- **Othar Hansson**

# The CICT Earth Science Systems Analysis Model

COMPUTING
INFORMATION &
COMMUNICATIONS
TECHNOLOGY

Barney Pell, Joe Coughlan,
Bryan Biegel, Ken Stevens,
Othar Hansson, Jordan Hayes

NASA Ames Research Center
& Thinkbank, Inc.
April 2004

## The ESSA Team

- Task leads:
  Barney Pell (Lead), Bryan Biegel (Co-lead),
  Joe Coughlan (Science Lead),
  Walt Brooks (Science Co-Lead)
- Subcontractor:
  Othar Hansson & Jordan Hayes, Thinkbank
- ARC team:
  Ken Stevens, Peter Cheeseman, Chris Henze,
  Samson Cheung, et al.

## Enough About Me

- Research collaborations with NASA Ames since 1989 (heuristic search, data-mining, planning/scheduling).
- PhD (Computer Science), Berkeley.
  Using decision analysis techniques for search control decisions in science planning/scheduling systems.
- Thinkbank:
  custom software development,
  software architecture consulting,
  technology due-diligence for investors.

## Agenda

CICT Systems Analysis

Our modeling approach

- a 3-part schematic investment model of technology change, impact assessment and prioritization

A whirlwind tour of our model

Lessons learned

## Systems Analysis in CICT

- Demonstrate "systematic and thorough investment decision process" to HQ, OMB and Congressional Decision Makers
- Increase awareness and substantiate CICT's impact to missions. Road map CICT projects to missions and measurement systems
- 4 teams in FY03:
  - 2 pilot studies (Earth Science [me]; Space Science [Weisbin]): explore models for ROI of IT.
  - TEAM: map from NASA Strategic Plan to IT capability requirement; technology impact assessment
  - Systems Analysis Tools (COTS/GOTS)

## Earth Science Pilot Study

How do we characterize and quantify a science process?

Can we build a model of how CICT technology investments impact ROI in a NASA science process?

What modeling approach is suitable for making such analyses understandable and repeatable?

## Current State

What have we learned? (FY03)

- Decision analysis modeling techniques can be applied to systems analysis of CICT project areas. Built model of weather-prediction data pipeline.

What don't we know? (FY04)

- How much time/expense needed to build a full model
- How such a full model fits into a real NASA program context (CDS: Collaborative Decision Systems)

## Pilot Study Focus

- Criteria for science process to study
  - Important to a major customer base,
  - Significantly drives technology investments
  - Generalizes to a class of related processes
  - Amenable to quantitative analysis.
- 2010 Weather Prediction process
  - Critical Earth Science process with relevance not only to NASA scientists but to the nation at large.
  - Stretch goals require technology breakthroughs.
  - Strong technology driver for other science problems
  - Starting point: analyses from ESE computational technology requirements workshop (4/02)

## Pilot Study Accomplishments

- Identified modeling formalism (influence diagrams)
  - Clear semantics accessible to both ES & CICT experts
  - Tools exist for sensitivity analysis, decision-making, etc.
    We chose Analytica as our modeling tool.
  - Successfully transferred/applied to Space Science pilot study as well.
- Built a model with an understandable, simple structure (after much research and many iterations).
- Demonstrated the kinds of analyses made possible by the model

## Agenda

CICT Systems Analysis

Our modeling approach

- a 3-part schematic investment model of technology change, impact assessment and prioritization

A whirlwind tour of our model

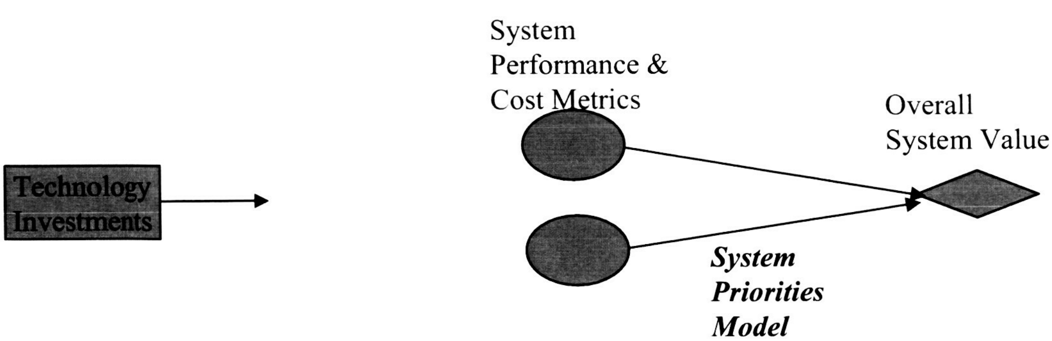
Lessons learned

## Methodology: Decision Model


Q1: Which technology investments should I make?

Q2: How does each technology investment improve overall system/mission value (including cost considerations)? Choose investments with highest value.

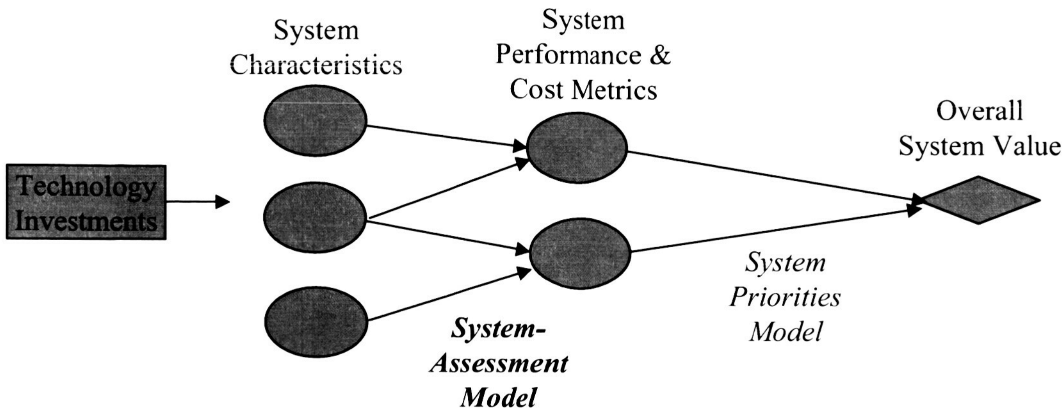
## Filling in the Decision Model




System value is a function of a set of metrics (accuracy, fidelity, cost, etc.). We can model the priority among the metrics independent of the technologies used.

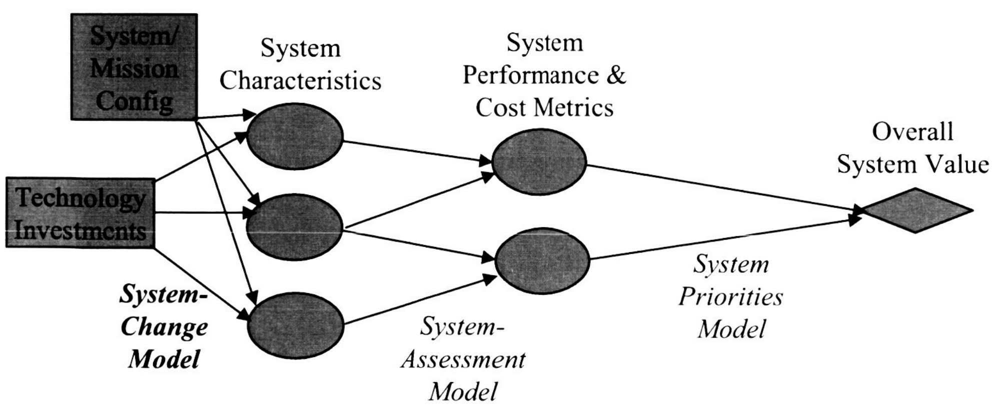
Technology investments have value in that they improve these metrics.

## Filling in the Decision Model



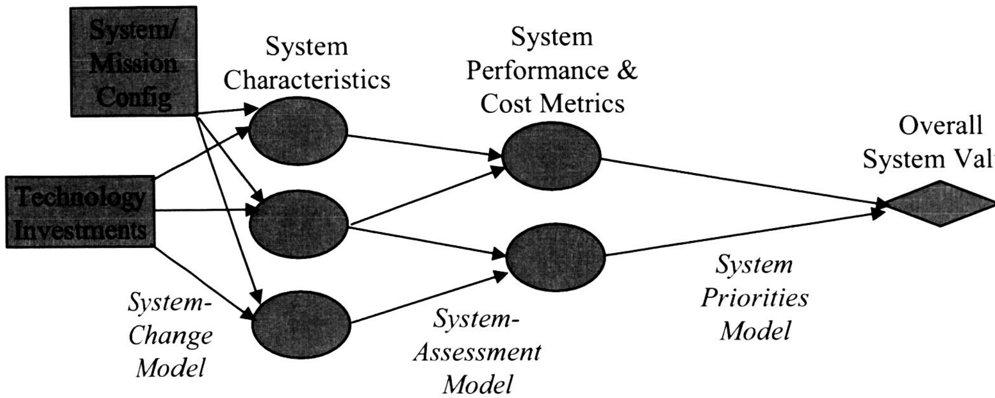

The metrics can be modeled in terms of abstract system characteristics (data volume, algorithm accuracy, processing speed, model fidelity, ...).

## Filling in the Decision Model




Technology investments, together with some mission-specific parameters, influence the system characteristics. A technology investment (such as data visualization research) has value in that it improves system characteristics (such as model fidelity).

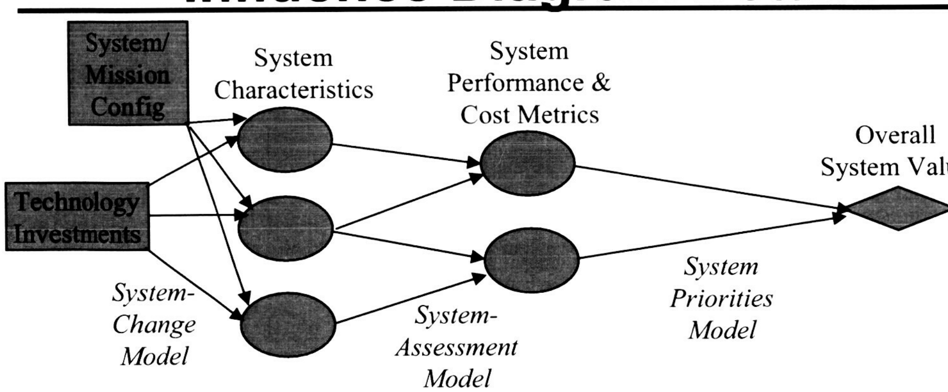
## Methodology: Influence Diagrams




We've sketched an "influence diagram" model of the decision.

Q: What tech. investments maximize expected overall system value?
Q: Value of model refinement: How sensitive to assumption A?
Q: Value of information: what if we knew that project P would succeed?
Q: Value of control: what if we could reduce risk of project P failing?

## Influence Diagram Details




Influence diagram tools (such as Analytica) allow you to specify and evaluate these models. Diagram structure and decision analysis techniques speed specification of required parameters.

"What-if" and optimization questions reduce to the problem of computing functions of conditional prob. distributions:
"best" technology investment is:

**argmax [E(Overall System Value | Technology Investments)]**

## Agenda

CICT Systems Analysis

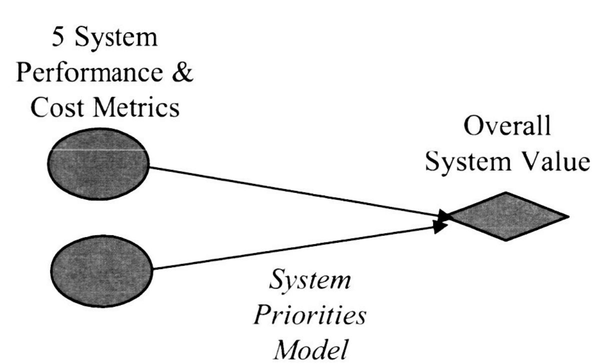
Our modeling approach

- a 3-part schematic investment model of technology change, impact assessment and prioritization

A whirlwind tour of our model

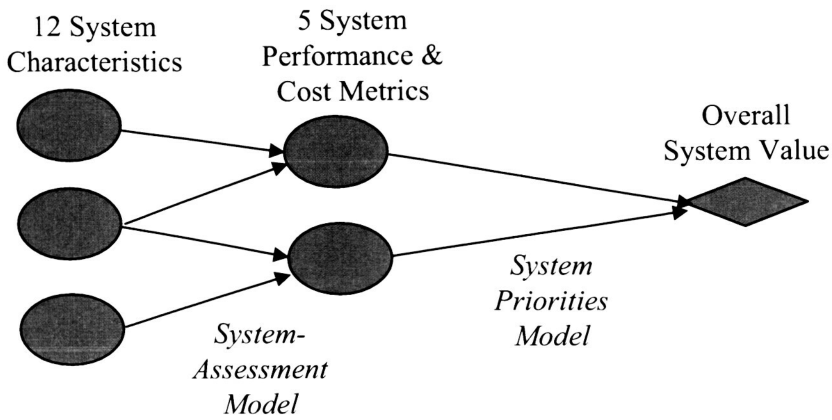
Lessons learned

## The ESSA Model



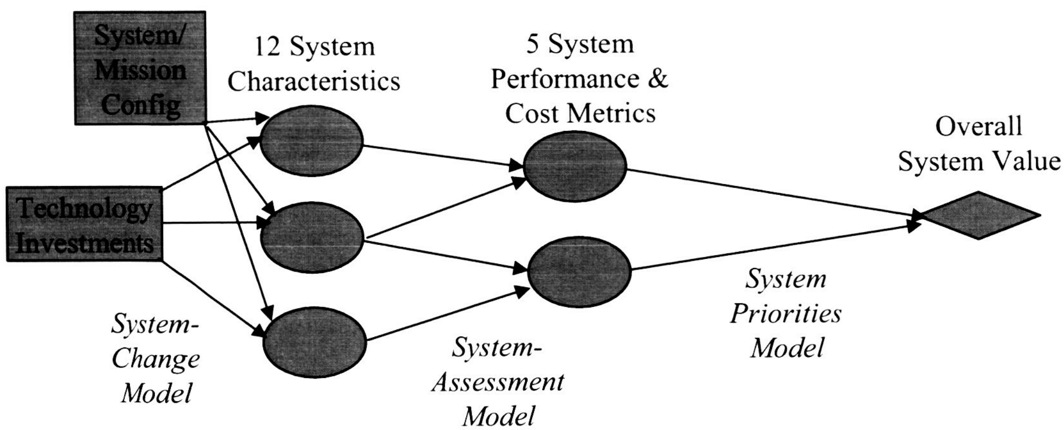

Our set of 5 metrics include:
development cost, operations cost, accuracy, model fidelity, etc.

# The ESSA Model




Our 12 System Characteristics include:
observation density, assimilation efficiency, cpu efficiency, etc.

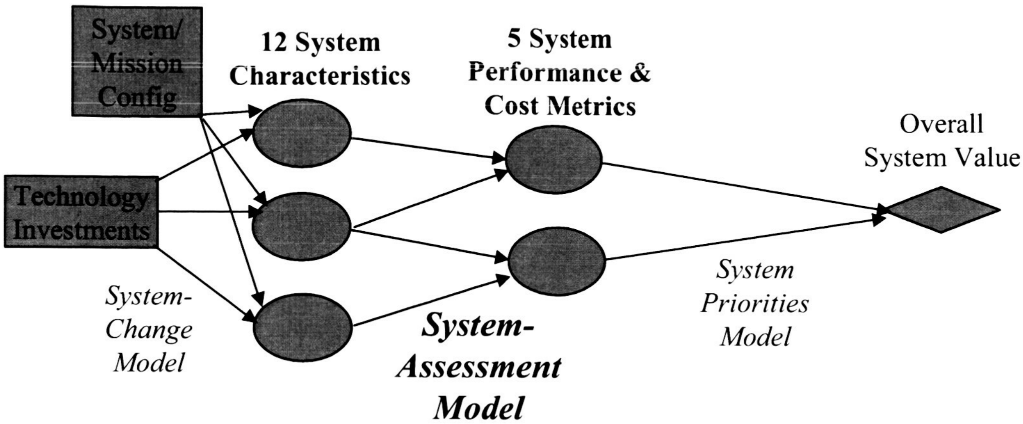
# The ESSA Model




Our 13 technology investments include: data-mining, launching a new data source, targeted observing, etc.

Each represents a research area, summarizing a range of individual research tasks or proposals.

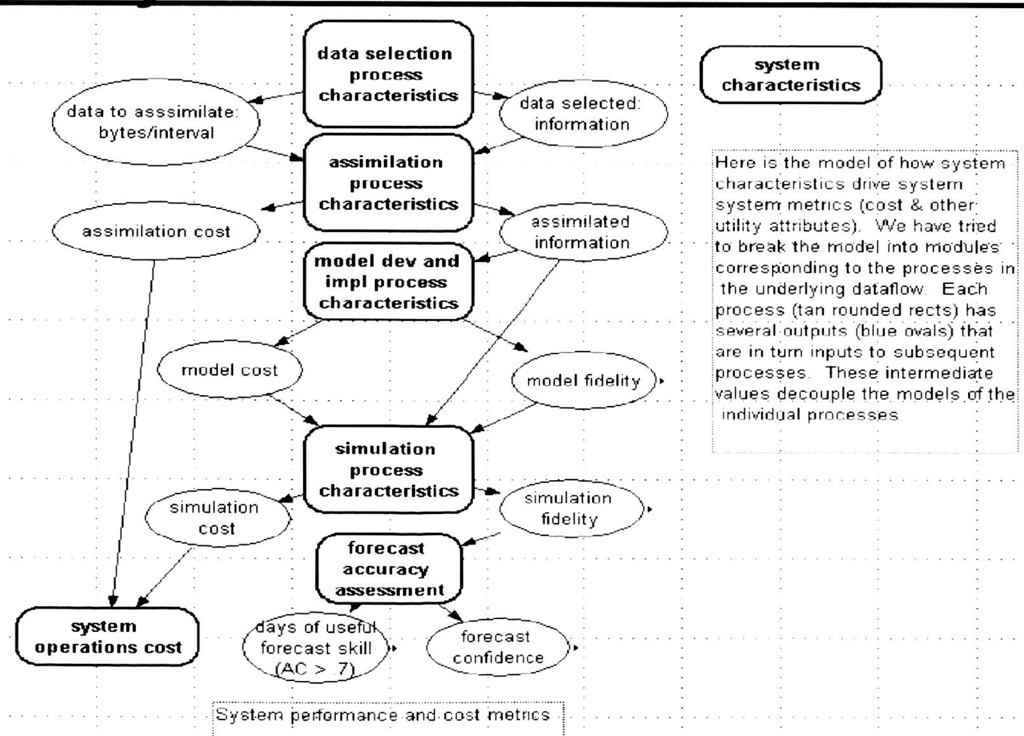
# Diving Down into the Model




System-Assessment Model: the most stable part of the model, owned/designed by a customer domain expert who understands the behavior of the system/mission being analyzed.

System-Assessment model computes System Metrics from System Characteristics

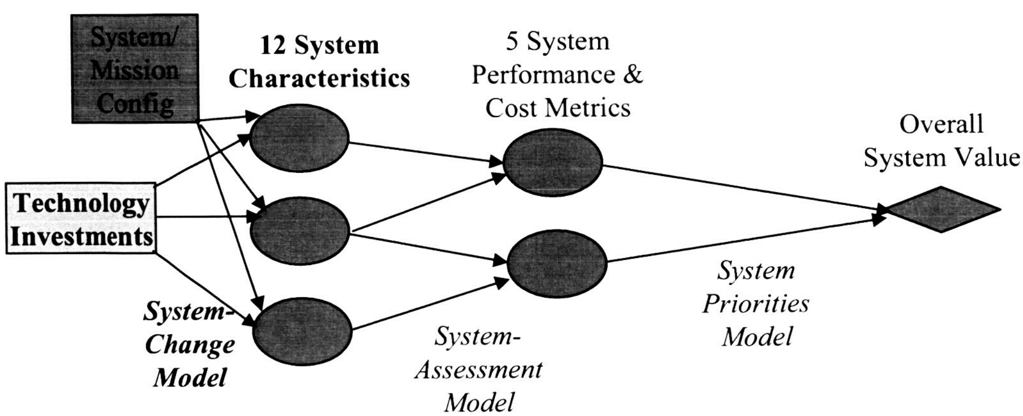
# System-Assessment Model



# Example System Characteristics

| | |
|---|---|
| Assimilation efficiency | 0-1 scale: how much information is retained despite approximations in data assimilation? |
| CPU efficiency | >0 : percentage speedup in CPUs due to R&D investments |
| Data efficiency | 0-1 scale: how much information is present in each bit of data selected? |
| Ensemble efficiency | 0-1 scale: how much improvement in forecast skill do we get from using ensemble algorithms? |
| Model framework | 0-1 scale: how much fidelity is present in our models? |
| Observation density | 0-1 scale: how many of the available observations do we make? |
| Postprocessing effectiveness | 0-1 scale: how much improvement in forecast skill do we get from using post-processing? |
| Simulation efficiency | > 0: percentage speedups in simulation due to R&D investments |

# Instantiating the Model




System-Change Model: owned/designed by a program manager who understands the feasibility and impact of different research areas.

System-Change model computes System Characteristics from the set of Technology Investments chosen (and system/mission config parameters)

## System-Change Model

- "Impact matrix" quantifies the changes to system characteristics that will occur if individual research projects succeed.
- "Cost matrix" quantifies cost breakdown for each research area.
- Portfolio of research areas determines what impacts will be felt.
- (In an extended model, cost and impact could vary over time.)

## System-Change: Research Areas

- **Data-efficient simulations (same data size)**
  choose a more informative set of observations to improve forecast skill at the same computational cost
- **Data-efficient simulations (less data)**
  reduce number of observations (and reduce computational cost) w/o reducing forecast skill
- **Targeted Observing**
  ditto, but also gather more targeted observations based on ensemble accuracy estimates (e.g., the SensorWeb concept)
- **Adaptive grid methods**
  reduce number of grid points by using regional forecast as boundary conditions
- **Improvements in ensemble methods**
  reduce number of ensembles needed to get similar accuracy estimates (e.g., through use of particle filter technology)
- **Data-mining of model outputs**
  increased skill from same model output via data analysis & visualization (intelligent data understanding)

## System-Change: Research Areas

- **Modeling tools**
  ESMF and other initiatives to make modeling efforts more productive
- **System Management/Tuning tools**
  Auto or Semi-Automatic Parallelization tools, Benchmarking, Cluster management, etc.
- **Instrument models**
  tools for creating more accurate instrument models.
- **Launch new data source**
  collect additional types of observation data by launching a new instrument.
- **Launch replacement data source**
  collect a new type of observation data, but keep the total amount of data processed the same.
- **Higher resolution models**
  develop higher resolution models and move to higher resolution simulation

## Research Area Impact

Impact matrix has a value for each pair (13 research areas x 12 system characteristics): 156 possible, but only 18 are nonzero.

Impact can be positive or negative:

Impact(targeted observing, observation density) = low neg.

Impact(launch new data source, observation density) = low

Some more examples:

Impact(targeted observing, targeting efficiency) = low

Impact(system mgmt/tuning, cpu efficiency) = low

Impact(adaptive grid, simulation efficiency) = medium

# Impact Matrix

| | Assimilation efficiency | Assimilation density | Cpu efficiency | Data efficiency | Downlink density | Ensemble efficiency | Model framework | Observation density | Observation efficiency | Postprocessing effectiveness | Simulation efficiency | Targeting efficiency |
|---|---|---|---|---|---|---|---|---|---|---|---|---|
| data-efficient simulations (same data size) | | | | | | | | | | | | |
| data-efficient simulations (less data) | | | | | | | | (lo) | | | | |
| targeted observing | | | | | | | | (lo) | | | | lo |
| adaptive grid methods | | | | | | | | | | | | |
| improved ensemble methods | | | | | | | | | | | | |
| data-mining of model outputs | | | | | | | | | | | | |
| modeling tools | | | | | | | | | | | | |
| system mgmt/tuning | | | lo | | | | | | | | | |
| launch new data source | | | | | | | | lo | | | | |
| launch replacement data source | | | | lo | | | | | | | | |
| instrument models | lo | | | | | | | | | | | |
| higher resolution models | | lo | | | | | lo | | | | (lo) | |

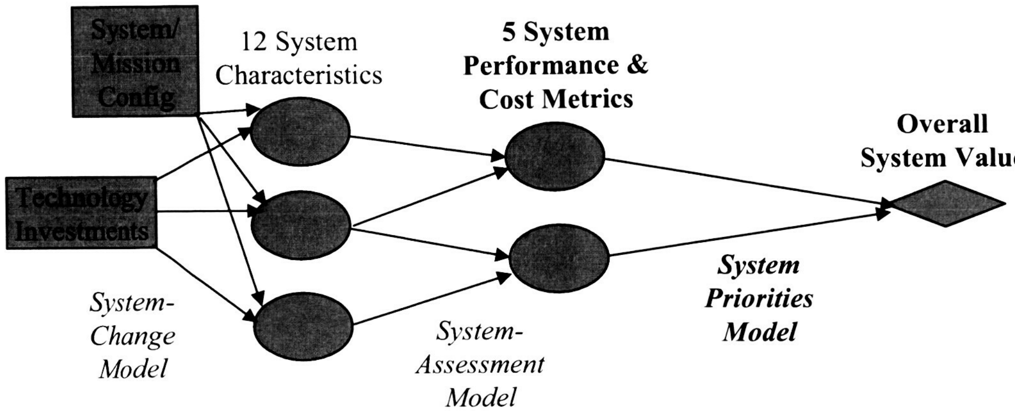
# Qualitative → Quantitative

Impact is parameterized qualitatively (lo, med, hi). This qualitative scale is then quantified inside the model.

Each of the parameters has a different interpretation under the four scenarios (pessimistic, consensus, optimistic, ideal). This allows us to compare in a best-case vs. worst-case manner.

| | pess. | cons. | optim. | ideal |
|---|---|---|---|---|
| Lo | .05 | .1 | .15 | 1.0 |
| Med | .2 | .3 | .4 | 1.0 |
| Hi | .3 | .5 | .7 | 1.0 |

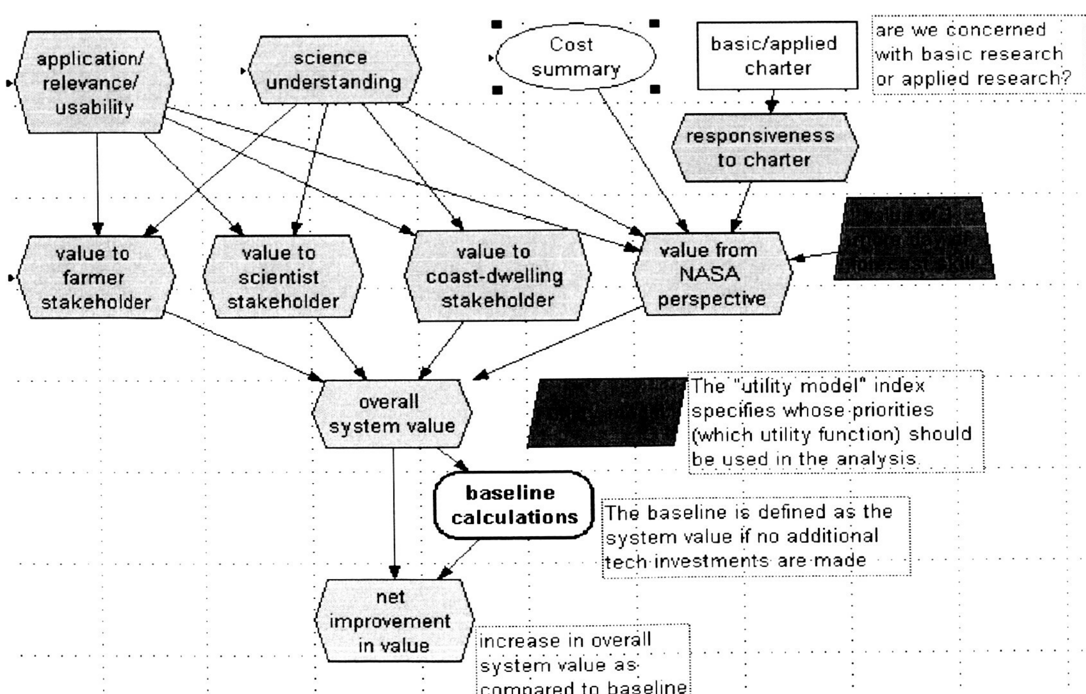
# Instantiating the Model



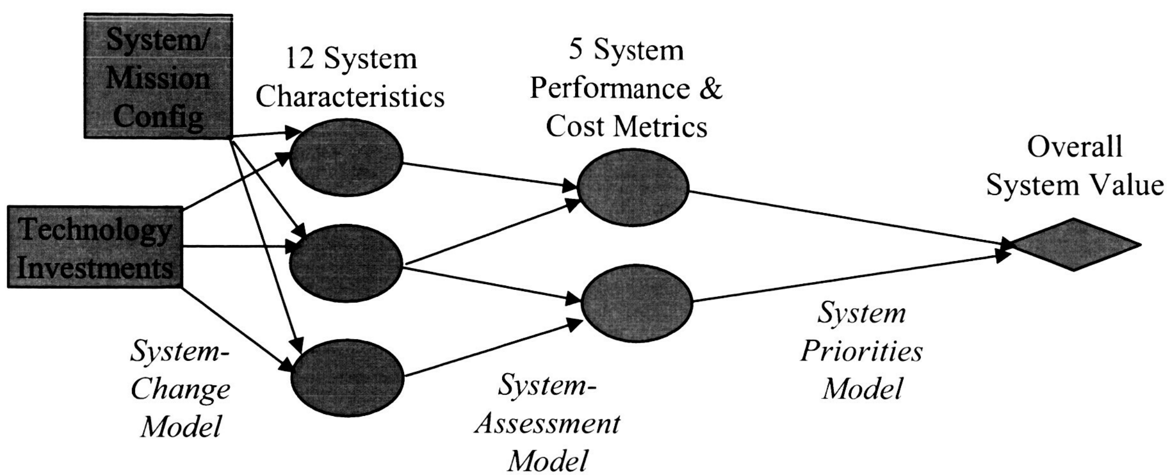

System Priorities Model: designed/owned by program manager cognizant of NASA priorities

System Priorities Model computes overall System Value given the System Metrics.

# System Priorities Model



## Review: Combining the Models




## Results: Caveat

Remember: results (evaluations, ROI, etc.) must be understood as a function of the inputs used to calculate the results:

$f$(model, assumptions, priorities)

Priorities depend on perspective: we model basic (science value only) versus applied (economic value only)

# Evaluating Research Areas

Result - net improvement in value

Mid Value of net improvement in value

utility model: NASA

optimism over research outcomes: consensus

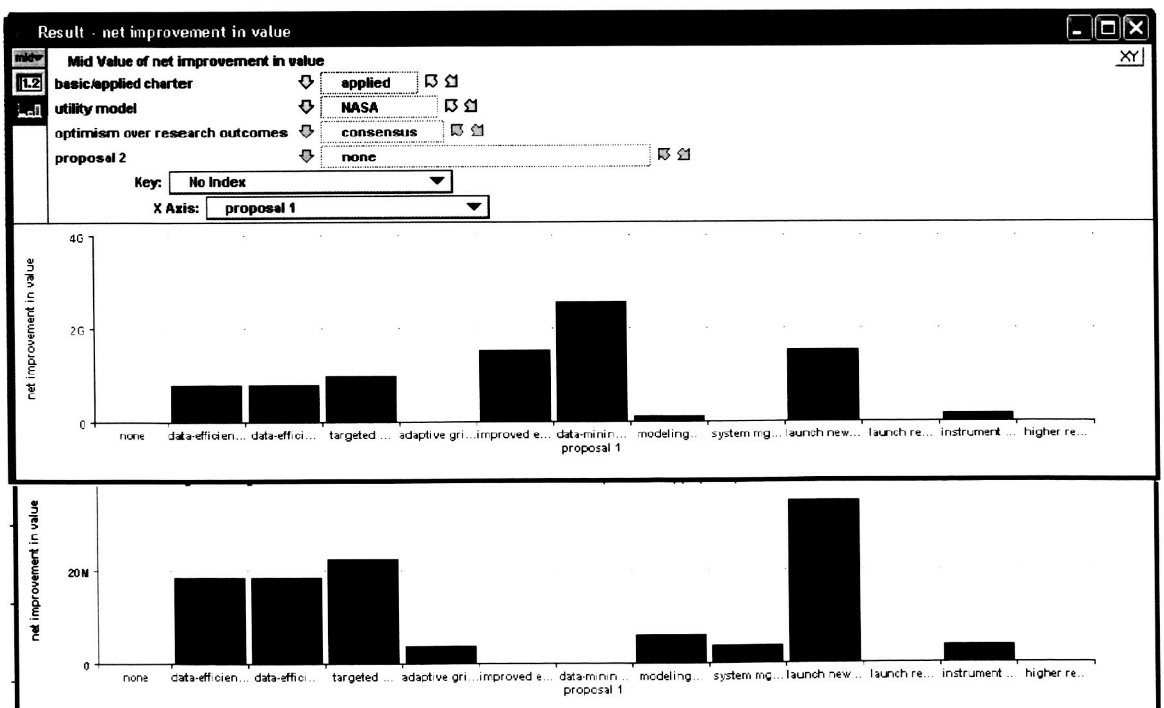
proposal 2: none

proposal 1 / basic/applied charter

| | basic | applied |
|---|---|---|
| none | 0 | 0 |
| data-efficient simulations (same data size) | 18.54M | 806.9M |
| data-efficient simulations (less data) | 18.54M | 806.9M |
| targeted observing | 22.56M | 982.6M |
| adaptive grid methods | 3.577M | 3.577M |
| improved ensemble methods | -10 | 1.536G |
| data-mining of model outputs | -10 | 2.56G |
| modeling tools | 6.098M | 108.1M |
| system mgmt/tuning | 3.577M | 3.577M |
| launch new data source | 35.01M | 1.528G |
| launch replacement data source | -20 | -20 |
| instrument models | 3.67M | 159.2M |
| higher resolution models | 0 | 0 |

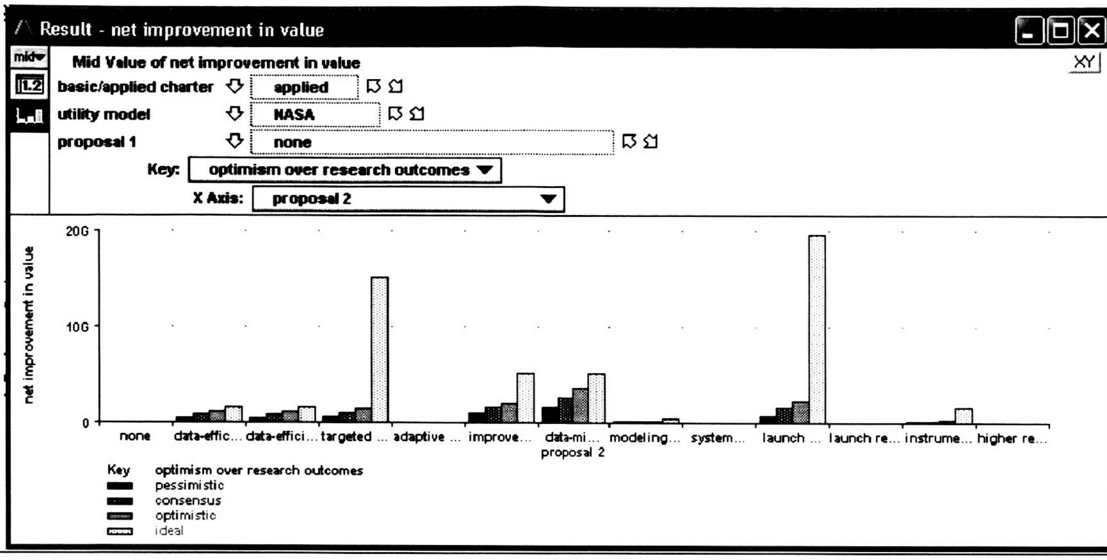
Basic: launch new data source (35M) & targeted observing (22M)
Applied: data-mining (2.5B) & improved ensemble methods (1.5B)

# Evaluating Research Areas



# Sensitivity Analysis

Sensitivity to "optimism" variable: two research areas have vastly higher potential impact under ideal assumptions. Pessimistic view of data-mining exceeds optimistic assessment of other areas.




# Synergy Between Research Areas

We can look for synergies by finding pairs of research areas with much higher value than the two areas individually...

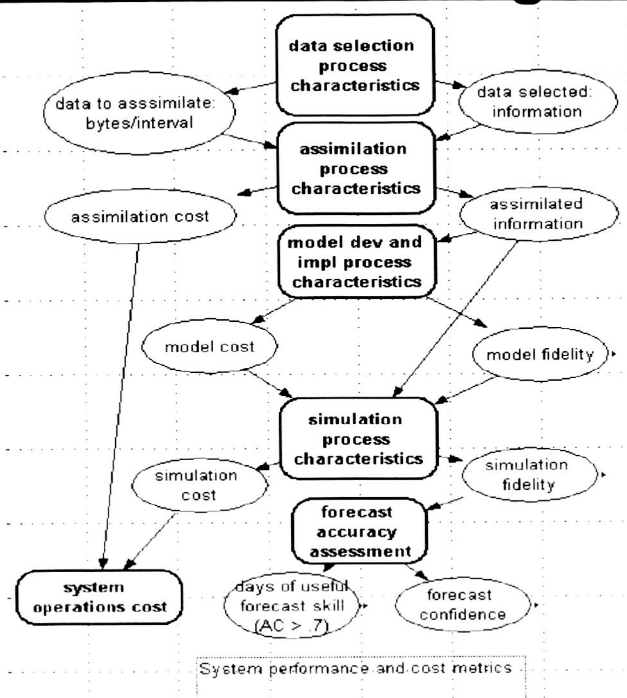
Under the applied research focus:

Biggest synergies

Launch new data source ($1.5B)
+ targeted observing ($1B)
yields a synergy of $700MM

Launch new data source ($1.5B)
+ data-efficient simulations ($800MM) yields a synergy of $400MM

# Understanding the Model




BLUE OVALS summarize the way that system changes flow through the assessment model. We can diagnose our assumptions by analyzing how these variables vary as we vary research area.

# Agenda

CICT Systems Analysis

Our modeling approach
- a 3-part schematic investment model of technology change, impact assessment and prioritization

A whirlwind tour of our model

Lessons learned

## Modeling lessons learned...

*Model and modeling technology should be:*

- understandable and easy to use

*and should support:*

- varying levels of detail (qualitative→quantitative)
- varying scope
  (cross-cutting value as well as mission-specific value)
- development of models by distributed stakeholders
- multiple uses / answer multiple questions
- varying assumptions/priorities
- communication/debate/collaboration

## Lessons learned...

- Model preferences of different stakeholders explicitly
- Allow for easy variation in assumptions ("what if our model is wrong? ...our estimates overly optimistic?")
- Compare impact of each technology to a no-investment baseline
- Make models modular and decoupled:
  technology investments →
    system characteristics →
      performance metrics →
        "return" or "mission value"
  (three arrows == three submodels)

*End of workshop talk…*

*Full report is available at http://support.thinkbank.com/essa-final*