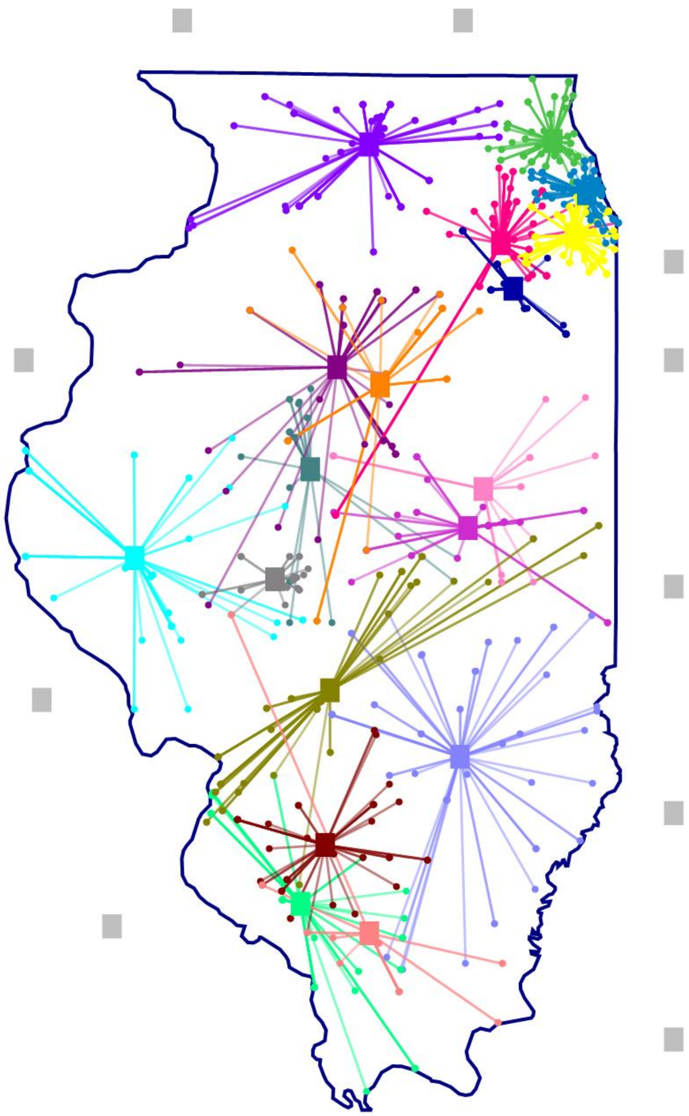
interface	fromNode	toNode	voltagekV	numOfCircuits	circuitReactance	rateACapijMW	rateACapjiMW	rateACc
18001_30825	18 001	30 825	345	1	0.003	949	949	0.0001
18001_32339	18 001	32 339	345	1	0.026	1195	1195	0.0001
18008_30282	18 008	30 282	500	1	0.257586	10 000	10 000	0.0001
18008_27618	18 008	27 618	500	1	0.185553	10 000	10 000	0.0001
18008_25387	18 008	25 387	500	1	0.451021	10 000	10 000	0.0001
18008_22668	18 008	22 668	500	1	0.873388	10 000	10 000	0.0001
18008_22660	18 008	22 660	765	1	0.146822	10 000	10 000	0.0001
18008_18401	18 008	18 401	500	1	0.0585579	10 000	10 000	0.0001
18008_31230	18 008	31 230	500	1	0.254055	10 000	10 000	0.0001
18008_18037	18 008	18 037	500	1	0.775199	10 000	10 000	0.0001
18008_27561	18 008	27 561	500	1	0.411563	10 000	10 000	0.0001
18008_30886	18 008	30 886	500	1	0.255704	10 000	10 000	0.0001

Transmission Network Equivalenced to Eastern Interconnect



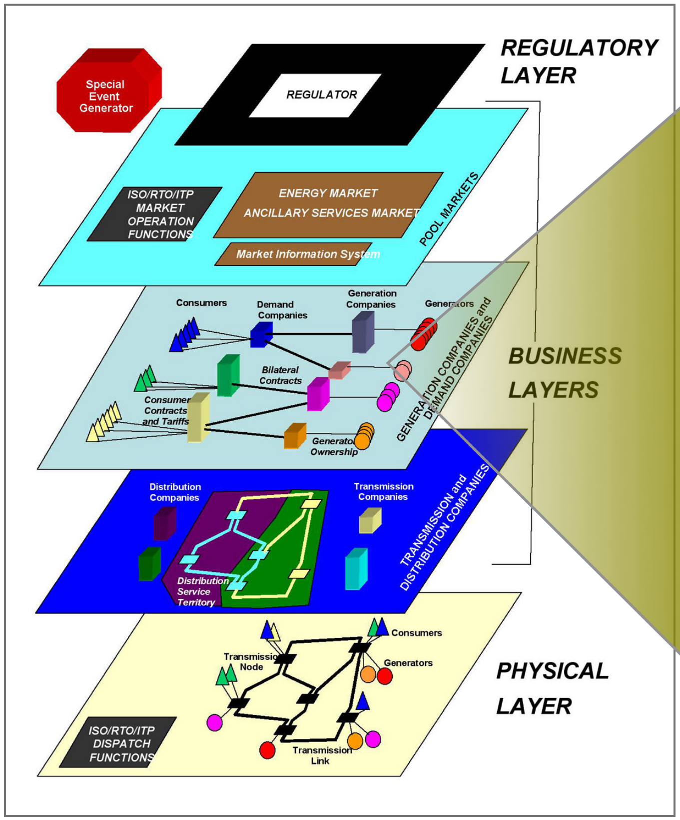
- 11 out-of-state connections to equivalenced network

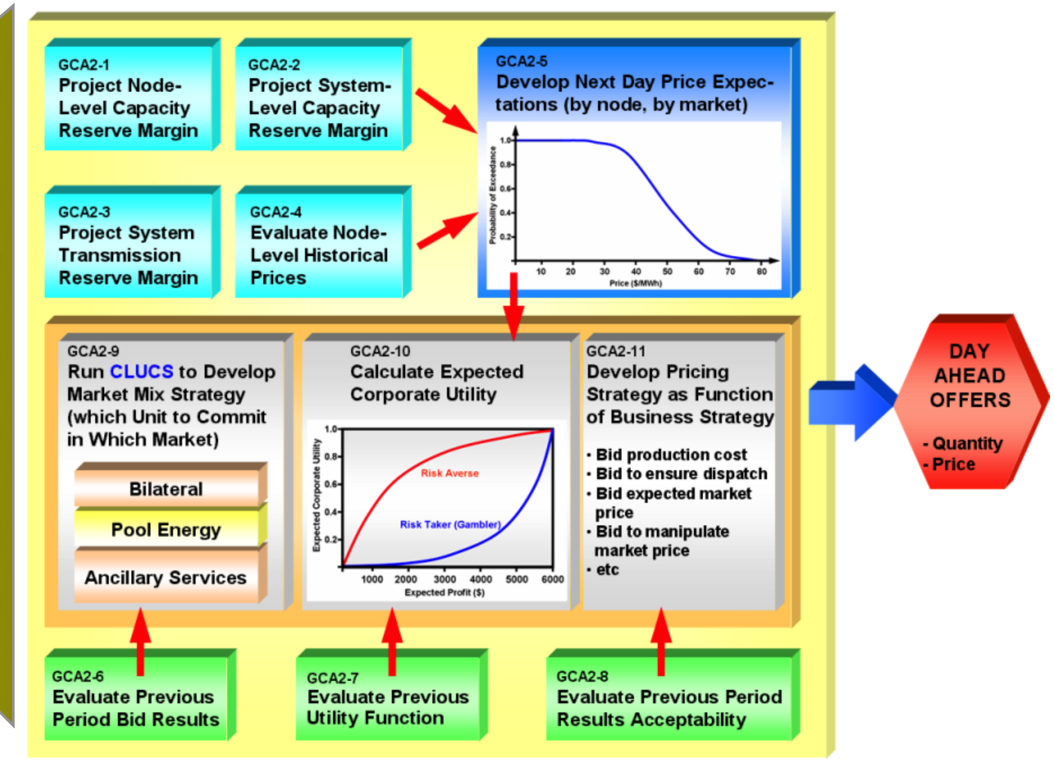
Zones – LMP Pricing



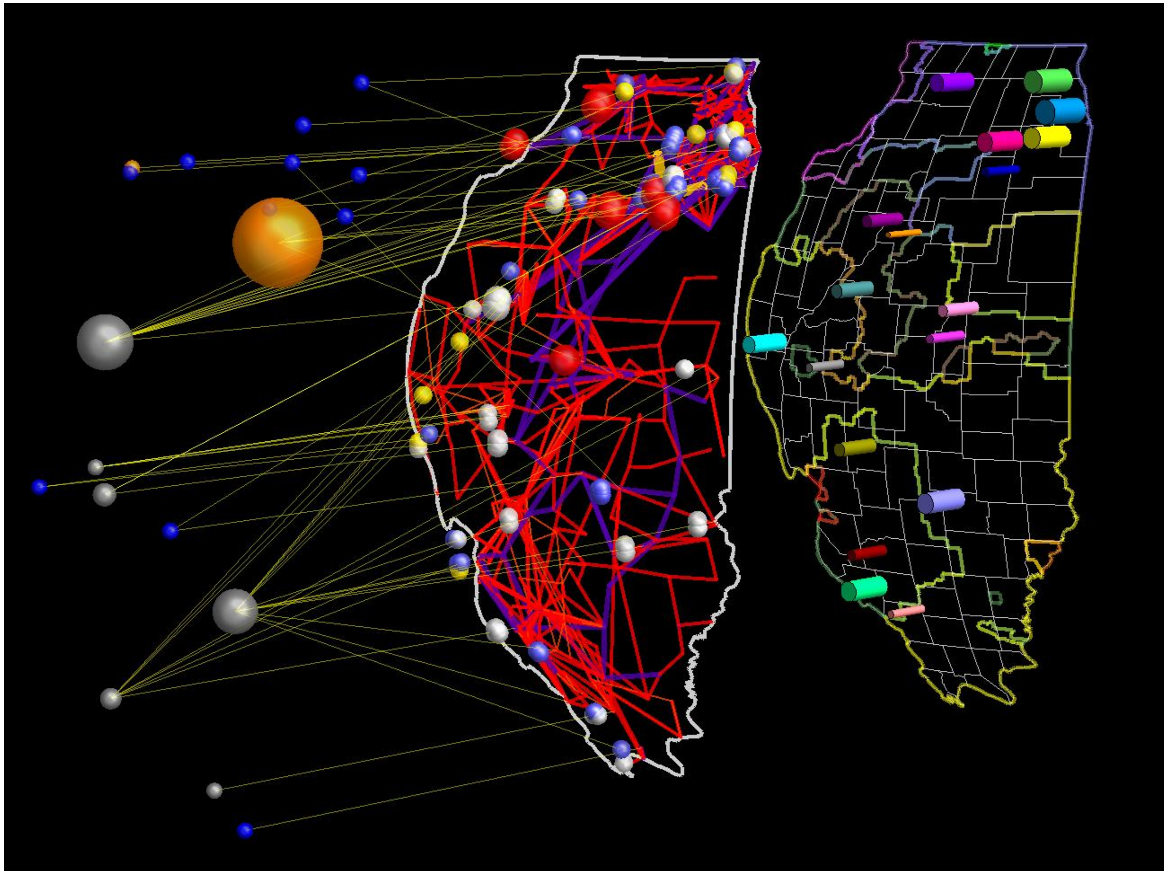
- 19 zones with Locational Marginal Prices (LMP) averaged across associated buses

Generation Company Agents Consider Many Factors in Proposing Bids for the Day-Ahead Market





Illinois Electric Power: EMCAS Agent-based Application



- Generating companies and ownership relationships (left)
- Electric generators and transmission network (center)
- Service area loads, consumers (right)

Validation Case Study: EMCAS

- Data
  - Checking the currency of the data with the original data sources
  - Cross-checking data with third parties having a vested interest in the data
- Subject Matter Experts (SME)
  - Model was developed by a team of experienced domain experts
  - Independent electric utility SMEs provided critical industry experience
- Participatory Simulation
  - Ability to place themselves in the positions of agents in the deregulated markets
- Replication of Special Cases
  - Model runs constructed to replicate special case for the previously validated *regulated* power market
- Comprehensive Model Cases for the Agent Parameter and Strategy Space
  - Not possible to draw general conclusions from only a handful of model runs: non-linear, dynamic aspects of the agent behaviors and interactions
  - Extensive cases verified expected model behaviors and discovered model unexpected model behaviors
  - Unexpected cases created focal points for further model runs and in-depth analysis
  - Comprehensive testing of plausible agent strategies
  - Extensive use of data visualization techniques
- Model-to-Model Validation
  - Validation of the simplified DC model to the complete and validated AC model was done by comparing results for extensive number of cases

Data Validation

- A single data gap or inconsistency can invalidate the results of any model and destroy the model's credibility (if not formally invalidate the model itself)
- Data Issues
  - Heterogeneous data bases
  - Data currency
  - Data gaps and inconsistencies
- Cross-checking data with third parties having a vested interest in the accuracy of the data
- Likely sources of objections to the data used in the model
- Proprietary data provided by vested third parties

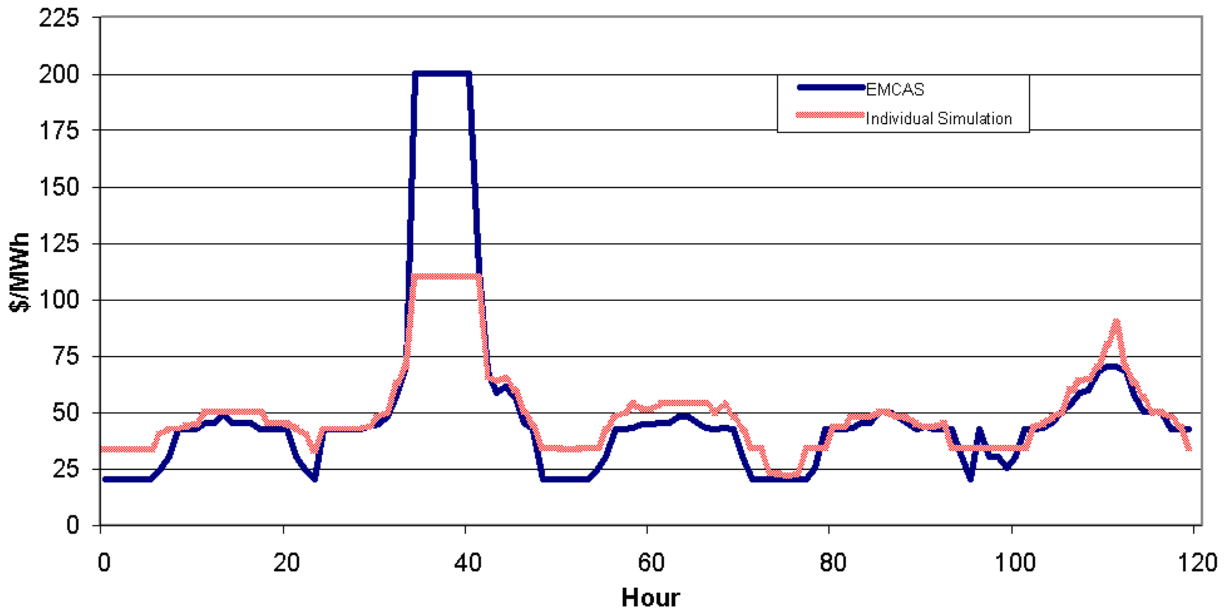
Subject Matter Experts

- EMCAS developed by a team of experienced subject matter experts (SME) consisting of electric power engineers, systems engineers, and economists.
- Independent subject matter experts evaluated the scope and detail of the model during the critical design phase, reviewing model assumptions, preliminary results, and purported agent behaviors.
  - SME workshop included former utility operators, industry consultants, and electric power market traders.
  - Independent SMEs had little vested interest in the correctness of the model’s assumptions if they were at odds with personal and industrial experience.
  - Independent SMEs had the ability to place themselves in the positions of agents in the deregulated markets, based on their work experience.
- The SMEs provided constructive critiques of relevant agent behaviors and validated (or invalidated) previous assumptions.
- In later phase of model application, SMEs raised questions on the EMCAS system-wide results and provided insights for explaining model behaviors

Participatory Simulation

- Proved to be very insightful to run a "participatory simulation" with real people playing the roles of the agents in the deregulated electric power market.
- Served to validate key assumptions on agent behaviors in a deregulated environment
- Identified likely strategies the agents would use in various situations
- Useful to include non-SMEs for the sake of increasing the diversity of the strategies that participants create during the simulation

Comparison of results from a 6-player participatory simulation and EMCAS simulation



Critical Tests and Key Indicators

- One mode of validation is to see how well the model performs at reproducing system behaviors for a specific set of critical test cases and examine key indicators upon which to judge model validity.
- Three critical test cases were identified:
  1. Replication of the known results for the current regulated electric power market configuration,
  2. Emergence of so-called hockey stick strategies adopted on the part of generating company agents as the agents in the deregulated simulation learned to responded to market incentives, and
  3. Replication of elements of the system behaviors reminiscent of the California electric power deregulation situation.

Model-to-Model Validation

- The EMCAS model had two other models against which its results could be compared in the interest of validation
  1. A model of a centrally planned, regulated electric power market configuration, as discussed above
  2. A more detailed and complete alternating current (AC) model of the physical aspects of the electric power grid.
- The EMCAS model subsumed the regulated electric power market model.
- Validation of the EMCAS reduced form direct-current (DC) model of the electric power grid was also necessary

Comprehensive Testing

- Testing of plausible agent strategies revealed it was not possible to draw general conclusions from only a handful of model runs
- The entire (price-quantity) space of plausible agent strategies and parameter settings had to be mapped
- Agent strategies particularly important to understand
  - Agent strategies that are observed in practice
  - Strategies that could be logically expected to be followed, for example, based on the rational choice model of the agent
- Strategies:
  - Bid all capacity at production cost (long-run theoretical market clearing strategy)
  - Physical withholding (reducing bid quantity at given price levels)
  - Economic withholding (increasing price levels for given quantities)
  - Price probing (exploring the effects of raising prices incrementally above recent bid price levels)
  - Price discovery (exploring the effects of raising prices incrementally above hypothesized levels for marginal suppliers)
- Within each of these strategies, a family of plausible agent sub-strategies

Invalidation Exercises

- Since no model can ever be validated with complete certainty, validation exercises essentially consist of a series of attempts to "invalidate" a model
- But this invalidation process can be done systematically rather than on an ad hoc basis to avoid validation bias
- Extensive model runs served the multiple purposes of verifying model behaviors that were expected, thereby increasing the confidence in the model, discovering model behaviors that were outside the range of what was expected, and invalidating the model's assumptions
- Discovering cases for which model behaviors were unexpected created focal points for more in-depth analysis and explanation

How Would We Go About Using Such A Model to Answer the Question of Whether There Is Market Power?

- Establish a base case that excludes strategic behavior – bids based on production cost
- Establish reasonable assumptions on scenario variables
  - Weather and effect on peak load
  - Unit outage assumptions
  - New capacity additions and transmission network adjustments
  - Agree on definitions (short-run, long-run), e.g., production cost
  - Identify credible strategies that agents might employ
- Test the effect of these strategies incrementally with the model and compare with the base case

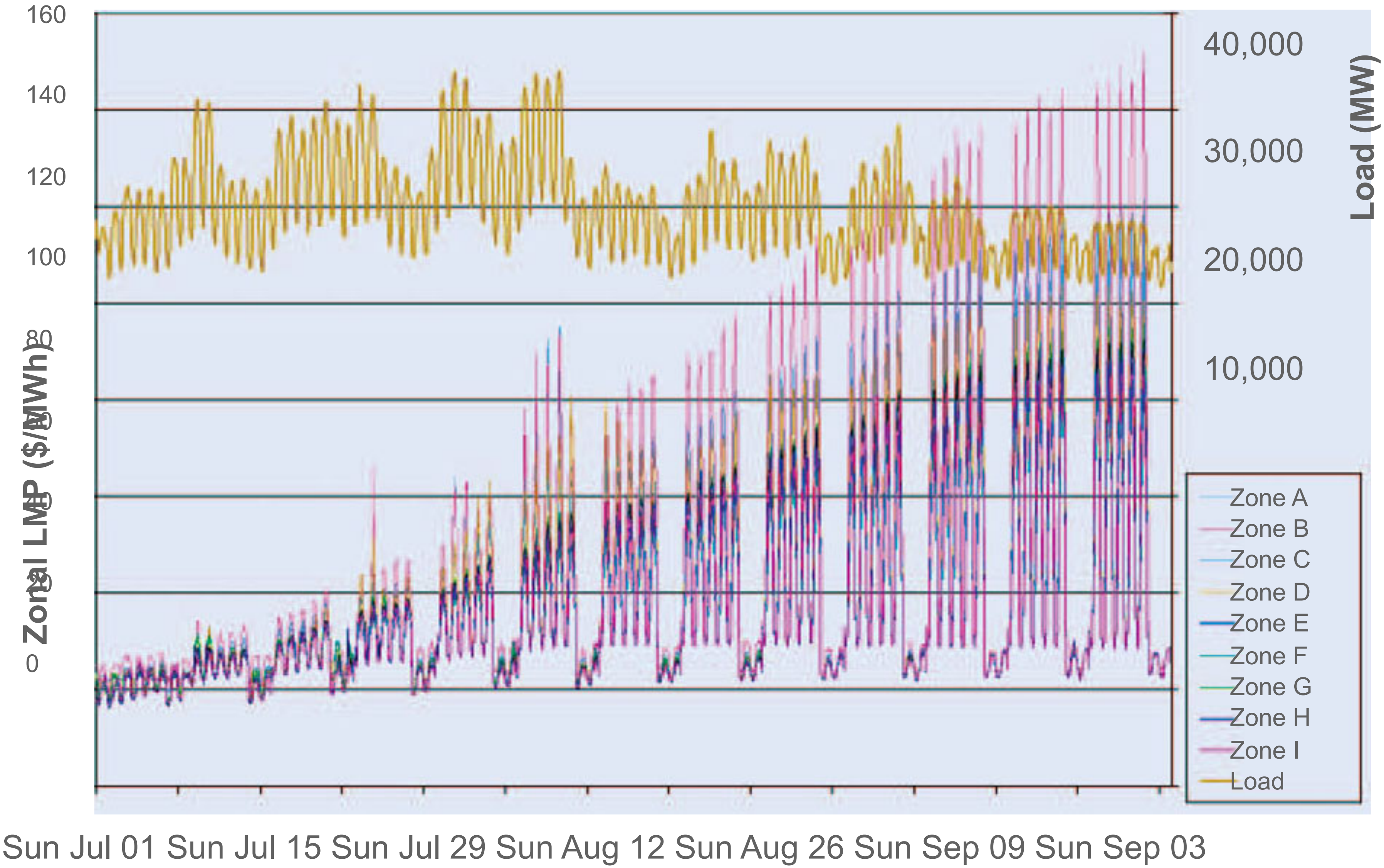
Production-Cost Analysis Case: Assumptions

1. Locational marginal prices (LMP) are paid to GenCos;
2. DemCos pay the load-weighted average zonal price;
3. There is a day-ahead market (DAM) administered by the ISO/RTO for both energy and ancillary services where GenCos and DemCos can participate;
4. Bids are unregulated;
5. Bilateral contracts between suppliers (GenCos) and purchasers (DemCos) are not allowed; all power is purchased in the spot markets;
6. Generation companies sell electricity based entirely on the production costs of the generators, including fixed costs;
7. Demand companies serve firm uninterruptible load; unserved energy is due only to forced outage conditions and not to any market considerations; consumers are assumed to have no response to electricity prices; and
8. Transmission companies do not employ any strategic business behavior; their revenue comes in two forms: a transmission use charge of $3 per MWh and a transmission congestion payment.

Strategies for Generating Company Agents

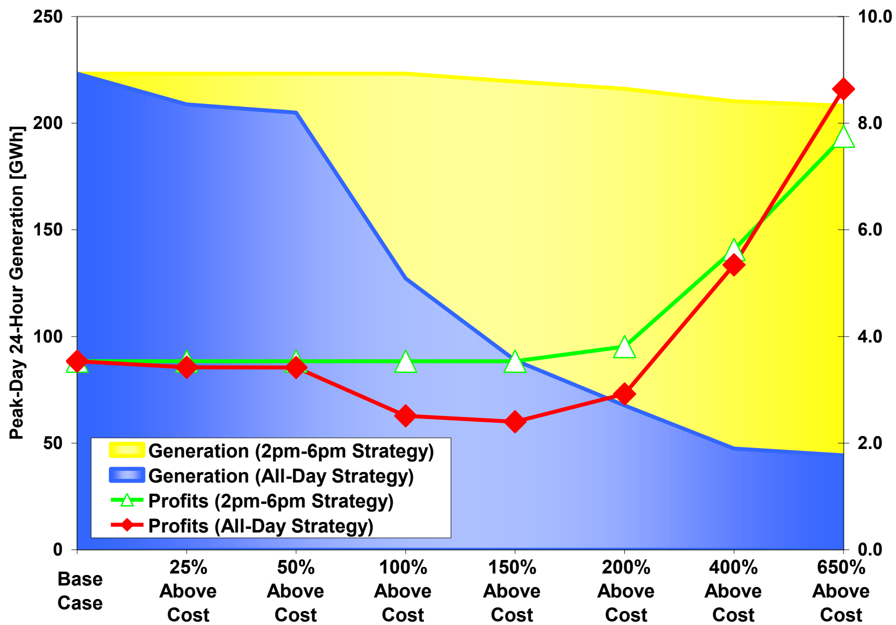
- GenCo agents have two levers upon which to base their strategies
  - Capacity to offer (bid) into the market
  - Price to offer the capacity
- Withholding
  - Physical Withholding
  - Economic Withholding
- GenCo Strategy: Price Probing: probe the market for weaknesses
  - Discover if you are the marginal supplier
  - Discover who is the marginal supplier
- Other Possible GenCo Strategies
  - incremental pricing
  - Bid production cost
  - Bid low to ensure dispatch
  - Bid high to increase the market clearing price
  - Adjust prices based on past performance (bid price based on moving average price)

Results of Scenario 2 – Price Probing

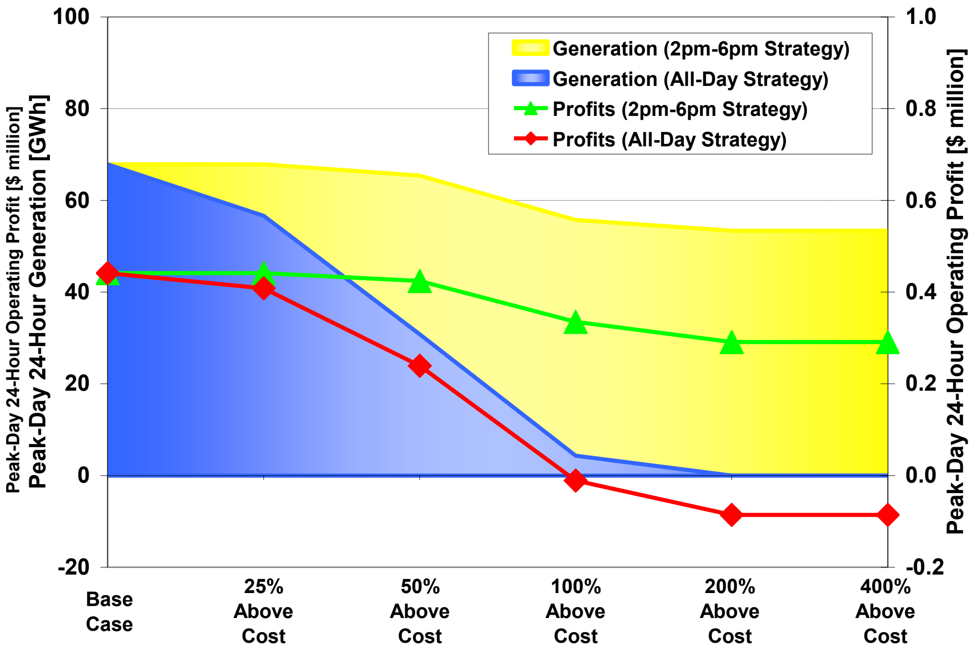


Identifying Market Power

- GenCo bid prices are increased to various levels above production cost
  - All-day
  - Afternoon only



GenCo with Market Power



GenCo with No Market Power

EMCAS Impact

- The findings: *There is the potential for some companies to exercise market power (i.e., raise prices and increase profitability by unilateral action) and raise consumer costs under selected conditions, particularly when there is transmission congestion.*
- EMCAS results* have been entered into the public record of the Illinois Commerce Commission (ICC), 6 June 2006.
- Report available from the ICC web site http://www.icc.illinois.gov/
- EMCAS is an example of an agent-based model that has been successfully applied to a real-world policy issue and provided information that would otherwise have not been available using any other modeling approach.

*Cirillo, R., P. Thimmapuram, T. Veselka, V. Koritarov, G. Conzelmann, C. Macal, G. Boyd, M. North, T. Overbye and X. Cheng. 2006. *Evaluating the Potential Impact of Transmission Constraints on the Operation of a Competitive Electricity Market in Illinois*, Argonne National Laboratory, Argonne, IL, ANL-06/16 (report prepared for the Illinois Commerce Commission), April.

Lessons Learned from EMCAS Validation

- As each step in the validation process was completed, the model was better accepted as a valid tool for answering important questions on aspects of electric power deregulation
- Lessons Learned
  - Easy to convince decision makers that an agent-based model could be relevant and useful. Difficult to credibly characterize agent behavior, explain it, and validate or explain the results
  - Agent models contain more assumptions and data than traditional models
  - There is a lot of explaining to do to decision makers about agent behavior. This adds to the burden of validation
  - The model validation phase ended up taking as long as the model development phase. In the end, however, it was generally accepted that the model was a valid one for answering a wide range of important questions pertaining to electric power deregulation

Lessons Learned Cont'd.

- When explaining the agent strategies to decision makers, certain questions seem to repeatedly arise:
  - Are all the relevant agent strategies considered?
  - Why is the model attributing certain strategies to some agents and not to others?
  - Why do all the agents adopt the same strategy (clearly not realistic)?
- In the final analysis, all model results have to be explainable in plain English and supportable or they were not credible or useful to decision makers

Practical Issues Related to Validation

- Verification
  - A separate set of test runs is specified to ensure the model's program is operating as intended
  - Unit testing is very important
  - Logging a subset of model computations as results
- Calibration

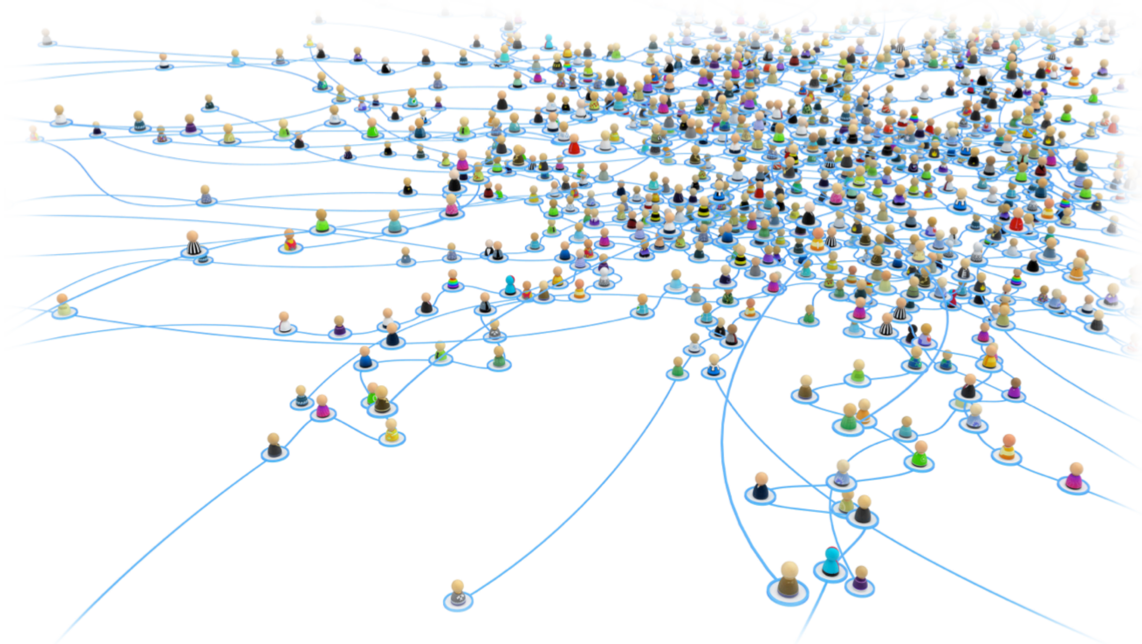
  After a long week of computer runs: …are we validating the model, yet? – or just calibrating the model?
- Portability

  Is your model and data transferable to third parties? – Should it be?
- Replication

  If *your* model is to be used to do good science, should the results from *your* model be replicable? -- Answer yes or no (hint: this is a debatable issue)

Conclusions

- An enormous number of problems seem to be amenable to agent modeling
- Perhaps, in the future all models and simulations will be agent-based or have agent-based components
- There is a big beautiful world of potential agent applications out there – *go forth and model!*

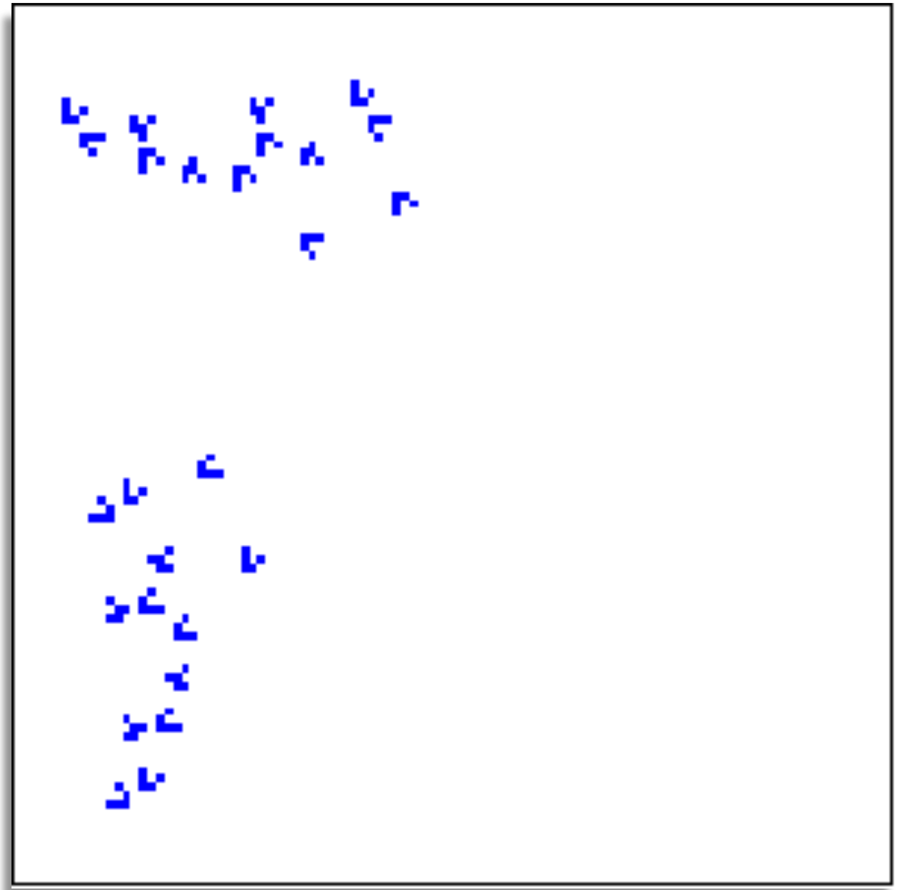


Questions?

Charles Macal

Argonne, U. Chicago

macal@anl.gov



3 Simple Rules.

Two sets of 13 **Life** *Gliders* create *Glider Guns* each of which creates a *Glider Stream*, which eventually destroy the *Glider Guns.*