Data-driven discovery and approaches to model evaluation

Louis J. Gross

National Institute for Mathematical and Biological Synthesis

Departments of Ecology and Evolutionary Biology and Mathematics and The Institute for Environmental Modeling

University of Tennessee, Knoxville

Supported by NSF Award DBI-1300426, DOE ORNL#4000110008

NIMBioS

National Institute for Mathematical and Biological Synthesis

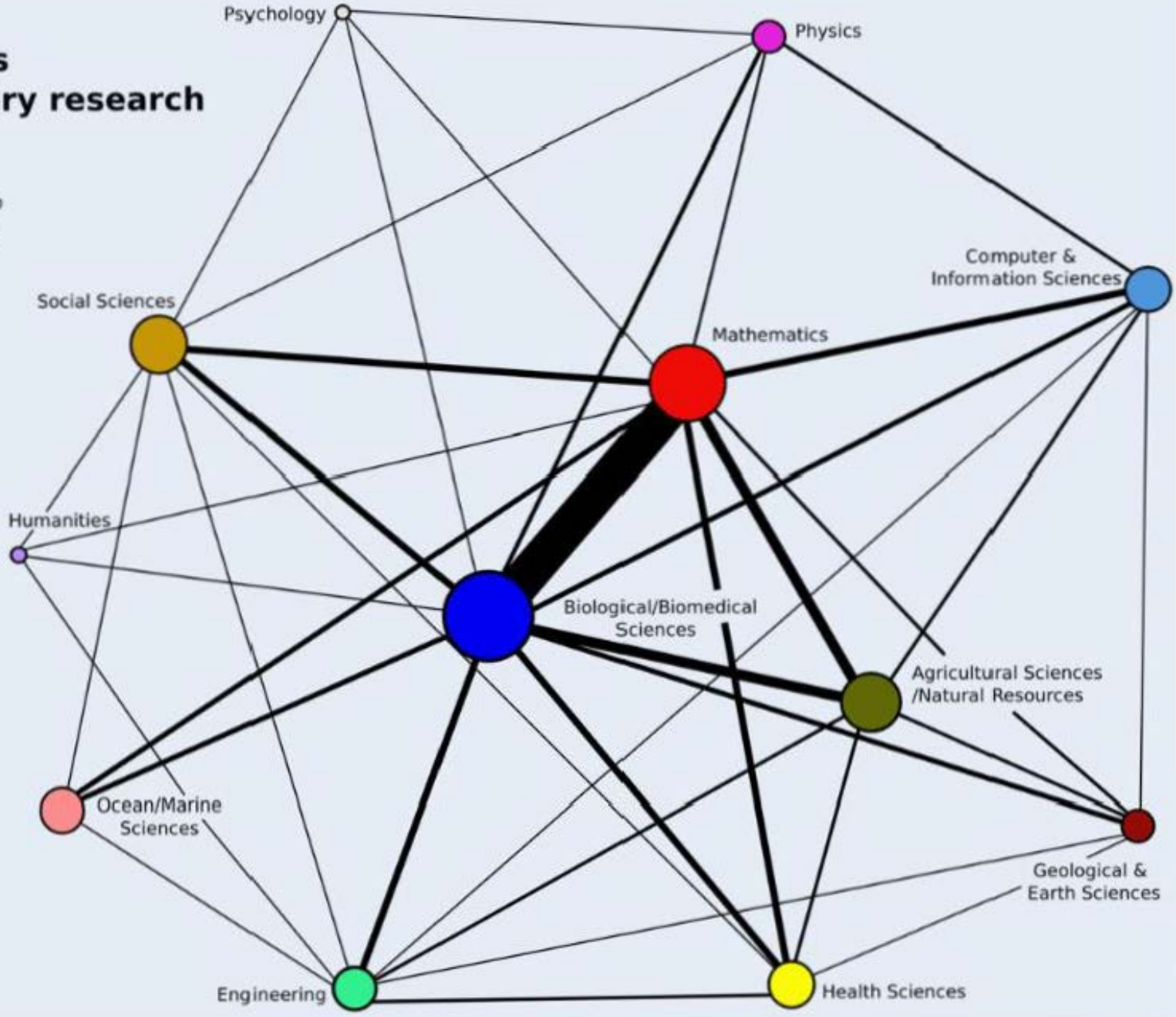
- Foster new collaborative efforts to investigate fundamental and applied questions arising in biology using appropriate mathematical and computational methods
- Enhance the essential human capacity to analyze complex biological questions and develop necessary new mathematics
- Encourage broader public appreciation of the unity of science and mathematics.

Deadlines for support requests: March 1, Sept. 1

NIMBioS fosters cross-disciplinary research

Node size
number of Working Group participants in a given research area, where the node radius is the log number of participants

Line size
number of collaborations between research areas within Working Groups



Working Groups focus on major scientific questions at the interface between biology and mathematics that require insights from diverse researchers who meet several times over a two-year period.

NIMBioS

National Institute for Mathematical and Biological Synthesis

Predictive Models for Ecological Risk Assessment

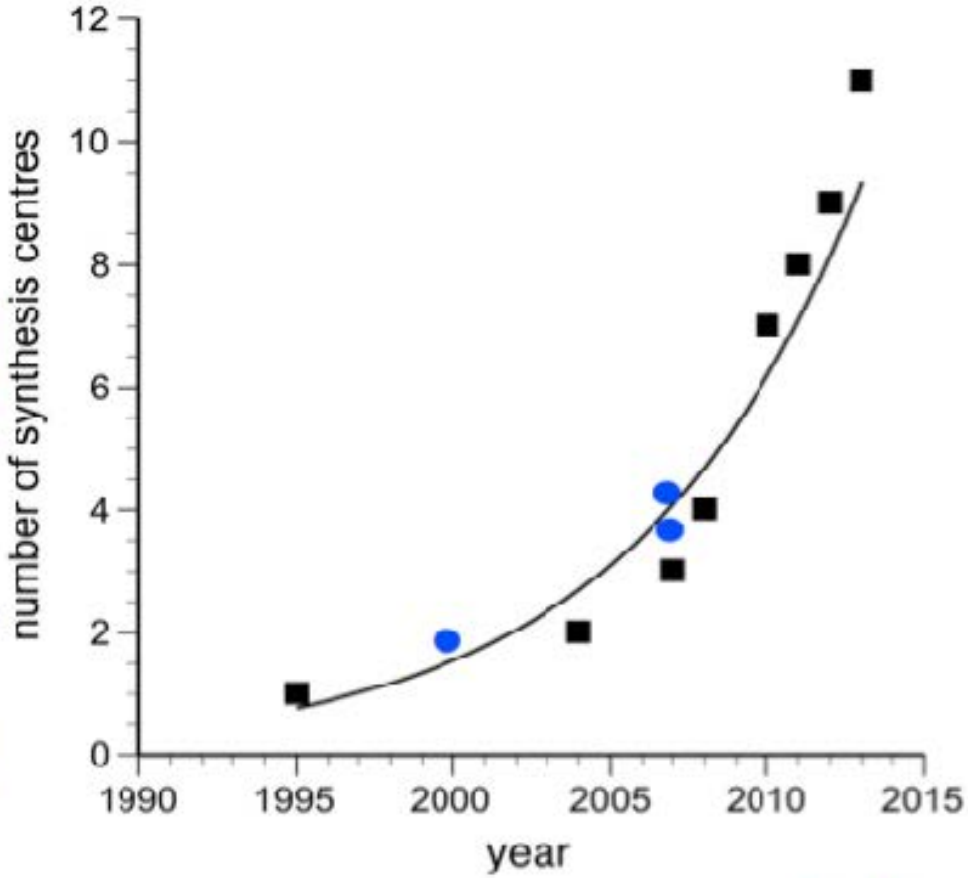
A NIMBioS Investigative Workshop

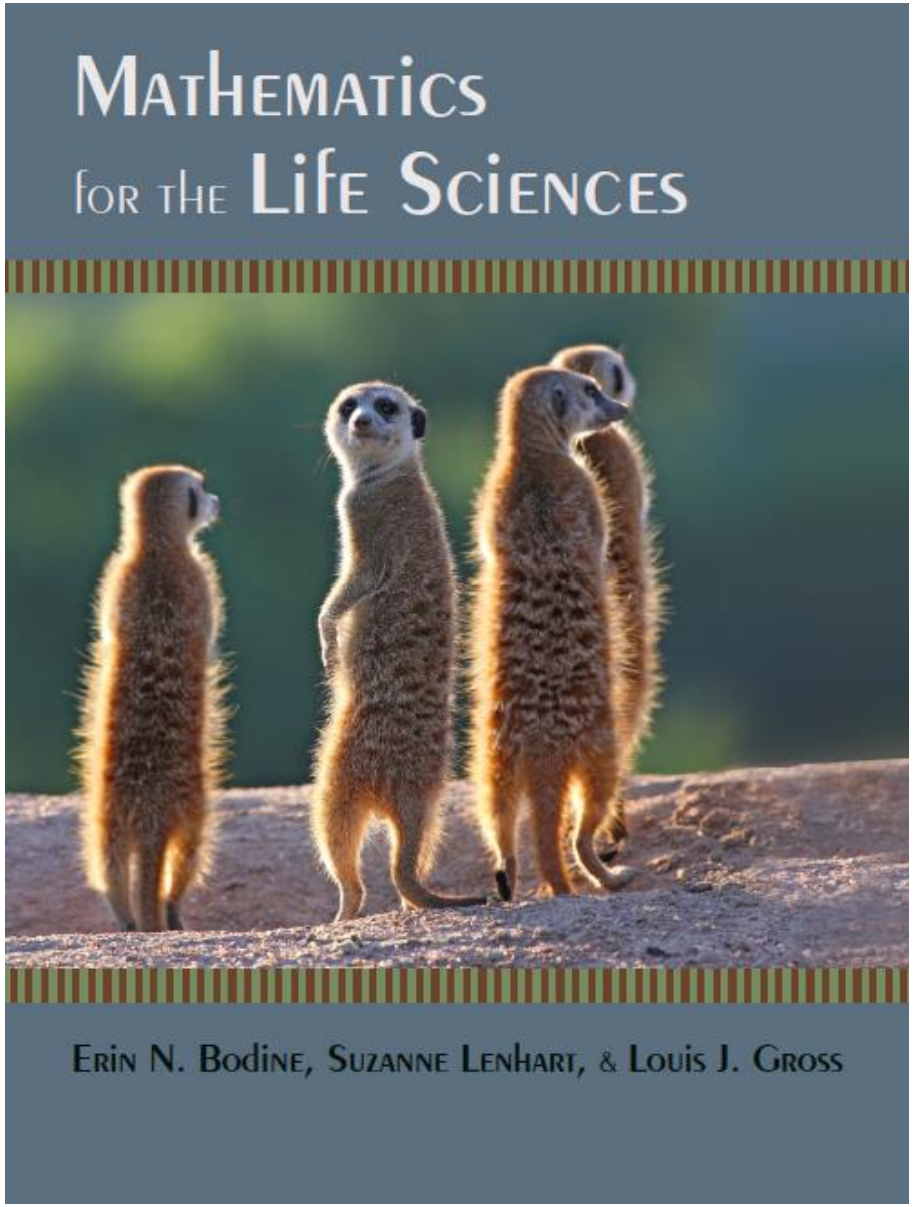
April 28-30, 2014

NIMBioS at the Univ. of Tennessee, Knoxville

This workshop will bring together a multi-disciplinary group of molecular and cell biologists, physiologists, ecologists, mathematicians, computational biologists, and statisticians to explore the challenges and opportunities for developing and implementing models specifically designed to mechanistically link between levels of biological organization so as to inform ecological risk assessment and ultimately environmental policy and management. The focus will be on predictive systems models in which properties at higher levels of organization emerge from the dynamics of processes occurring at lower levels of organization.

Synthesis Centers around the world





MathForTheLifeSciences.com

Princeton University Press – Aug. 2014

Overview

- **Models and science**
- **Objectives of models**
- **Data driven discovery**
- **Constraints on models**
- **Model evaluation**
- **Risk assessment examples**
  - **RAIS – Risk Assessment Information System**
  - **SADA - Spatial Analysis and Decision Assistance**
- **Take-home lessons**

What is science?

Science is thought to be a process of pure reductionism, taking the meaning out of mystery, explaining everything away, concentrating all our attention on measuring things and counting them up. It is not like this at all. The scientific method is guesswork, the making up of stories. The difference between this and other imaginative works of the human mind is that science is then obliged to find out whether the guesses are correct, the stories true. Curiosity drives the enterprise, and the open acknowledgement of ignorance.

Lewis Thomas - Sierra Club Bulletin, March/April 1982, P. 52

What is science?

Science is thought to be a process of pure reductionism, taking the meaning out of mystery, explaining everything away, concentrating all our attention on measuring things and counting them up. It is not like this at all. The scientific method is guesswork, the making up of stories. The difference between this and other imaginative works of the human mind is that science is then obliged to find out whether the guesses are correct, the stories true. Curiosity drives the enterprise, and the open acknowledgement of ignorance.

Models

Lewis Thomas - Sierra Club Bulletin, March/April 1982, P. 52

What is science?

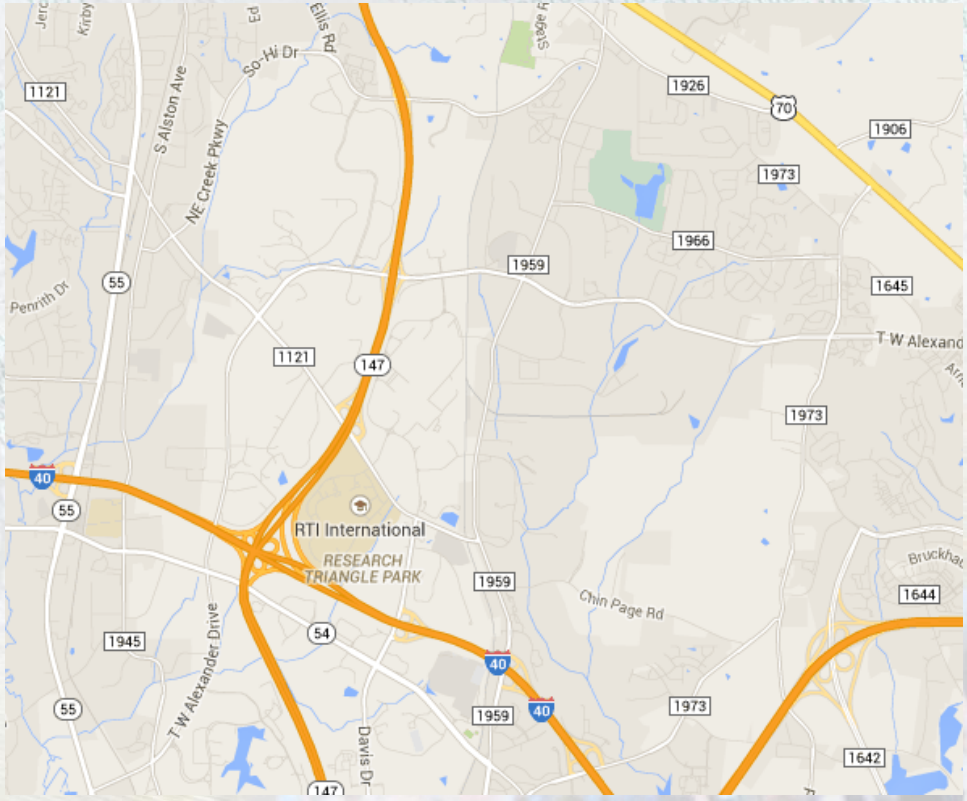
Science is thought to be a process of pure reductionism, taking the meaning out of mystery, explaining everything away, concentrating all our attention on measuring things and counting them up. It is not like this at all. The scientific method is guesswork, the making up of stories. The difference between this and other imaginative works of the human mind is that science is then obliged to find out whether the guesses are correct, the stories true. Curiosity drives the enterprise, and the open acknowledgement of ignorance.

Data

Lewis Thomas - Sierra Club Bulletin, March/April 1982, P. 52

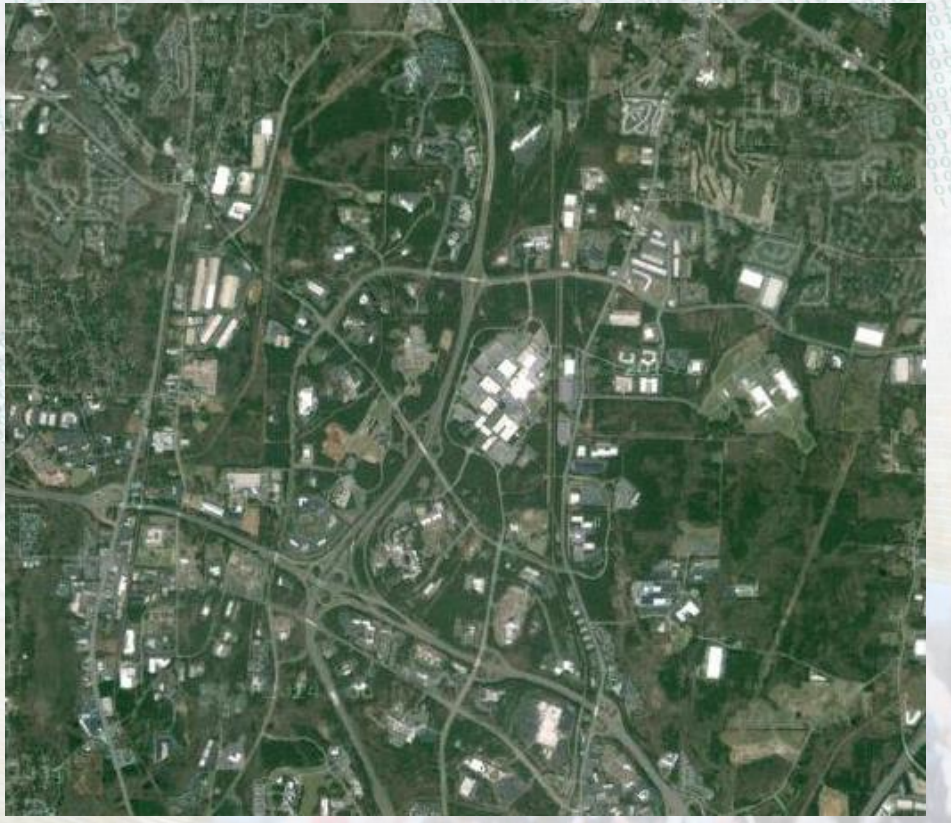
The “stories” in science are models

A model is a simplification of reality. Think of it as a map - it includes some features that represent what we observe but not others. Modeling is the process of selective ignorance - we select what to include and what to ignore.



The “stories” in science are models

A model is a simplification of reality. Think of it as a map - it includes some features that represent what we observe but not others. Modeling is the process of selective ignorance - we select what to include and what to ignore.



You make models all the time:

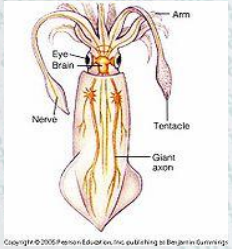
What decision do you make when faced with:

The "best" model for you may not be the "best" model for someone else.

Models in Biology

Physiology

Disease



Neurobiology

Microbiology

Development

Genetics

Ecosystems

Spatial-epidemiology model with vaccination

$$(1)\quad \frac{dS_i}{dt} = -\beta_i S_i I_i - \gamma\sigma_i S_i + \sum_{j,j\neq i}^{n} a_{ji} S_j - \sum_{j,j\neq i}^{n} a_{ij} S_i - \mu_S S_i$$

$$(2)\quad \frac{dI_i}{dt} = \beta_i S_i I_i + \sum_{j,j\neq i}^{n} c_{ji} I_j - \sum_{j,j\neq i}^{n} c_{ij} I_i - \mu_I I_i$$

$$(3)\quad \frac{dR_i}{dt} = \gamma\sigma_i S_i + \sum_{j,j\neq i}^{n} a_{ji} R_j - \sum_{j,j\neq i}^{n} a_{ij} R_i - \mu_R R_i$$

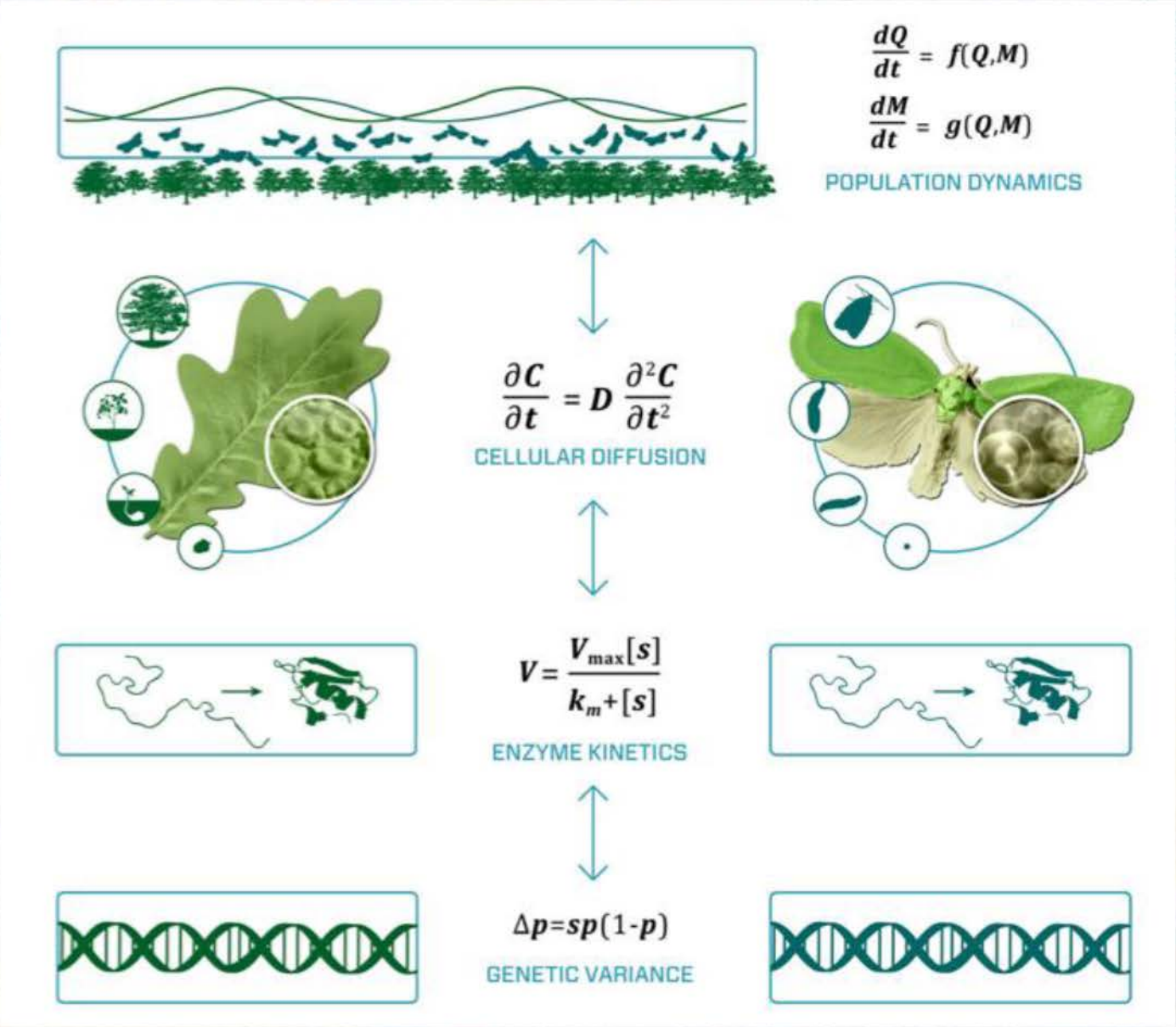
ICs :

$$(4)\quad S_i(0) = S_{i0}, \quad I_i(0) = I_{i0}, \quad R_i(0) = R_{i0}$$

Control(Vaccination) :

$$(5)\quad 0 \le \sigma_i \le \sigma_{\max} \qquad \text{for } i = 1, 2, \ldots, n.$$

Models across multiple scales



Environmental Modeling

Species densities

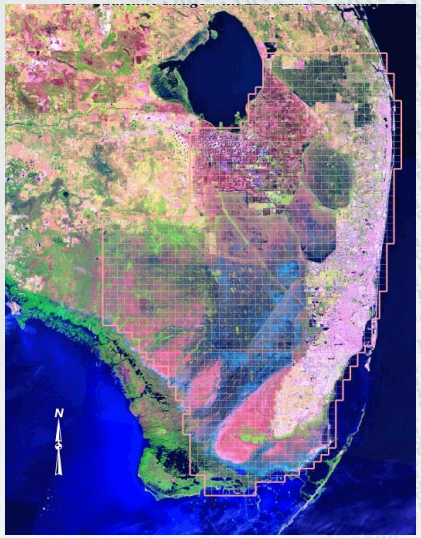
Animal telemetry

Physical conditions

Data sources

GIS map layers (Vegetation, hydrology, elevation),Weather, Roads, Species densities

Monitoring



Models

Management input

Statistical

Differential equations

Matrix

Agent-based

Harvest regulation

Water control

Reserve design

Evaluation/Analysis

Simulation

Visualization, corroboration, sensitivity, uncertainty

Matlab, C++, Distributed, Parallel

Objectives of Models

There are many reasons to use a model aside from prediction:

1. Suggest observations and experiments

2. Provide a framework to assemble bodies of facts/observations - standardize data collection

3. "Allows us to imagine and explore a wider range of worlds than ours, giving new perceptions and questions about how our world came to be as it is" F. Jacob - The Possible and the Actual, 1982

4. Clarifies hypotheses and chains of argument

5. Identifies key components in systems

6. Allow investigation while accounting for societal or ethical constraints

Objectives of Models

7. Allows simultaneous consideration of spatial and temporal change

8. Provides a means to extrapolate or interpolate to situations for which data can not easily be obtained

9. Prompts tentative and testable hypotheses

10. Serves as a guide to decision making in circumstances where action cannot wait for detailed studies or those studies are not feasible

11. Provides an antidote to the helpless feeling that the world is too complex to understand in any generality - provides a means to get at general patterns and trends

But this view is being challenged

WIRED MAGAZINE: 16.07

SCIENCE : DISCOVERIES

The End of Theory: The Data Deluge Makes the Scientific Method Obsolete

By Chris Anderson 06.23.08

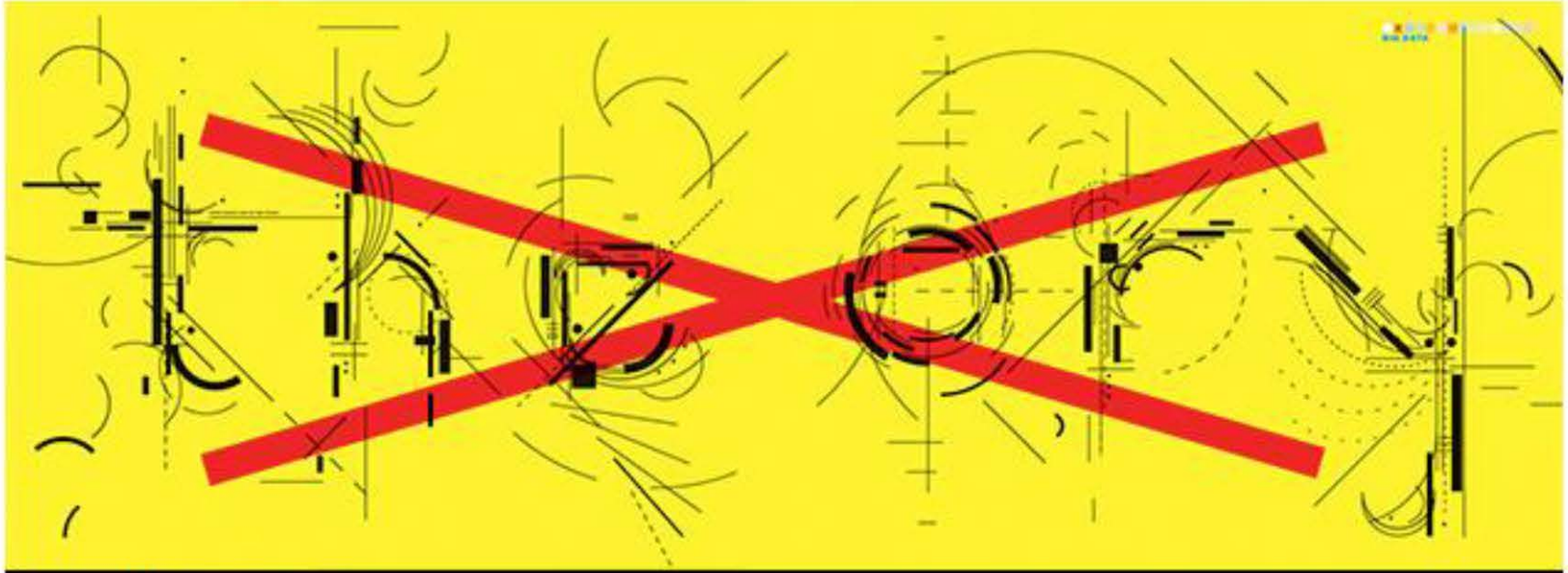


Illustration: Marian Bantjes

THE PETABYTE AGE:

"**All models are wrong,** but some are useful."

The essence is: Let the Data Tell it's Own Story – who needs generality!

Report | McKinsey Global Institute

Big data: The next frontier for innovation, competition, and productivity

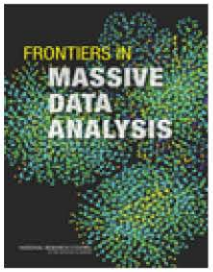
May 2011 | by James Manyika, Michael Chui, Brad Brown, Jacques Bughin, Richard Dobbs, Charles Roxburgh, Angela Hung Byers

ASA AMERICAN STATISTICAL ASSOCIATION

Discovery with Data: Leveraging Statistics with Computer Science to Transform Science and Society

July 2, 2014

A Working Group of the American Statistical Association[1]



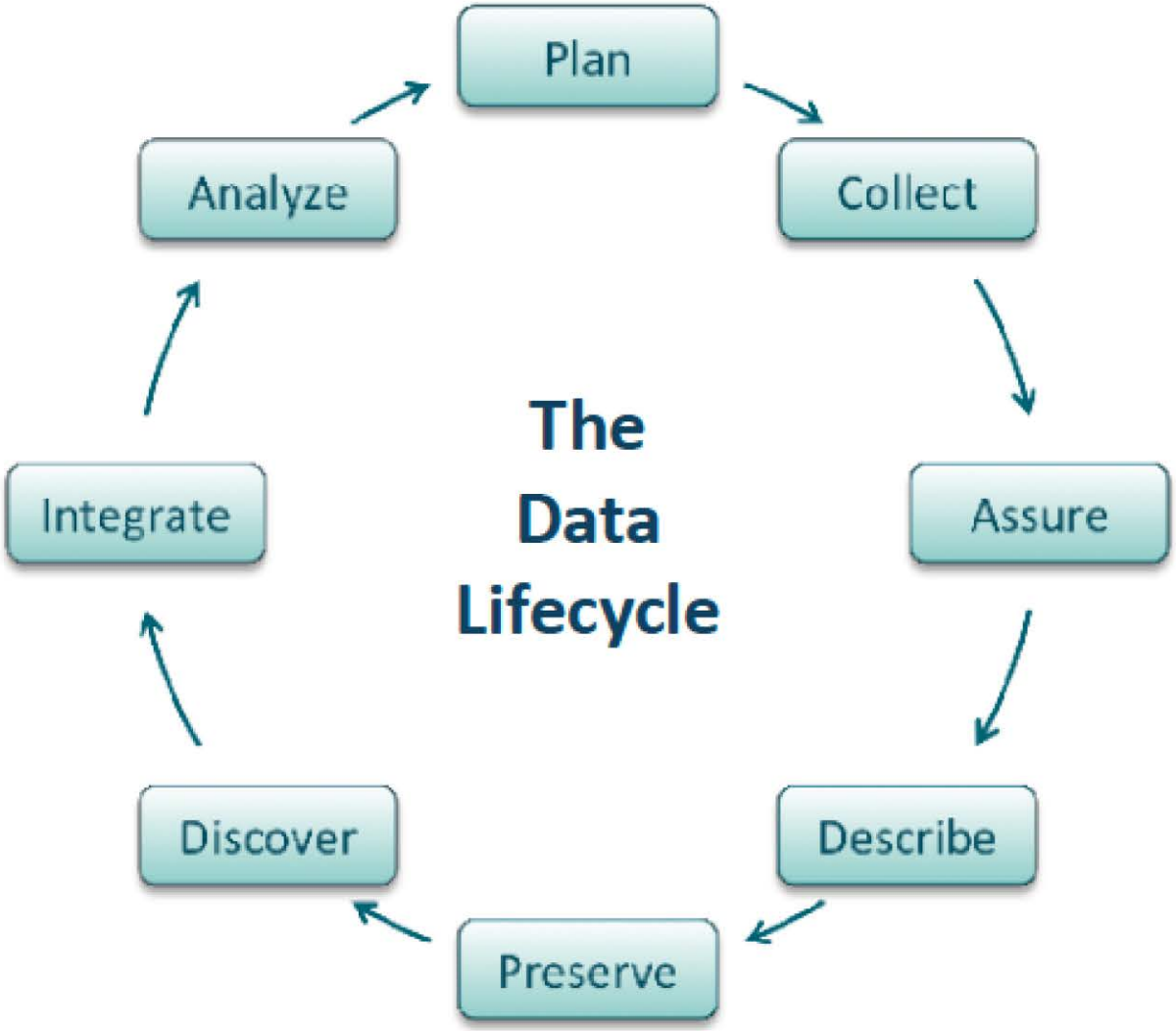
TECHNOLOGY FEATURE

THE BIG CHALLENGES OF BIG DATA

As they grapple with increasingly large data sets, biologists and computer scientists uncork new bottlenecks.

Cyclical Aspects of Data

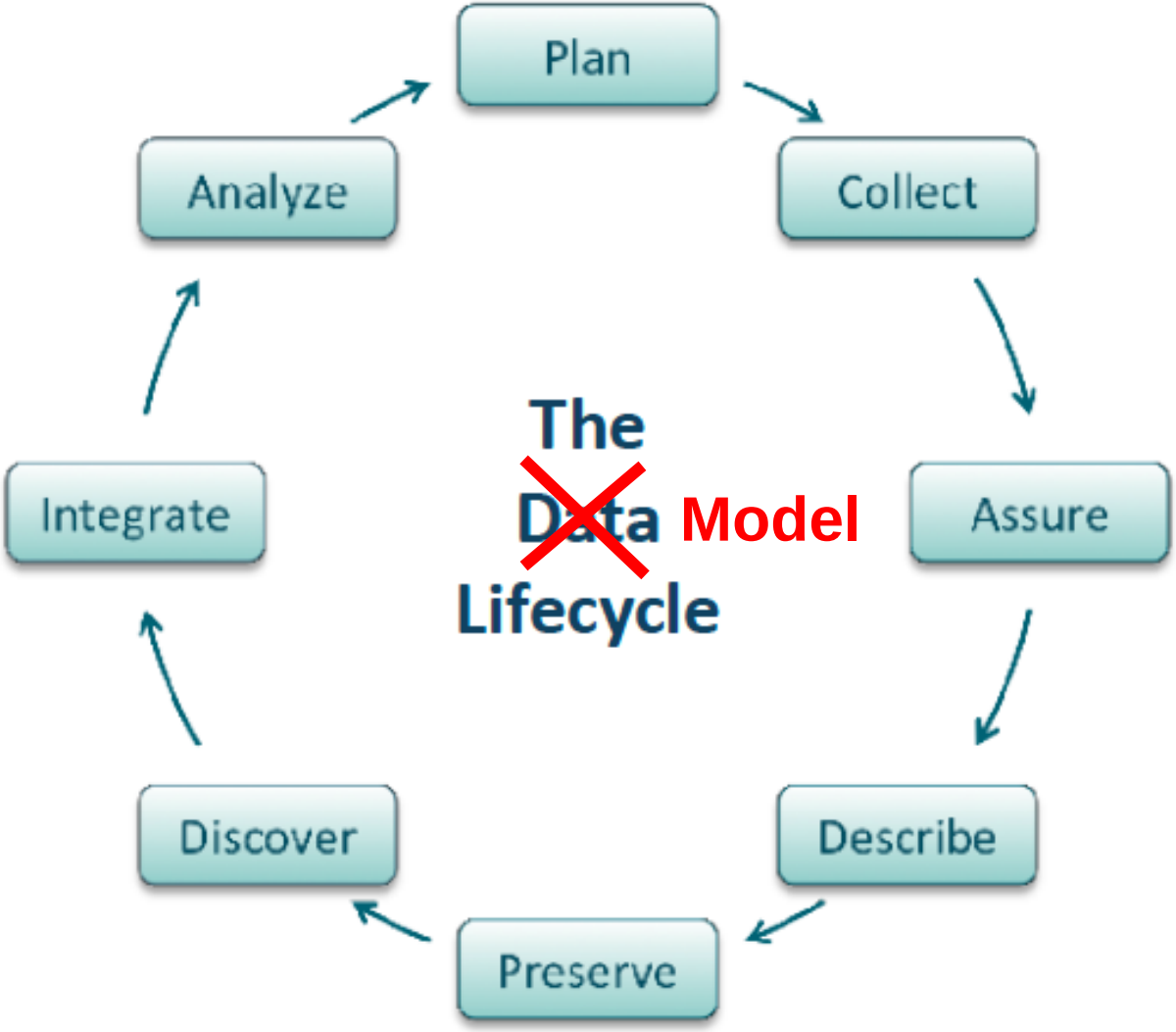
The stages through which well-managed data passes from project inception to conclusion.



From DataONE.org

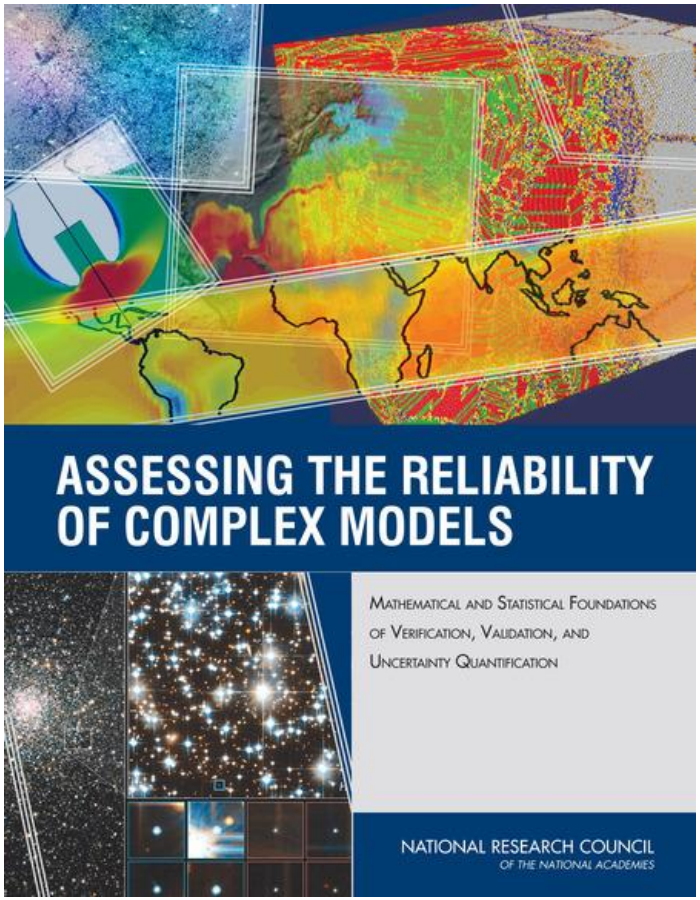
Cyclical Aspects of ~~Data~~ Models

The stages through which well-managed data passes from project inception to conclusion.



From DataONE.org

If you think Big Data is challenging, what about Big Models!



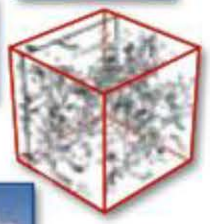
All of the concerns regarding big data (heterogeneity, data quantity, data quality, curation, metadata characterization) also apply to the complex models applied to provide regulatory guidance.

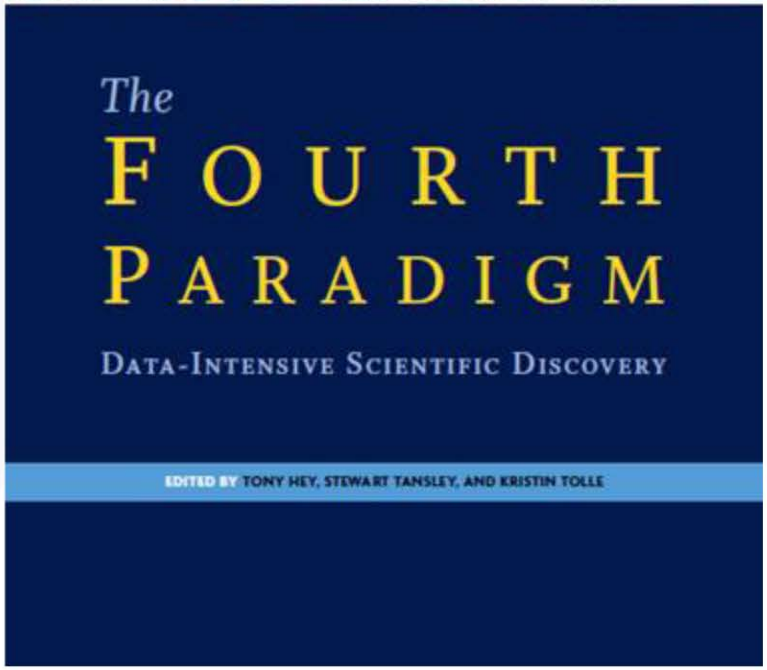
One Vision for Data-Exploration

Science Paradigms

- Thousand years ago:
  science was **empirical**
  describing natural phenomena
- Last few hundred years:
  theoretical branch
  using models, generalizations
- Last few decades:
  a **computational** branch
  simulating complex phenomena
- Today: **data exploration** (eScience)
  unify theory, experiment, and simulation
  - Data captured by instruments or generated by simulator
  - Processed by software
  - Information/knowledge stored in computer
  - Scientist analyzes database/files using data management and statistics

$$\left(\frac{\dot{a}}{a}\right)^2 = \frac{4\pi G p}{3} - K\frac{c^2}{a^2}$$





Constraints on models

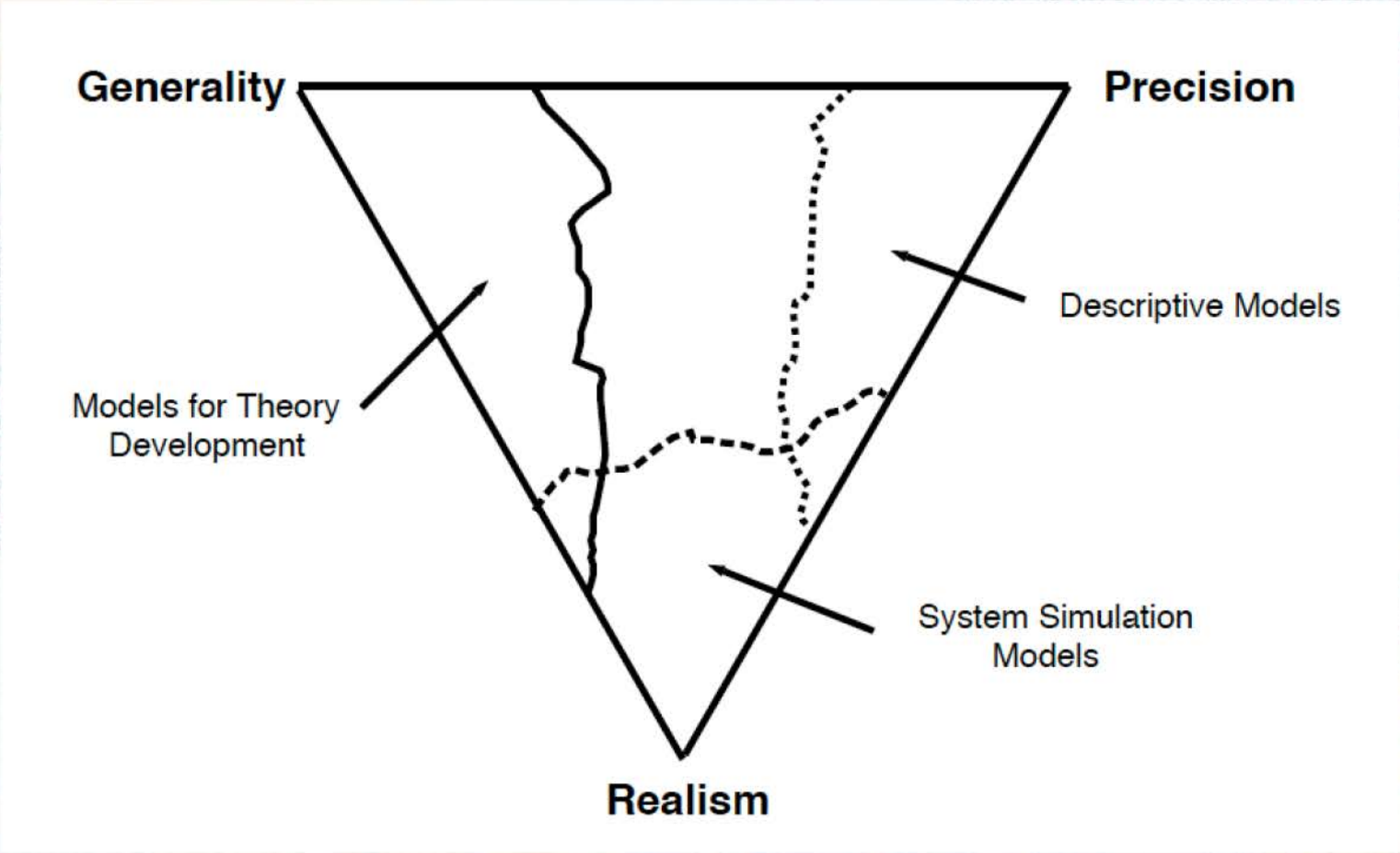
***Data constraints*: Available data may not be sufficient to specify appropriate functional forms, interrelationships, or parameters. May force aggregation of components. May not be sufficient to elaborate criteria for evaluation of model performance.**

***Effort constraints*: Resource constraints may limit the amount of detail it is feasible to include. Limits time modelers and collaborators may invest as well as pressure to produce results.**

***Computational constraints:* Despite great enhancements in computational resources, there are many problems still not feasible to carry out computationally.**

***Other constraints*: ethical or other societal considerations.**

Models and tradeoffs



No one model can do everything!

Model evaluation – some terminology

Verification - model behaves as intended, i.e. equations correctly represent assumptions; equations are self-consistent and dimensionally correct. Analysis is correct. Coding is correct - there are no bugs.

Calibration - use of data to determine parameters so the model "agrees" with data. This is specific to a given criteria for accuracy. Some call this Tuning or Curve-fitting.

Corroboration - model is in agreement with a set of data independent from that used to construct and calibrate it.

Validation - model is in agreement with real system it represents with respect to the specific purposes for which it was constructed. Thus there is an implied notion of accuracy and domain of applicability.

Evaluation (testing) - appropriateness to objectives; utility; plausibility; elegance; simplicity; flexibility.

Evaluating different types of models

Models for theory development –

General, some realism, little precision.

Make qualitative comparisons to patterns, not quantitative ones, over some parameter space. No calibration or corroboration performed, except theoretical corroboration (meaning that model agrees with the general body of theory in the field).

Evaluating different types of models

Descriptive models-

Precise, little realism, not general

Statistical hypothesis testing; time series analysis methods applied.

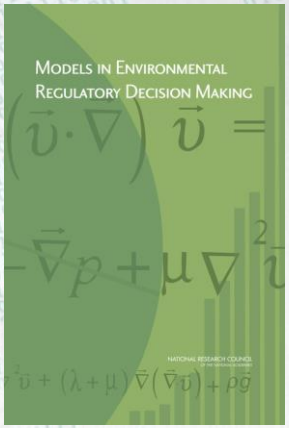
Models for specific systems -

Realism, some precision, not general

Quantitative comparisons, constrained by available data. Compare component-by-component if data are available.

Evaluation rather than "Validation"

The NRC report on *Models in Environmental Regulatory Decision Making* avoided the use of the term validation for several reasons including:



- **The confusion over the term in different communities**
- **The prevalent perception that a "valid" model exists outside of the objective for which it was developed – and these objectives may not be assessed through "domains of applicability"**
- **The implication that validation is "static" whereas the report recommends a life-cycle for ongoing model evaluation and a plan for carrying this out**

Models and evaluation

- **Given the many objectives for models, we should expect many diverse criteria for evaluating whether a model is useful**
- **Before developing a model in any detail, criteria should be established for evaluating its use**
- **Evaluation should account for constraints of Data Availability, Effort and Resources, Computation**
- **Include evaluation of alternative approaches based on these constraints to assess most appropriate methods, decide level of detail, scale, and what to ignore (e.g. modeling is a process of "selective ignorance" and the art is in deciding what to include and what to exclude).**

Animal models

The most frequent use of the term “model” in connection with biology concerns animal models used as proxies for humans to investigate medical questions. General guidance on evaluating animal models is prevalent throughout the literature (a specific disease model faithfully mimics the human disease, a model system is appropriate for the human system being modeled) but there is little direct methodology to evaluate a particular animal model.

Animal models

A recent detailed analysis of the use of animal models (Wall and Shani, 2008) argues that "on average, the extrapolated results from studies using tens of millions of animals fail to accurately predict human responses" despite the fact that these studies have been invaluable for investigating general processes and biological pathways. They note that the recommendations on animal models make theoretical sense but often lack practicality.

R. Wall and M. Shani (2008) Are animal models as good as we think? Theriogenology 69: 2–9

Why so little emphasis on evaluation?

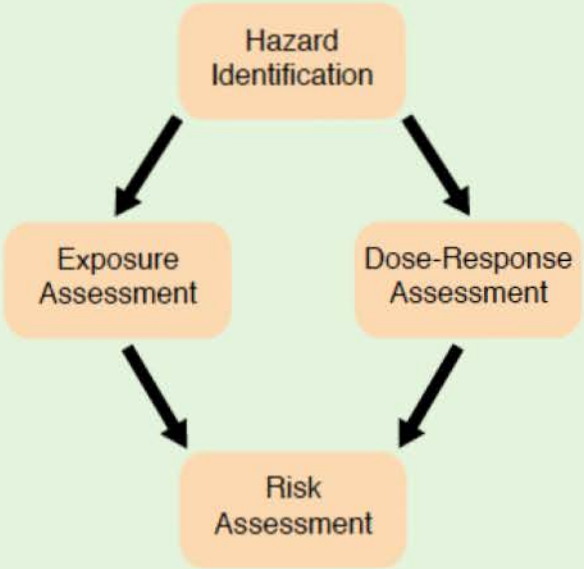
1. It's difficult and requires potentially different skill sets from those constructing and using models.
2. Science is very much a human enterprise and it is natural that once one has devoted considerable effort to developing a particular model, it is difficult to critique yourself.
3. Modern settings with a great amount of team effort to develop models or experimental protocols can constrain individuals who do not wish to be an outcast in a lab.

Take home lessons

- **Model evaluation for all types of biological models is relatively rare.**
- **Set criteria for model evaluation prior to expending a lot of effort on a model.**
- **Tie evaluation criteria to model objectives.**
- **Encourage consideration of evaluation in all your educational initiatives.**
- **Multiple models are good – encourage this.**
- **Consider whether an evaluation has been done or discussed whenever you review a paper or grant proposal.**

What is Risk Assessment?

Risk assessment is a process for characterizing the nature and magnitude of health risks to humans (e.g., residents, workers, recreational visitors) and ecological receptors (e.g., bird, fish, wildlife) from chemical contaminants and other stressors that may be present in the environment. Risk assessment involves four major steps:



1. ***Hazard Identification**—an examination of whether a stressor has the potential to cause harm to humans or ecological systems*
2. ***Dose-Response Assessment**—an examination of the numerical relationship between exposure and effects*
3. ***Exposure Assessment**—an examination of what is known about the frequency, timing, and levels of contact with a stressor*
4. ***Risk Characterization**—an examination of what is known about the frequency, timing, an levels of contact with a stressor*

NRC. 2014. Advancing Risk Assessment with Systems Biology

Risk Assessment Information System (RAIS) - Public access website for all things concerning environmental risk assessment: toxicity values, chemical parameters, PRG calculation, risk calculation, ARARs, ecological benchmarks.

Fred Dolislager

Leslie Galloway

Debra Stewart

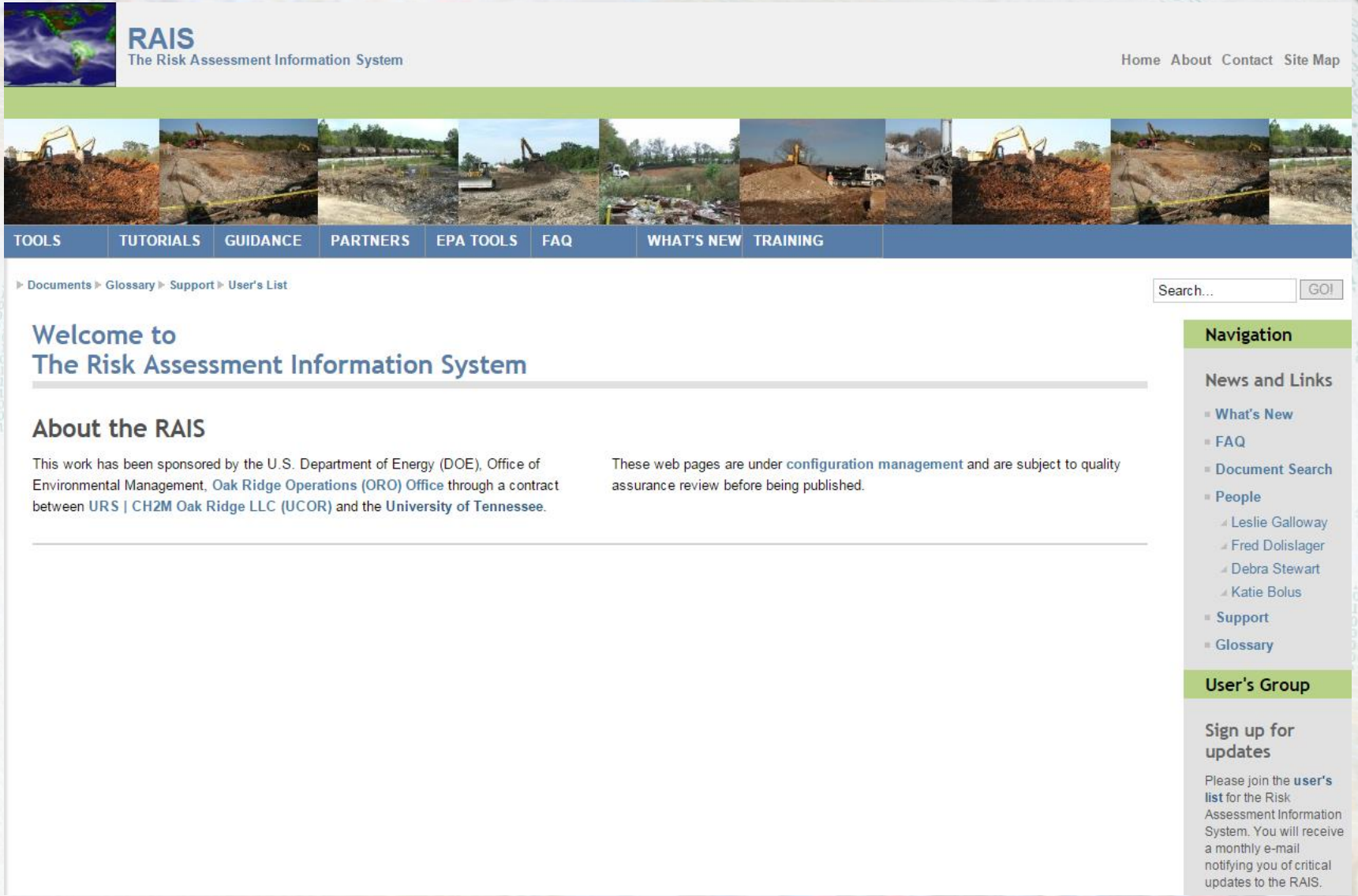
The University of Tennessee

The Institute for Environmental Modeling

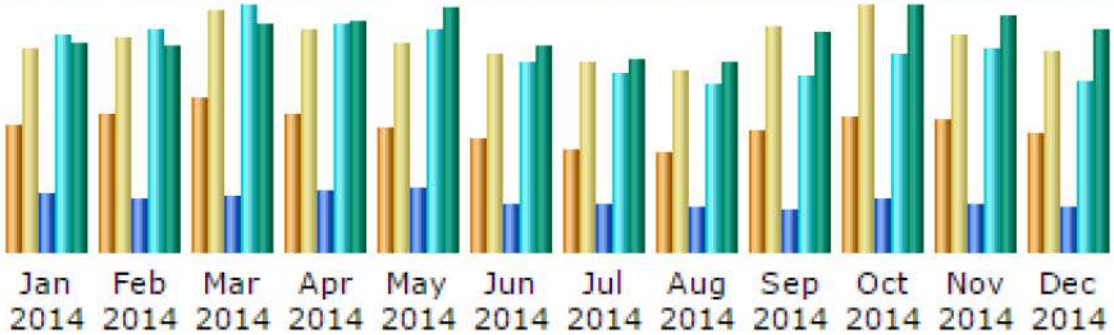
RAIS Progression

- ORNL developed the first instance with sponsorship from DOE in 1996.
- RAIS consolidated and unified risk assessment procedures and practices by hosting documents, equations and databases on the Oak Ridge Reservation.
- Developed our tools through IAG with EPA to provide national guidance.
- Gave the tools utility for international use by allowing users to modify all of our parameters.

RAIS Homepage http://rais.ornl.gov/

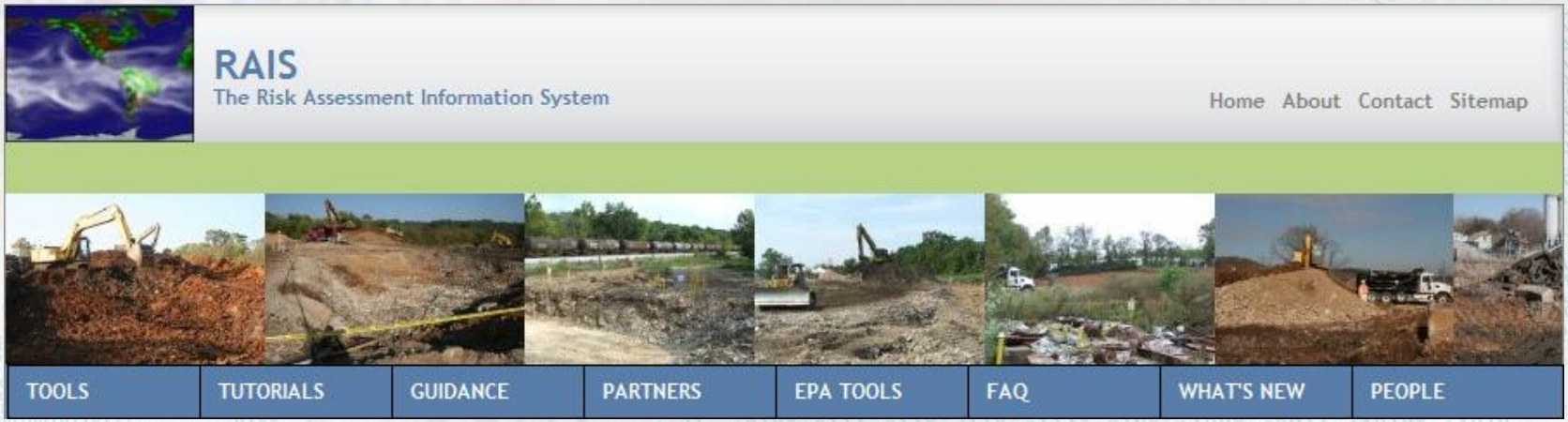


RAIS Usage



Month	Unique visitors	Number of visits	Pages	Hits	Bandwidth
Jan 2014	5,365	8,652	59,963	225,362	29.08 GB
Feb 2014	5,823	9,151	56,396	230,608	28.56 GB
Mar 2014	6,575	10,304	59,400	254,756	31.69 GB
Apr 2014	5,925	9,476	63,577	236,714	31.81 GB
May 2014	5,266	8,886	66,772	229,446	33.83 GB
Jun 2014	4,811	8,433	48,588	197,129	28.53 GB
Jul 2014	4,319	8,077	48,542	186,123	26.75 GB
Aug 2014	4,276	7,698	45,936	173,105	26.13 GB
Sep 2014	5,190	9,579	42,845	183,718	30.46 GB
Oct 2014	5,753	10,462	54,556	203,850	33.99 GB
Nov 2014	5,671	9,280	49,101	209,911	32.63 GB
Dec 2014	5,039	8,537	47,103	175,697	30.71 GB
Total	64,013	108,535	642,779	2,506,419	364.17 GB

RAIS Tool Bar Highlights



- Toxicity Profiles
- Toxicity Values»
- Toxicity Metadata»
- Parameters»
- Chemical Data Profiles
- PRGs»
- Risk Models»
- ORO Risk Information»
- Ecological Benchmarks
- Background Values»
- ARAR Search
- Gamma Radiation Instrument Response Tool»
- Radionuclide Decay Chain Tool
- Transport Calculators»

- **Toxicity Profiles (70 Analytes)**
- **Toxicity Values (800 Chemicals and 1368 Radionuclides)**
- **Toxicity Metadata (37 Pieces of Study Information)**
- **Chemical Factors (22 variables)**
- **PRGs (5 land uses) (5 media)**
- **Risk models (5 land uses)**
- **Ecological Benchmarks (4 media) (82 benchmarks)**
- **ARAR (Federal + 13 US States)**
- **Soil to Groundwater**
- **Background Values for metals**
- **Radionuclide decay chain generator**
- **Air and soil transport models**

Risk Assessment Steps: Data Evaluation

RAIS Preliminary Remediation Goals (PRGs) Calculator

Fields that are highlighted are required.

Select Scenario

- (•) Resident
- ( ) Indoor Worker
- ( ) Outdoor Worker
- ( ) Recreator
- ( ) Excavation Worker
- ( ) Farmer

Select PRG type

- (•) Defaults
- ( ) Site Specific

RAIS Soil Screening Levels Calculator

$$SSL(mg/kg) = C_w\left(\frac{mg}{L}\right) \times \left[K_d\left(\frac{L}{kg}\right) + \left(\frac{\theta_w\left(\frac{L_{water}}{L_{soil}}\right) + \theta_a\left(\frac{L_{air}}{L_{soil}}\right) \times H'}{\rho_b\left(\frac{1.5\ kg}{L}\right)}\right)\right]$$

where:

$$\theta_a\left(\frac{L_{air}}{L_{soil}}\right) = n\left(\frac{L_{water}}{L_{soil}}\right) - \theta_w\left(\frac{0.3\ L_{water}}{L_{soil}}\right);$$

$$n\left(\frac{L_{pore}}{L_{soil}}\right) = 1 - \left(\frac{\rho_b\left(\frac{1.5\ kg}{L}\right)}{\rho_s\left(\frac{2.65\ kg}{L}\right)}\right) \text{ and}$$

$$K_d\left(\frac{L}{kg}\right) = K_{oc}\left(\frac{L}{kg}\right) \times f_{oc}\ (0.002 \text{ unitless})$$

Inorganic Soil Background Selection

Results

All units are ppm except where noted

Back

Chemical	Soil-type	Range
Aluminum	Soils over granites and gneisses	7.2-8.2%
Aluminum	Sandy and lithosols on sandstones	2.5-4.3%
Aluminum	Various soils	0.45-10%
Aluminum	Soils over limestones and calcareous rocks	0.43-1.3%
Aluminum	Soils over Volcanic rocks (or ash*)	6.9-8.1%

Risk Assessment Steps: Exposure Assessment

RAIS Preliminary Remediation Goals (PRGs) Calculator

Fields that are highlighted are r

Select Scenario

- (•) Resident
- ( ) Indoor Worker
- ( ) Outdoor Worker
- ( ) Recreator
- ( ) Excavation Worker
- ( ) Farmer

Select PRG type

- (•) Defaults
- ( ) Site Specific

Soil Non-Carcinogenic Ingestion Equation

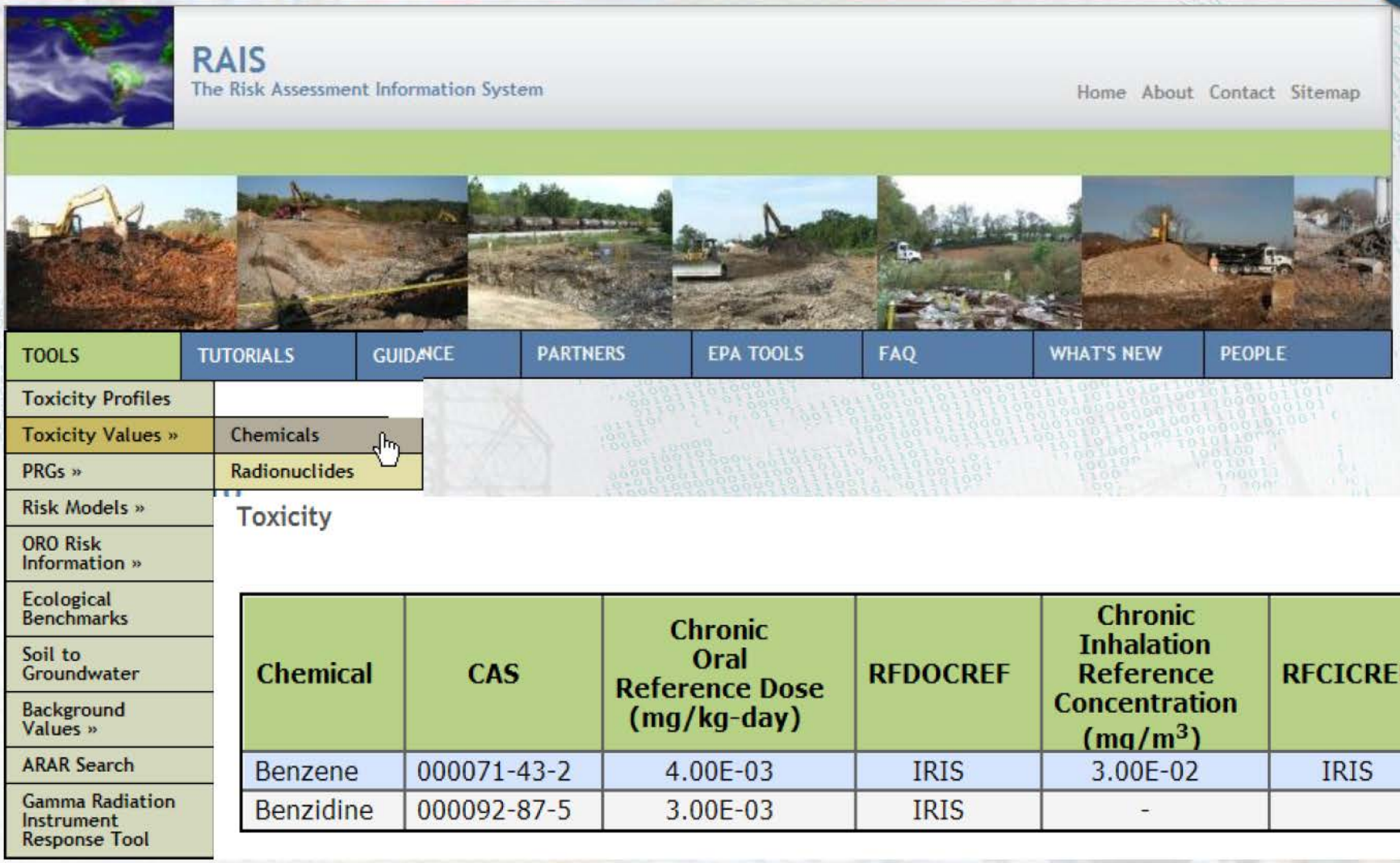
$$PRG_{res\text{-}soil\text{-}nc\text{-}ing}(mg/kg) = \frac{THQ \times AT_{ress}\left(\frac{365\ days}{year}\right) \times ED_{ress}(26\ years)}{\left(\frac{1}{RfD_o\left(\frac{mg}{kg\text{-}day}\right)}\right) \times IFS_{res\text{-}adj}\left(\frac{36{,}750\ mg}{kg}\right) \times \frac{10^{-6}\ kg}{1\ mg}}$$

where:

$$IFS_{res\text{-}adj}\left(\frac{36{,}750\ mg}{kg}\right) = \frac{ED_{ressc}(6\ years) \times EF_{ressc}\left(\frac{350\ days}{year}\right) \times IRS_{ressc}\left(\frac{200\ mg}{day}\right)}{BW_{ressc}(15\ kg)} + \frac{\left(ED_{ress}(26\ years) - ED_{ressc}(6\ years)\right) \times EF_{ressa}\left(\frac{350\ days}{year}\right) \times IRS_{ressa}\left(\frac{100\ mg}{day}\right)}{BW_{ressa}(80\ kg)}$$

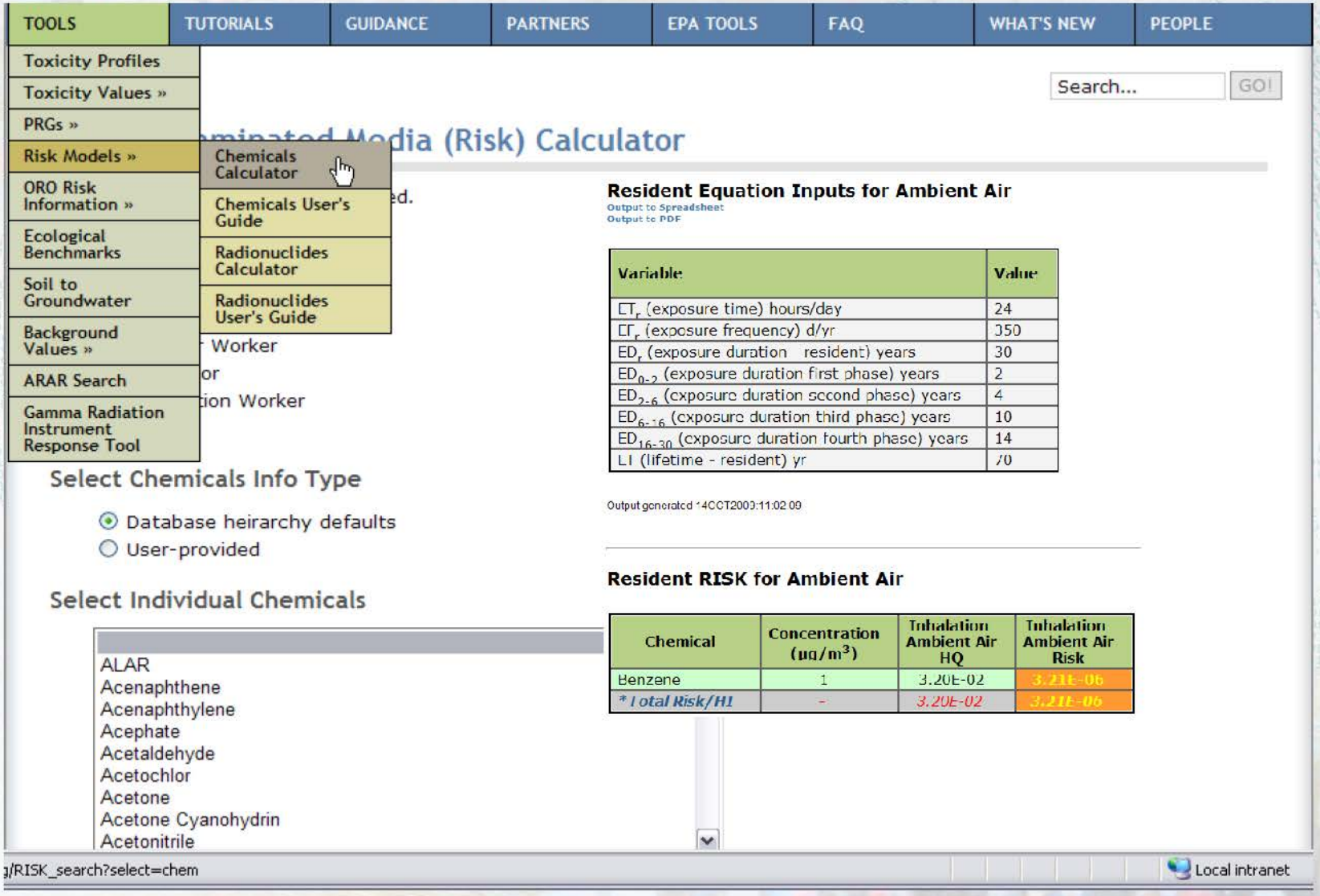
Age Segment (yr)	AF (mg/cm^2)	BW (kg)	ED (yr)	EF (day/yr)	ET (hr/event)	IRS (mg/day)	SA (cm^2/day)
0-2	0.2	15	2	350	24	200	2690
2-6	0.2	15	4	350	24	200	2690
6-16	0.07	80	10	350	24	100	6032
16-30	0.07	80	10	350	24	100	6032
Child (0-6)	0.2	15	6	350	24	200	2690
Adult (6-70)	0.07	80	20	350	24	100	6032

Risk Assessment Steps: Toxicity Assessment



Chemical	CAS	Chronic Oral Reference Dose (mg/kg-day)	RFDOCREF	Chronic Inhalation Reference Concentration (mg/m^3)	RFCICREF
Benzene	000071-43-2	4.00E-03	IRIS	3.00E-02	IRIS
Benzidine	000092-87-5	3.00E-03	IRIS	-	

Risk Assessment Steps: Risk Characterization



RAIS Working with SADA

- **RAIS Toxicity database shared with SADA**
- **RAIS Chemical parameters database shared with SADA**
- **RAIS (and EPA) Risk and PRG models shared with SADA.**

Spatial Analysis and Decision Assistance SADA

Robert Stewart

Fred Dolislager

Tom Purucker

The Institute for Environmental Modeling, University of Tennessee

Geographic Information Science and Technology Group, Oak Ridge National Laboratory

Spatial Analysis & Decision Assistance

- SADA project engages research and development at the nexus of geospatial analytics, risk assessment, and decision analysis.
- Goals are to embed risk assessment, uncertainty modeling, and downstream decision processes entirely within a spatial context
- Two lanes define project activities
  - Advancing methods in a variety of areas particularly well connected to environmental regulatory community, characterization, remediation, RCRA, Superfund, MARSSIM, etc.
  - Freeware desktop application (SADA) integrating environmental risk analytics, spatial modeling, and decision sciences

Questions That SADA Addresses

What exposure scenarios are likely dangerous?

What contaminants are driving the risk?

What pathways (ingestion, inhalation, etc)?

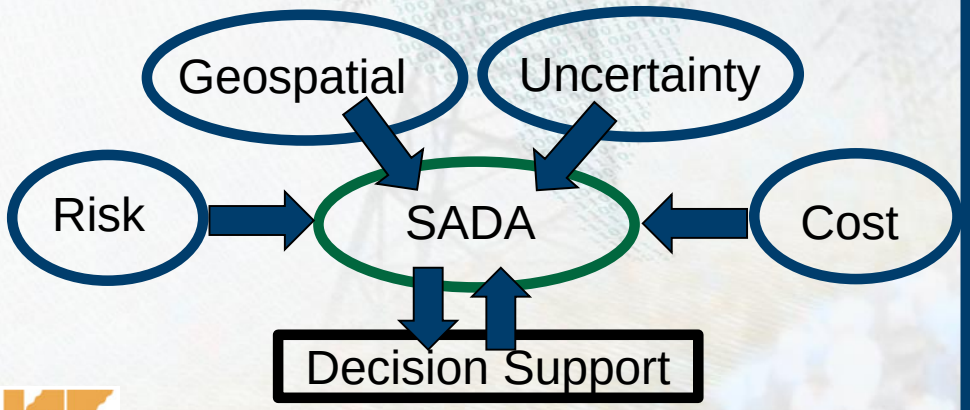
What is the risk or concentration limit for an exposure time of 30 years?, 1 day? 1 hour?

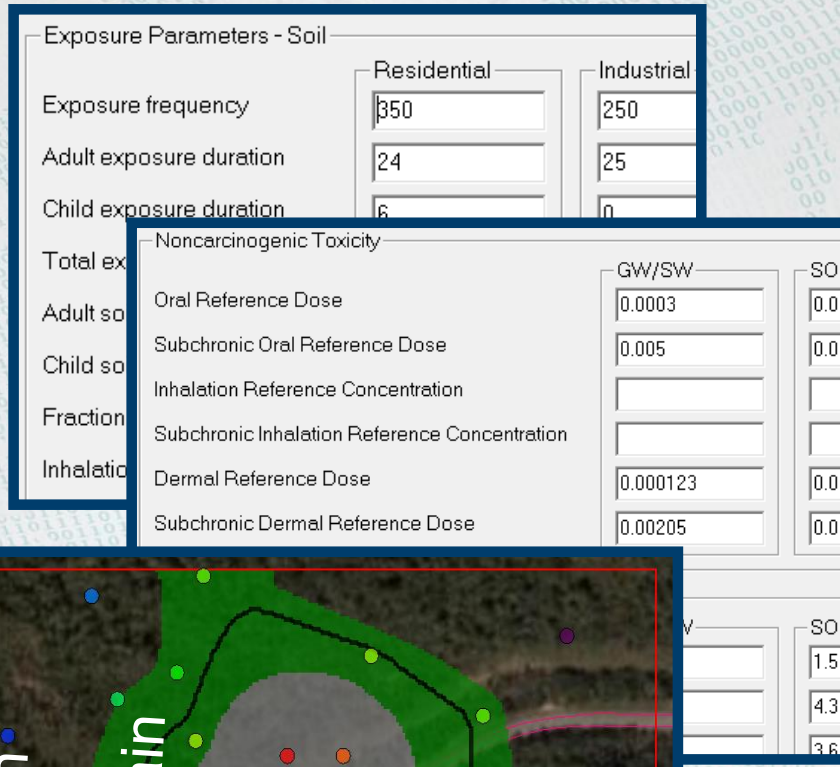
Where is exposure unsafe? Who might be in harms way? How sure are we?

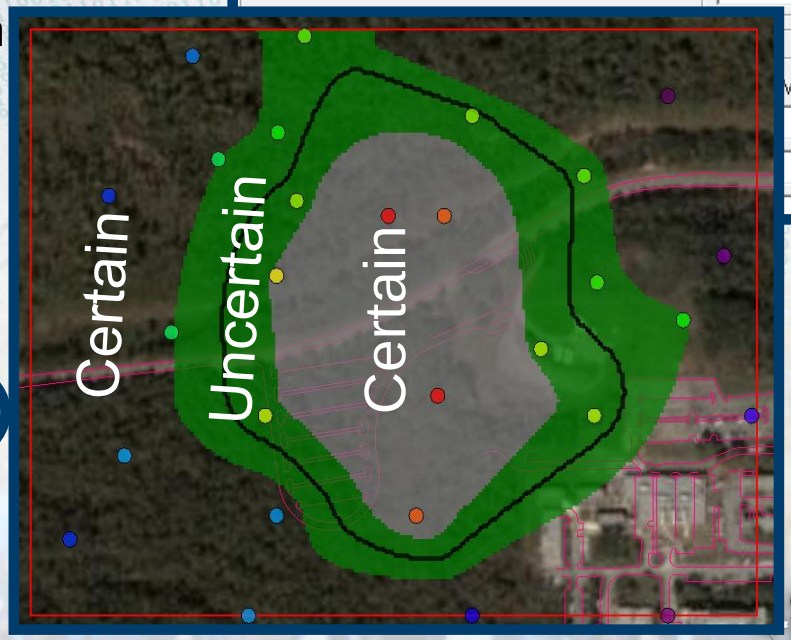
Where should we apply risk mitigation measures?

Where and what type of additional information would support the model?

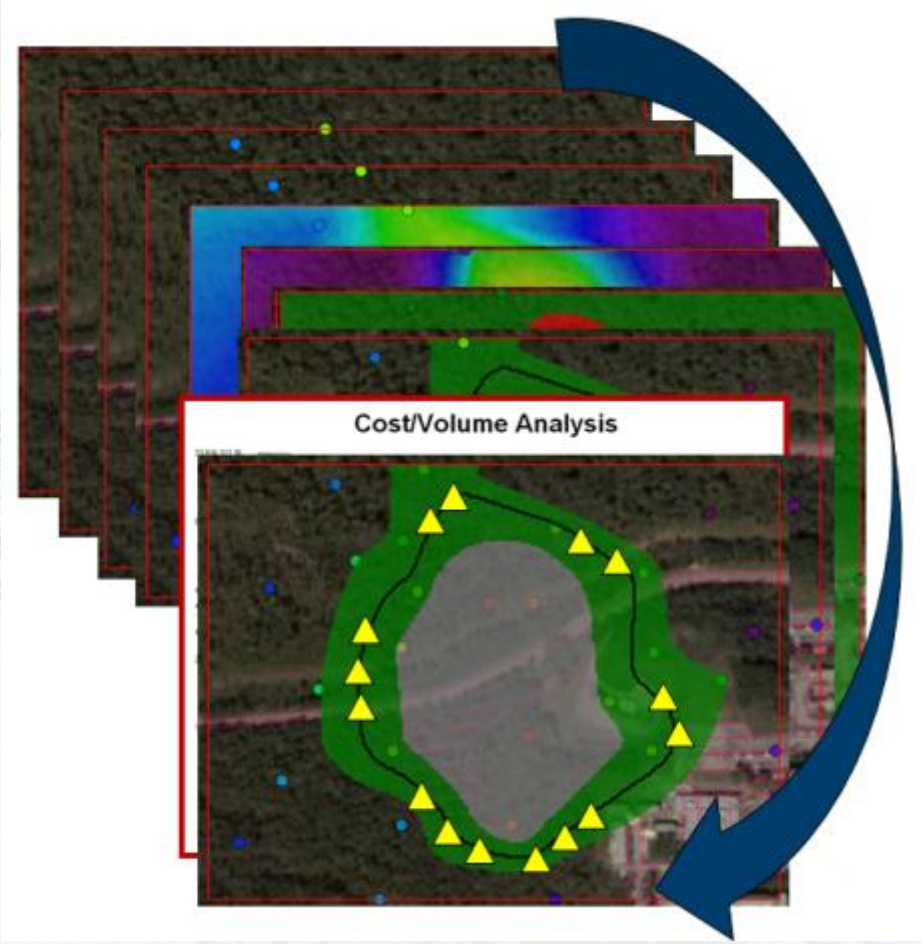
What are our decision risks?







Answers that SADA V5 Provides



Initial Sample Designs

Judgmental

Random

Simple Grid

Standard Grid

Unaligned Grids

Search Grids

MARSSIM

3D Search

Cost Benefit Analytics

Built on risk-space models

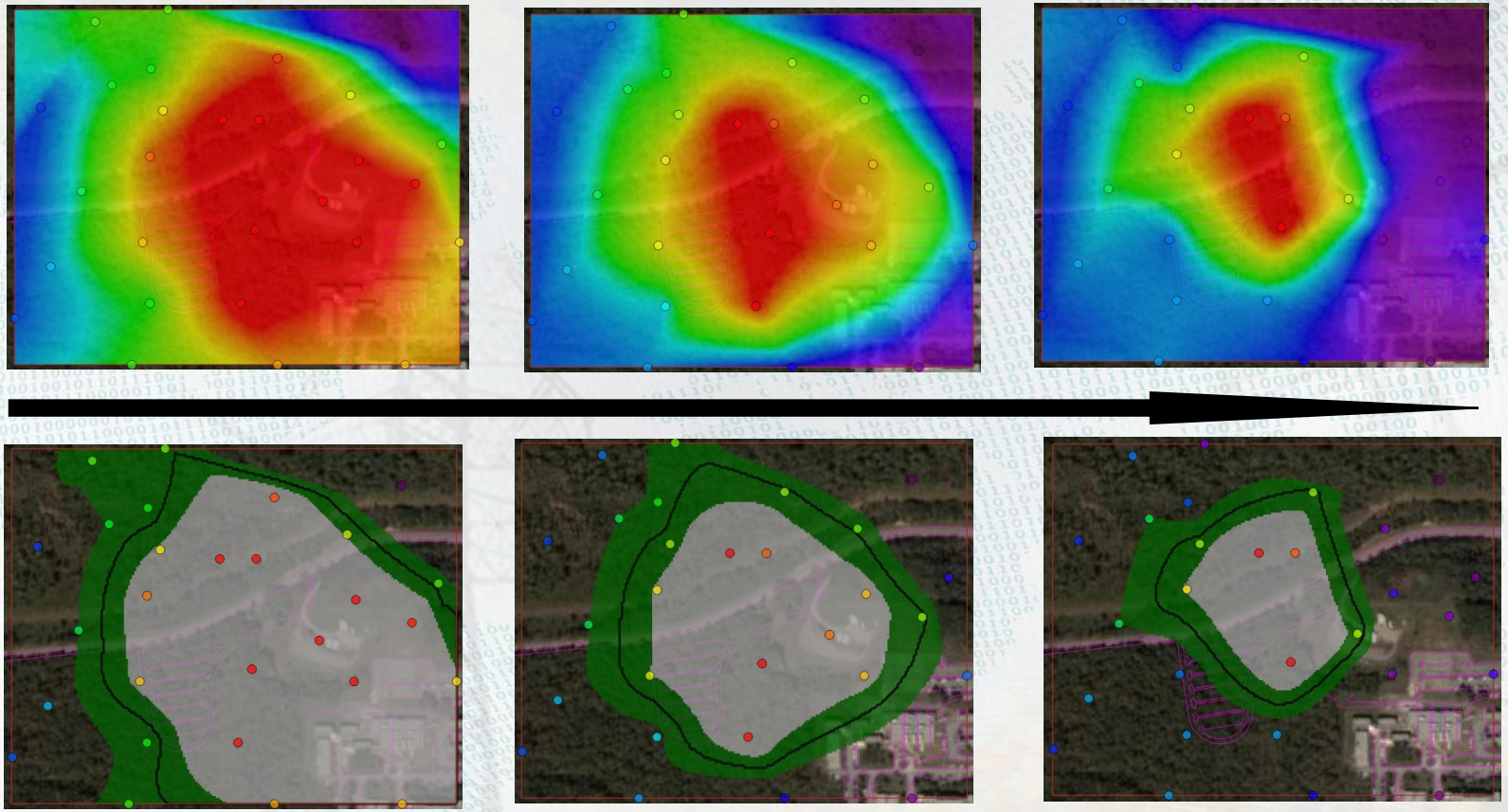
Permitwhat if's

Quantify cost and *decision* risk reduction

Secondary Sample Designs

Sample where model needs most support....

Risk Based Decisions Over Time



Day 1 Day 2 Day 3

But that's Not All

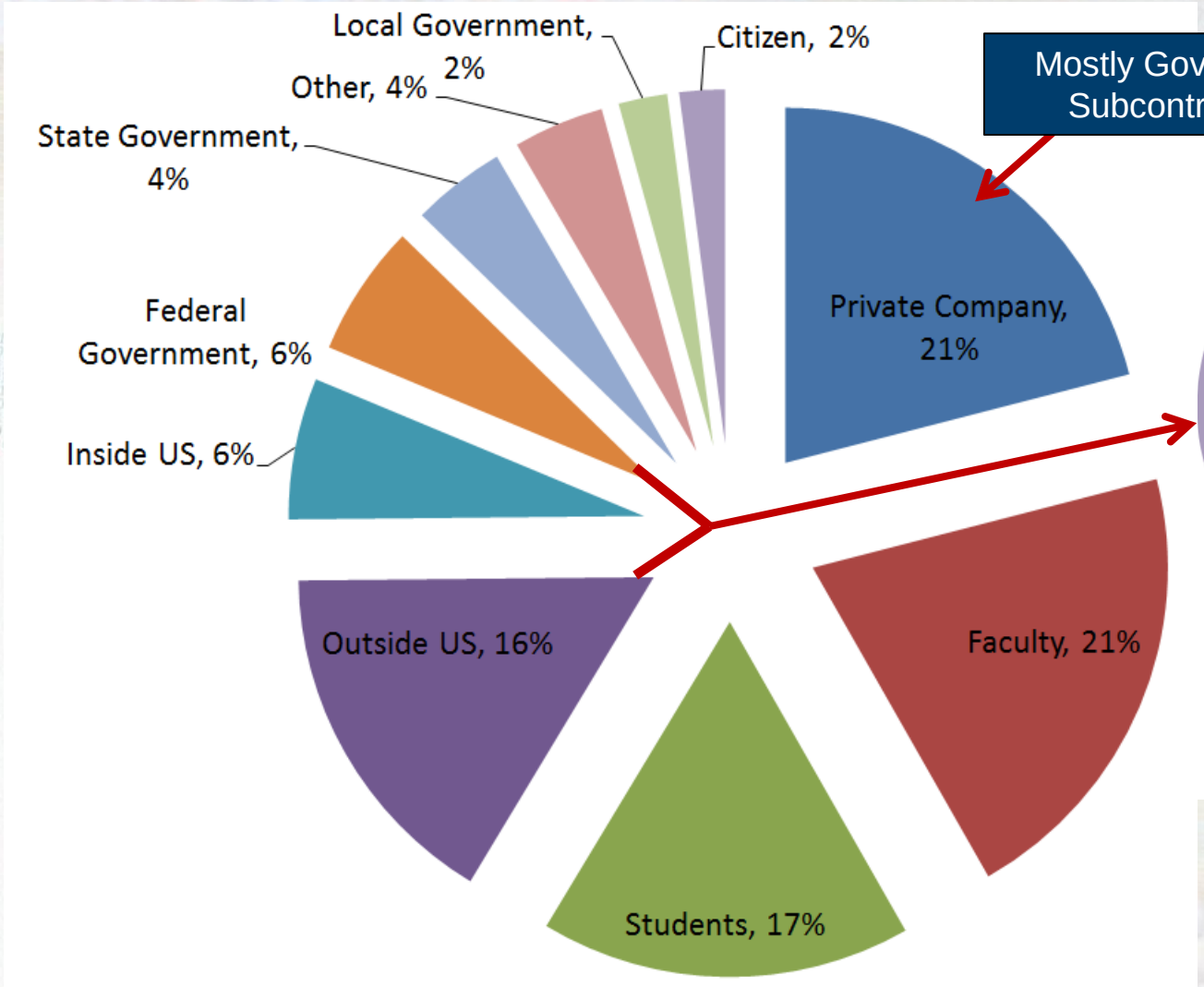
- Because SADA is an open spatiotemporal modeling environment, it can be used for numerous applications outside of toxicologial and radiological risk.
- Examples include, engineering, geophysical, geological, ecosystems monitoring, epidemiological

Where

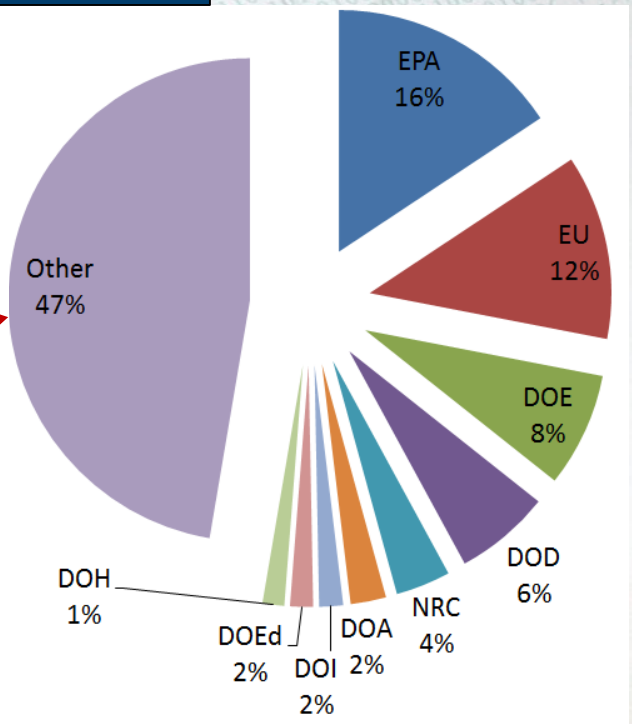
- In the 17th year of deployment (began ~1998)
- 18,000+ registered users
- 90+ scientific and regulatory communications (e.g. journal articles, reports, web pages, theses, etc.
- User group, workshops, conferences, international presence etc.

www.sadaproject.net

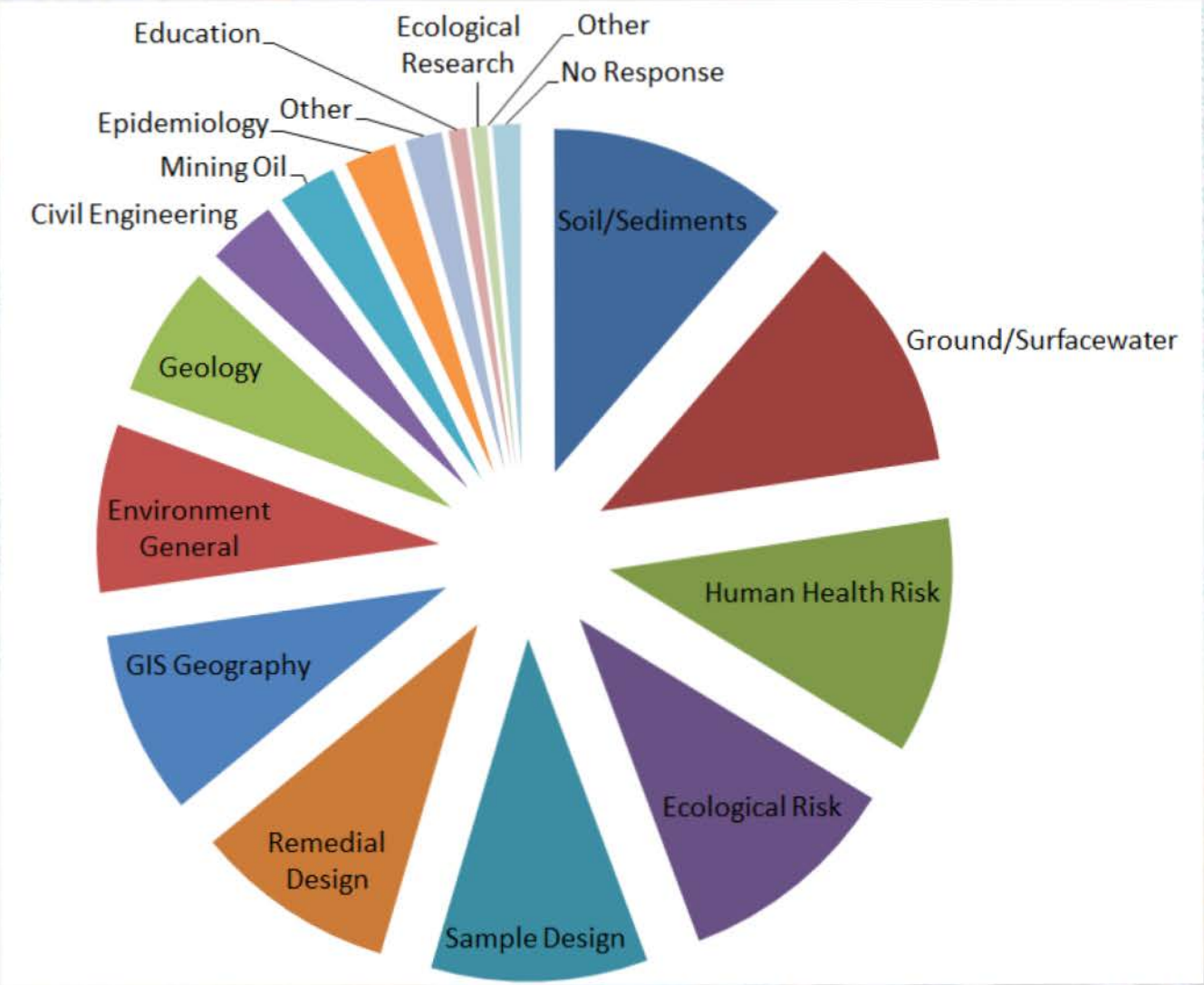
Who



Mostly Government Subcontractors



What


Education
Ecological Research
Other
No Response
Other
Epidemiology
Mining Oil
Civil Engineering
Soil/Sediments
Ground/Surfacewater
Geology
Environment General
Human Health Risk
GIS Geography
Ecological Risk
Remedial Design
Sample Design

SADA Contributors/Sponsors/Advisors

United States Nuclear Regulatory Commission
United States Environmental Protection Agency
United States Department of Energy
Environmental Measurements Laboratory
Pacific Northwest National Laboratory
Oak Ridge National Laboratory
Michigan State University
San Diego State Univeristy
Standford (GSLIB codes)
Swedish Geotechnical Institute
Numerous subcontractors and Private Sector Advising

Dr. Louis Gross (NIMBios/University of Tennessee)
Dr. Tom Purucker (USEPA ORD)
Dr. Christopher Welsh (NimBios/University of Tennessee)
Fred Dolislager (TIEM/UT)
Debra Stewart (TIEM/UT)
John Joseph Roberts-Niemann (UT/Lockheed Martin)
Dr. Pierre Goovaerts (U of Michigan/Biomedware/PGeostats)
Dr. George Powers(Regulatory Research USNRC)
Dr. Carl Gogolak(EML/Consultant)
Wilson McGinn (ORNL)
Katie Tucker (Ingenium)
Teresa Perry (DOE)

Michelle Burgess (USEPA, Superfund)
John Bing-Canar (USEPA, Region 5)
Brian Cooper(USEPA, Region 5)
Tim Drexler (USEPA, Region 5)
Dr. Brenda Jones(USEPA, ERAF)
Leslie Galloway(TIEM/University of Tennessee)
Dr. Bruce Ralston(Geography/UT)
John Wilson(PNNL)
Rich Pawlowitz (Bechtel-Jacobs)
Dr. George Christakos(San Diego State University)
Dr. Hwa-Lung Yu (San Diego State University)
Jim Davidson(PNNL/Davidson & Davidson)
Melinda Vzanna (student/University of Tennessee)
Rob Fletcher (student/University of Tennessee)
Peter Starzec(Swedish Geotechnical Institute)
Jerry Montgomery(US Army Corp Of Engineers)
Jenny Norman(Swedish Geotechnical Institute)
Tom Nicholson(USNRC)
Bobby Abueid(USNRC)
Jim Wulf(Tetratech)
Carl Stineman(Consultant)
Martin Bittens(University of Tuebingen, DE)

Project Manager: Dr. Robert Stewart, stewartrn@ornl.gov

Lessons - Doing the Modeling:

Work closely with those with long experience in the system being modeled.

Moderate the above based first on the availability of data to construct reasonable models, and secondly on the difficulty of constructing and calibrating the models.

Don't try to do it all at once - start small - but have a long-term plan for what you wish to include overall, given time and funding.

Lessons - Doing the Modeling:

Leave room for multiple approaches: don't limit your options.

In the face of limited or inappropriate data, use this as an opportunity to encourage further empirical investigations of key components of the system.

Build flexibility in as much as possible.

Be flexible about what counts as success.

Lessons - Personnel Matters:

Build a quality team who respect each others abilities and won't second guess each other, but who accept criticism in a collegial manner.

Keep some part of the team out of the day-to-day political fray.

Be persistent, and have at least one member of the team who is totally dedicated to the project and willing to stake their future on it.

Do whatever you can to maintain continuity in the source of long-term support for the project.

Lessons - Interacting with Stakeholders:

Constantly communicate with stakeholders.

Regularly explain the objectives of your modeling effort, as well as the limitations, to stakeholders. Be prepared to do this over and over for the same people, and do not get frustrated when they forget what you are doing and why.

Be prepared to regularly defend the scientific validity of your approach.

Lessons - Interacting with Stakeholders:

Don't limit your approach because one stakeholder/funding agency wants you to.

Be prepared for criticism based upon non-scientific criteria, including personal attacks.

Ignore any of the stakeholders at your peril.