

# The TEMOA Project:

Tools for Energy Model Optimization and Analysis

International Energy Workshop 2010

Stockholm, Sweden, June 21

Joe DeCarolis, Kevin Hunter, Sarat Sreepathi

Department of Civil, Construction, and Environmental Engineering

North Carolina State University

[jdecarolis@ncsu.edu](mailto:jdecarolis@ncsu.edu)

# Energy-economy optimization models



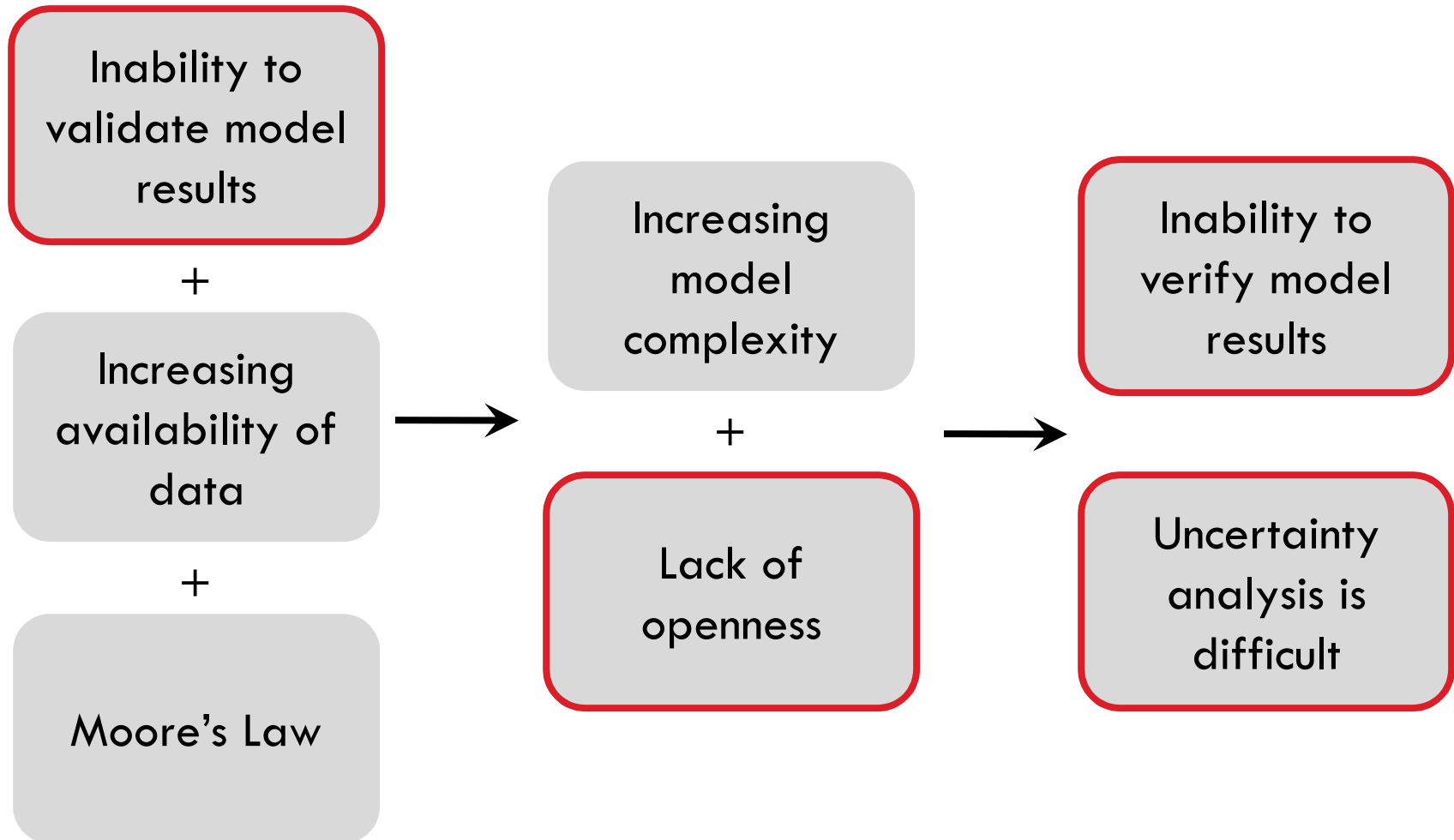
Energy-economy optimization models refer to partial or general equilibrium models that **minimize cost or maximize utility** by, at least in part, optimizing the energy system over multiple decades

Expansive system boundaries and multi-decadal timescales

Encoded with a set of structured, self-consistent assumptions and decision rules

Such models have emerged as a key tool for the analysis of energy and environmental policy

# Broad trends, bad outcomes



# Model validation is a problem



No practical way to validate energy-economy models → cannot be validated in the same way as models of physical processes

Three validation options:

1. Wait
2. Backcast
3. Compare results with other models

Little to guide the modeler and reign in efforts that do not improve model performance

# Lack of openness



Most EEO models and datasets remain closed source. Why?

- protection of intellectual property
- fear of misuse by uninformed end users
- inability to control or limit model analyses
- implicit commitment to provide support to users
- overhead associated with maintenance
- unease about subjecting code and data to public scrutiny

Model	Developers	Model Type <sup>a</sup>	Scope	Open Source	Comm. software required <sup>b</sup>	Collab tools <sup>c</sup>
ADAGE	Research Triangle Institute <a href="http://www.rti.org/">http://www.rti.org/</a>	CGE	global	No	Yes	No
AMIGA	Argonne National Laboratory <a href="http://amiga.dis.anl.gov/">http://amiga.dis.anl.gov/</a>	CGE	global	No	No	No
DICE-2007	Yale University <a href="http://nordhaus.econ.yale.edu/DICE2007.htm">http://nordhaus.econ.yale.edu/DICE2007.htm</a>	CGE	global	Yes	Yes	No
EPPA	MIT <a href="http://globalchange.mit.edu/igsm/eppa.html">http://globalchange.mit.edu/igsm/eppa.html</a>	CGE	global	No	Yes	No
GTAP	Purdue University <a href="https://www.gtap.agecon.purdue.edu/">https://www.gtap.agecon.purdue.edu/</a>	CGE	global	Yes	Yes	No

## Open Source Energy Modeling System (OS<sub>e</sub>MOSYS)

Utilizes the GNU Linear Programming Kit (GLPK); coded in Mathprog

See presentation by Mark Howells, 3:30pm this afternoon, Room E36

OS<sub>e</sub>MOSYS Workshop, tomorrow (6/22), 5:15-6:10pm, Room E1

GCAM	Joint Global Change Research Institute <a href="http://www.globalchange.umd.edu/models/gcam/">http://www.globalchange.umd.edu/models/gcam/</a>	CGE	global	No	No	No
NEMS	U.S. Energy Information Administration <a href="http://www.eia.doe.gov/oiaf/aeo/overview/">http://www.eia.doe.gov/oiaf/aeo/overview/</a>	CGE/ TE	U.S.	Yes	Yes	No
OS <sub>e</sub> MOSYS	International Atomic Energy Agency <a href="http://osmosys.yolasite.com/">http://osmosys.yolasite.com/</a>	TE/PE	any	Yes	No	No
SGM	Joint Global Change Research Institute <a href="http://www.globalchange.umd.edu/models/sgm/">http://www.globalchange.umd.edu/models/sgm/</a>	CGE	global	No	Yes	No

# Inability to verify model results



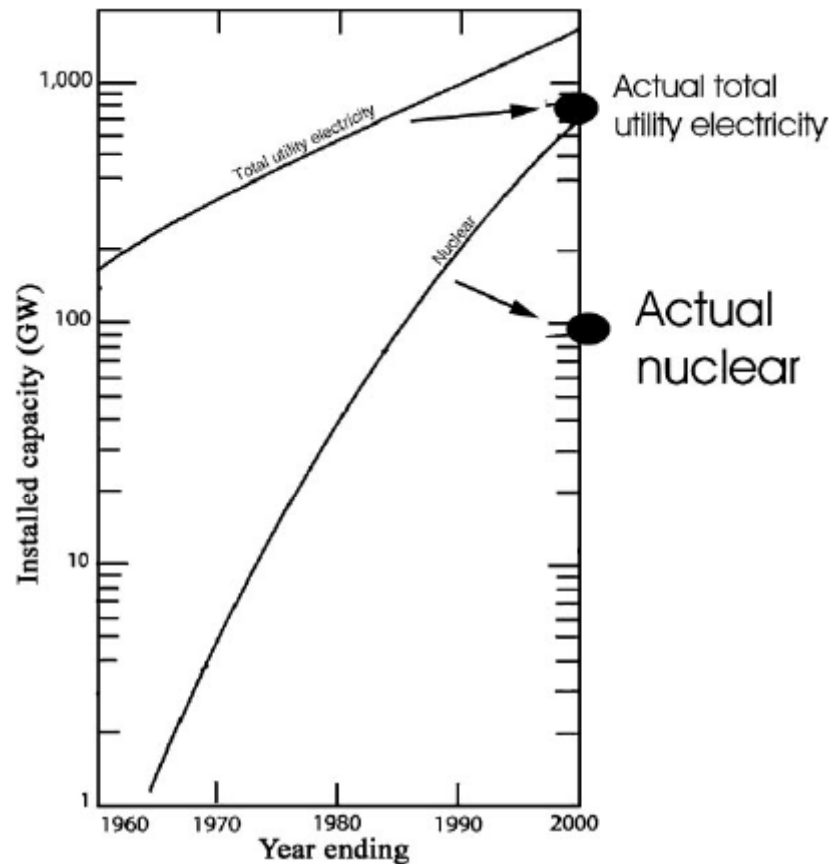
With a couple exceptions, **energy-economy models are not open source**

Descriptive detail provided in model documentation and peer-reviewed journals is insufficient to reproduce a specific set of published results

Reproducibility of results is fundamental to science

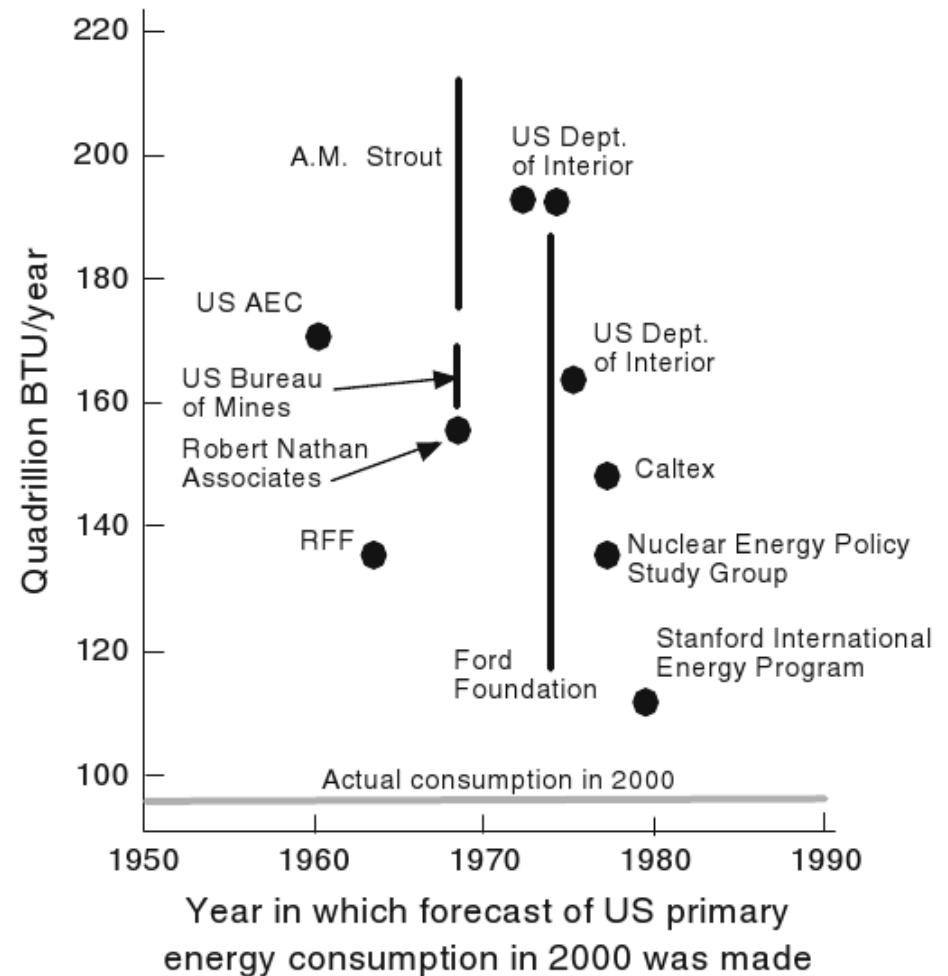
**Replication and verification of large scientific models can't be achieved without source code and input data**

# Our ability to predict is generally dismal



U.S. Atomic Energy Commission  
forecast from 1962

**Source:** Craig et al. (2002). "What Can History Teach Us? A Retrospective Examination of Long-Term Energy Forecasts for the United States." *Ann. Rev. Energy Environ.* 27:83-118.



**Source:** Morgan G, Keith D. (2008). "Improving the way we think about projecting future energy use and emissions of carbon dioxide." *Climatic Change.* 90: 189-215.

# Critique of scenario analysis



Stretch one's thinking about how the future may unfold

Shell Group (2005):

They are not forecasts, projections or predictions of what is to come. Nor are they preferred views of the future. Rather, they are plausible alternative futures: they provide reasonable and consistent answers to the 'what if?' questions relevant to business.

Without subjective probabilities  $p(X|e)$ , scenarios of little value

Cognitive heuristics play a role and can lead to misinterpretation of results

→ Availability heuristic

→ Anchoring and adjustment

A few highly detailed scenarios can create cognitively compelling storylines (Morgan and Keith, 2008)



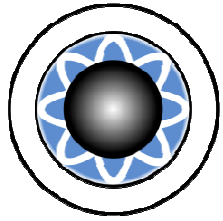
# The TEMOA Project

Tools for Energy Model Optimization and Analysis

Temoa also means “to seek something” in the Nahuatl (Aztec) language:

**TÊMOĀ** vt to seek something / buscar algo, o inquirir de algún negocio. This contrasts with TEMŌHUA, the nonactive form of TEMŌ ‘to descend.’

Taken from: *An analytical dictionary of Nahuatl*  
by Frances E. Karttunen



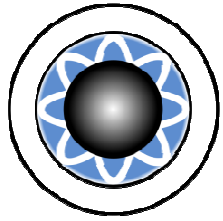
# The TEMOA Project

Tools for Energy Model  
Optimization and Analysis

**Goal:** Create a set of community-driven energy economy optimization models

**Our Approach:**

- Open source code (GNU Public License)
- Open source data (GNU Public License)
- No commercial software dependencies
- Input and output data managed directly with a relational DB
- Data and code stored in a web accessible electronic repository
- A version control system
- Programming environment with links to linear, mixed integer, and non-linear solvers
- Built-in capability for sensitivity and uncertainty analysis



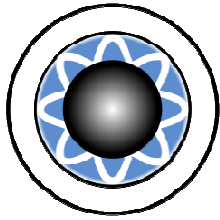
## Other large scientific software projects

We investigated the software engineering practices employed by several large scale scientific software projects:

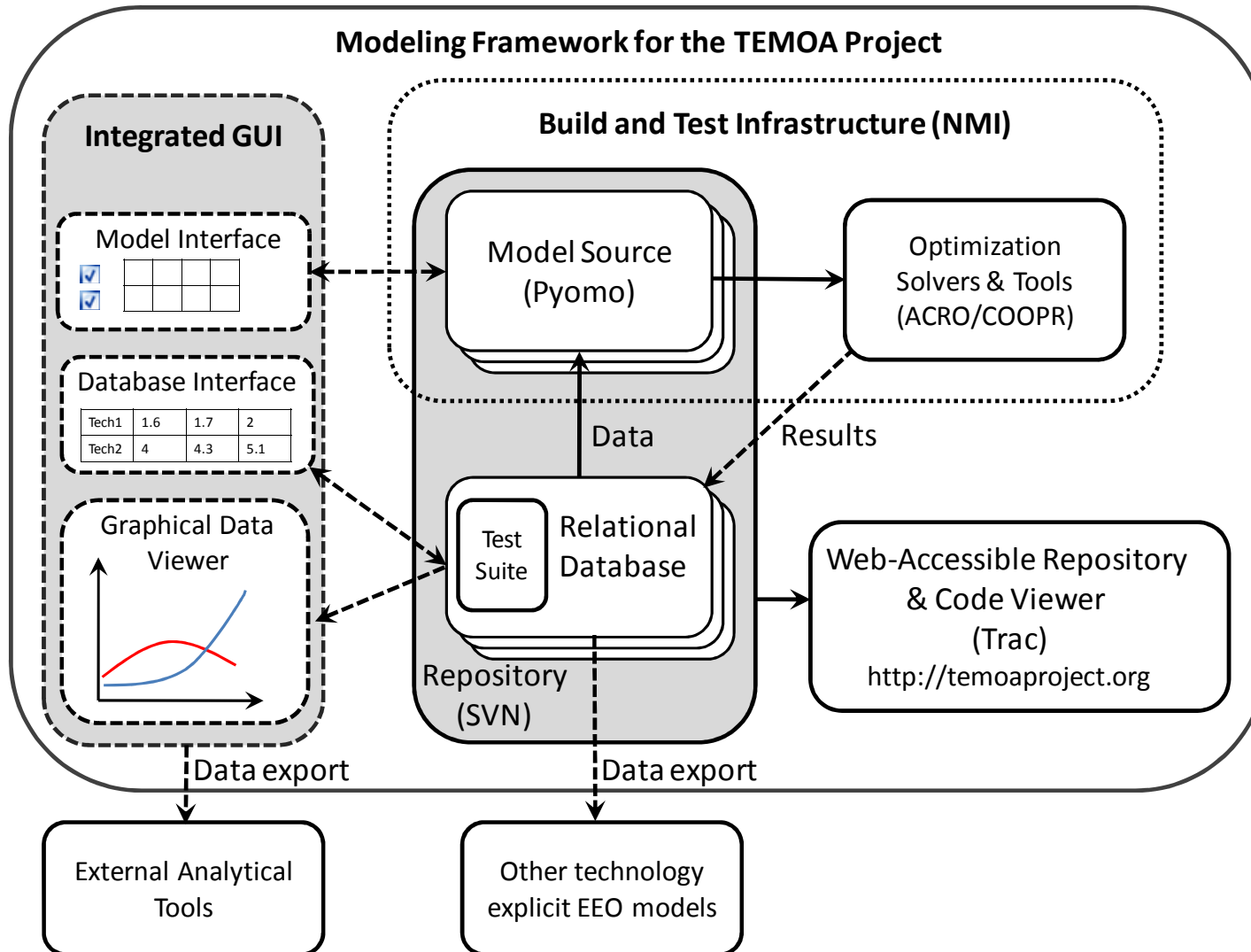
- ❑ MPICH2, a high performance parallel programming library  
<http://www.mcs.anl.gov/research/projects/mpich2/>
- ❑ Portable, Extensible Toolkit for Scientific Computation (PETSc), a parallel differential equations solver library  
<http://www.mcs.anl.gov/petsc/petsc-as/>
- ❑ Community Climate System Model (CCSM), a climate system model  
<http://www.ccsn.ucar.edu/>

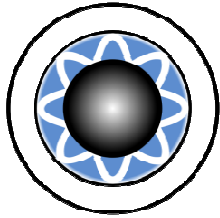
### Lessons:

- ❑ Use a revision control system
- ❑ Develop well-documented examples to introduce the diverse capabilities of the software
- ❑ Use online code browsing



# The Proposed TEMOA Framework





# Version control with Subversion



We are using a version control system called Subversion (SVN)

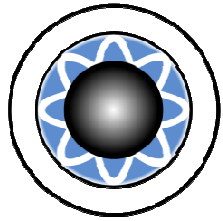
<http://subversion.apache.org/>  
<http://svnbook.red-bean.com/>

Why? Ensure the integrity, sustainability and traceability of changes during the entire software lifecycle.

## **SVN enables:**

- ❑ Multiple developers to work simultaneously on software components; automatic integration of non-conflicting changes
- ❑ Display the modifications to model source code
- ❑ Create software snapshots (releases) that represent well-tested and clearly defined milestones
- ❑ Utilize the release mechanism to take snapshots of the model code and data used to produce research publications.
- ❑ Public access to snapshots of the code and data

Works on all major (Unix, Windows, MacOS) platforms



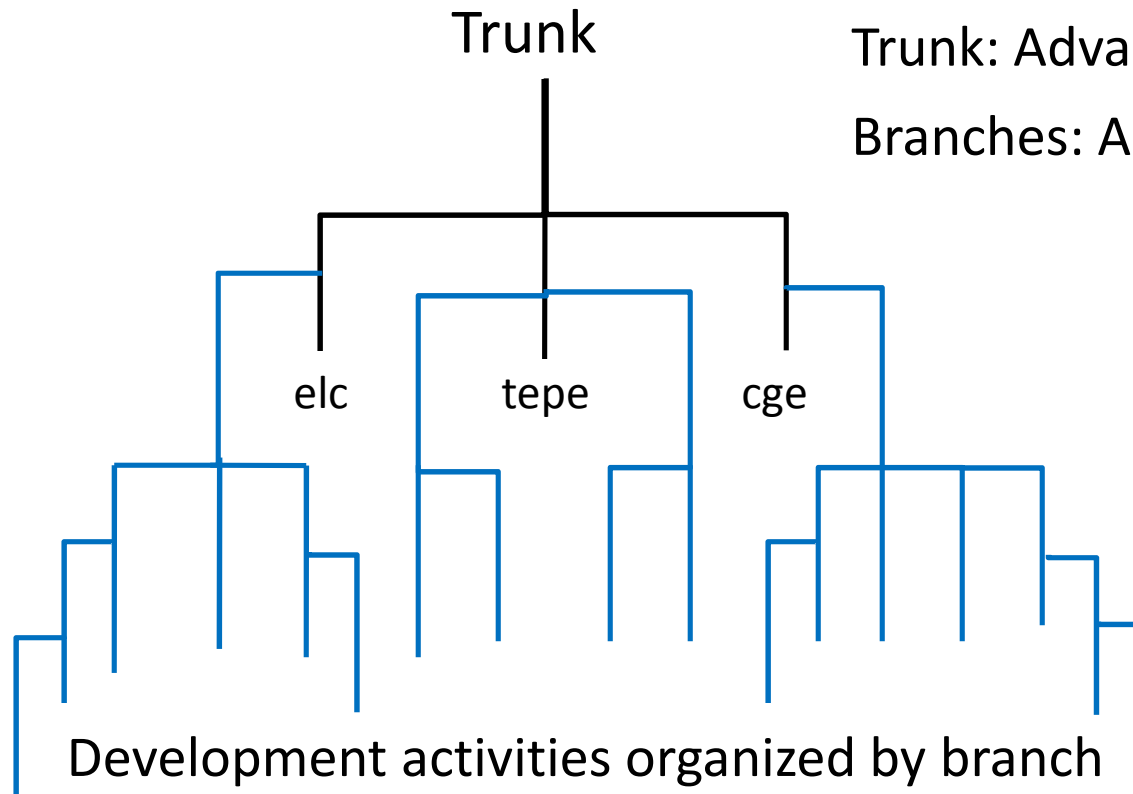
# Mock Repository Layout

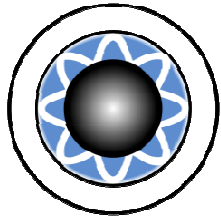
## Organization/Accessibility

Releases: All users

Trunk: Advanced users & developers

Branches: Active Developers





# Python Optimization Modeling Objects

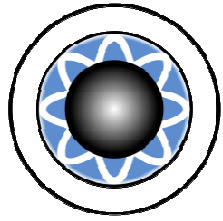
We're developing the model against the Pyomo API

Why Pyomo?

- Uses a full-featured modern programming language
- Rich set of Python libraries that cover nearly every task
- Active development; linkages between Pyomo and custom solvers are being developed within the COmmon Optimization Python Repository (COOPR)

Pyomo developed at Sandia National Laboratories:

<https://software.sandia.gov/trac/coopr/wiki/Pyomo>



# Pyomo vs AMPL

Algebraic Formulation:  $\sum_{t \in \text{tech}} \text{production}_{t,\text{seg}} = \text{dmd}_{\text{seg}}, \forall \text{seg} \in \text{segments}$

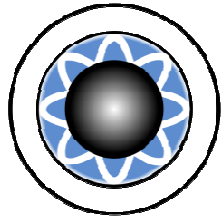
AMPL Formulation:

```
s.t. elc_demand{seg in segments}:  
sum{t in tech[seg]} production[t] = dmd[seg];
```

Pyomo Formulation:

```
def elc_demand (seg, model):  
    "Constraint: Electricity production >= demand each segment"  
    constraint_val = sum(  
        model.production[t]  
        for t in model.tech[ seg ]  
    )  
    return ( constraint_val == model.dmd[ seg ] )  
model.elc_demand = Constraint( model.segments, rule=elc_demand )
```

Use comment blocks to dynamically generate model documentation (via Sphinx). Can embed LaTeX formatting in comments.



# COmmon Optimization Python Repository

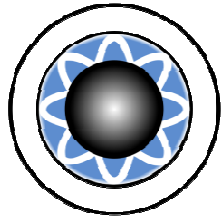
Pyomo is part of the COOPR package, which is in turn part of A Common Repository for Optimizers (ACRO)

Two-language approach: high-level language for model formulation and efficient low-level languages for numerical computations (e.g., C, C++, Fortran)

ACRO includes both libraries developed at Sandia and publicly available third-party libraries (e.g., GLPK and COIN-OR)

Gives us the capability to formulate linear, mixed integer, and non-linear model formulations without commercial solvers

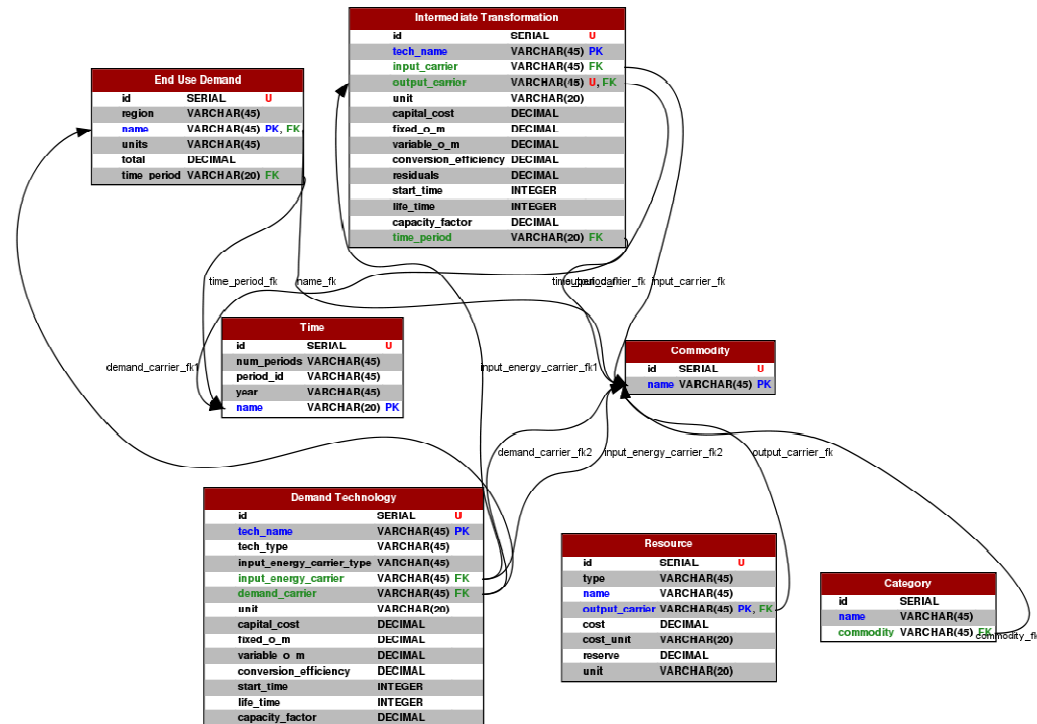
Active collaboration with Discrete Math and Complex Systems Department at Sandia National Laboratories



# Data Management

We will use a relational database to store input and output data → **SQLite**

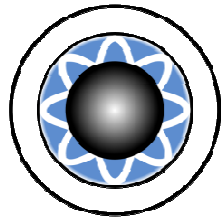
We plan to use the U.S. EPA MARKAL 9R database to populate an initial TEMOA database



Declarative access: Give me the average system-wide CO<sub>2</sub> emissions from 10000 runs in which the installed nuclear capacity exceeded 200 GW:

```

SELECT AVERAGE(co2_output) FROM impact_table, nuclear_table
WHERE impact_table.run = nuclear_table.run
AND nuclear_table.cap > 200;
  
```



# Addressing Uncertainty

We distinguish uncertainty in:

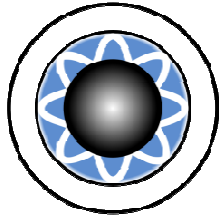
1. Parameters (data, value, stochastic)
2. Structure (how well model equations describe real world)

## **Parametric uncertainty:**

- Fuel prices (esp. crude oil, natural gas, coal)
- Assumed discount rates (technology-specific, global)
- Projected end-use demands (e.g., lighting, heating, mi traveled)

## **Structural uncertainty:**

All real-world factors not encoded in engineering-economic variables (e.g., politics, public attitudes, culture, risk)



## Approach to uncertainty analysis

Use the following techniques in series:

Sensitivity analysis and Monte Carlo simulation

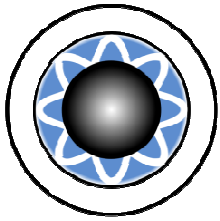
→ Determine key sensitivities

Multi-stage stochastic optimization

→ Develop a hedging strategy

Explore near-optimal, feasible region (MGA)

→ Test robustness of hedging strategy



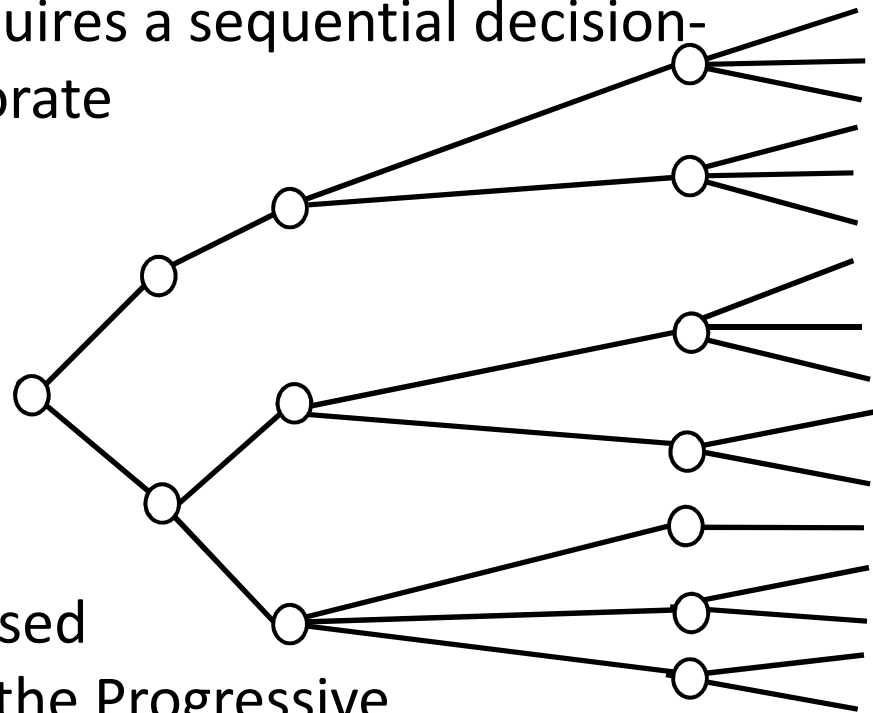
# Stochastic Optimization with PySP

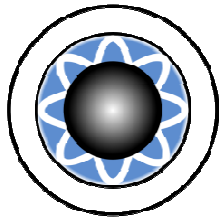
Decision-makers need an “act then learn” approach since uncertainty is not resolved at the time decision is made

Timescale of climate change requires a sequential decision-making process that can incorporate new information

PySP can write and/or solve the extensive form of the stochastic program

PySP includes a generic and highly customizable scenario-based decomposition solver based on the Progressive Hedging algorithm by Rockafellar and Wets (1991)





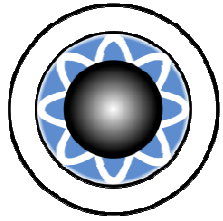
# Possible model formulations

## **Desirable features for energy models:**

- ❑ Multi-stage (greater than 2)
- ❑ Multi-objective (e.g., cost, risk, emissions)
- ❑ Mixed integer (esp. endogenous tech learning)

## **Potential stochastic parameters:**

- ❑ Fuel prices (esp. crude oil, natural gas, coal)
- ❑ Policy targets (e.g., CO<sub>2</sub> constraints, subsidies)
- ❑ Technology performance (e.g., capital cost, thermal eff)
- ❑ End-use demand projections (e.g., heating, cooling)



# Initial focus of the TEMOA project

## Reminder:

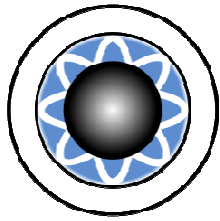
Models are most useful when built to answer specific questions

Initial efforts are focused on the development of a **technology explicit, partial equilibrium model with perfect foresight**, similar to the TIMES model generator: *TEMOA\_tepe*

## Questions to answer:

- ❑ How do different policy architectures affect future energy and environmental outcomes?
- ❑ Are regional variations in energy system development significant?
- ❑ Which energy technologies should be targeted for R&D funding?
- ❑ Given future uncertainty in climate change and policy, what are robust hedging strategies for energy system development?

All from the perspective of a decision-maker at the national level



# TEMOA objective function

$$\begin{aligned}
 \text{Min } C_{\text{TOT}} = & \sum_{tec} \sum_{per} \sum_{iper} \sum_{y=0}^{t(per+1)-t(per)} \left( \begin{aligned} & \left[ C_i(tec, iper) \cdot \frac{r(tec)}{1 - (1 + r(tec))^{-loan(tec)}} \cdot imat(tec, iper, per) + C_f(tec, iper, per) \right] \cdot xc(tec, iper) \\ & + C_m(tec, iper, per) \cdot xu(tec, iper, per) \end{aligned} \right) \\
 & \cdot \frac{1}{(1 + r_g)^{t(per)+y-t(per0)}}
 \end{aligned}$$

Fixed costs = investment + fixed O&M

Marginal cost = fuel + variable O&M

## Sets

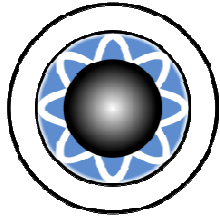
*per* = model time periods  
*iper* = model investment periods  
*tec* = energy technologies

## Decision Variables

*xc(tec, iper)* = installed tech capacity  
*xu(tec, iper, per)* = technology utilization

## Features:

- Separate specifications for the loan period and lifetime of each technology
- Technology-specific interest rates on investment and a separate global discount rate to bring future costs back to the present
- Automatic tracking of fixed and marginal costs by technology vintage
- User-defined model time periods can be different lengths



# Multi-core computation

## **NCSU Cluster “Cygnus”:**

11 nodes, each with 2 AMD quad-core Opteron processors (2.0 GHz with 512 KB Cache/core)

1.8 TB of storage

176 GB memory

OpenSuse 10.3 (Linux)

FLOPS = 704 Gigaflops

1 GigE interconnect

## **Current electric sector model:**

technologies: 19

Investment periods = 10

Model periods = 10

Capacity variables= 190

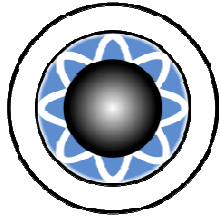
Activity variables =2600

Constraints: 2676

Pyomo+GLPK (model gen + solution) = 5.1 sec.

Regionalizing data and building a reference energy system will increase model size by a factor of at least 100.

Multi-region MARKAL model is generated + solved in ~20 mins on a fast PC



## Near-term development timeline

**Now:** Working electric sector model

**Summer:** Incorporate progressive hedging and MGA  
Build constraints to represent energy system

**Fall:** Functioning schema and test DB

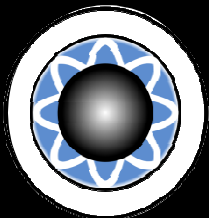
Looking for interested modelers to help with our model formulation or build your own “branch”

Follow progress and/or join the effort at: <http://temoaproject.org>

# Thank You For Your Time

“It is important not to let the perfect become the enemy of the good, even when you can agree on what perfect is. Doubly so when you can't. As unpleasant as it is to be trapped by past mistakes, you can't make any progress by being afraid of your own shadow during design.”

—Greg Hudson, *Subversion* developer



<http://temoaproject.org>