Visual HPC Workflows for the Analysis of System Dynamics Models

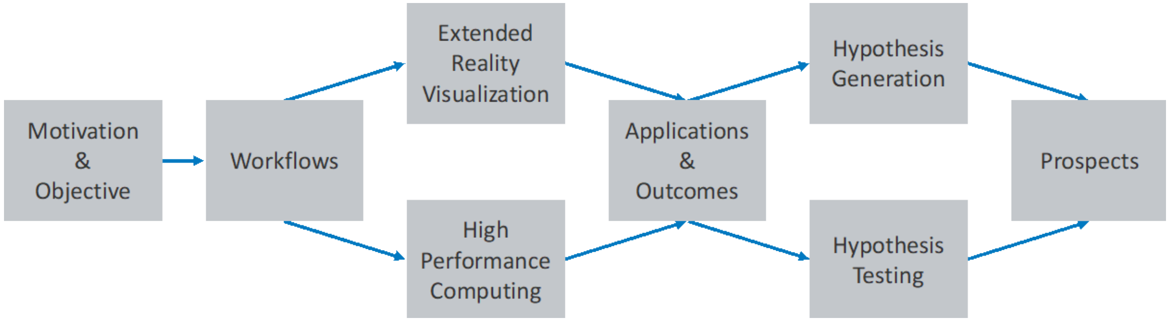
Brian Bush and Danny Inman
National Renewable Energy Laboratory
2021 International System Dynamics Conference
27 July 2021

Abstract

Visual analytics supported by high performance computing (HPC) accelerates and enhances the discovery, exploration, and analysis of causal patterns in complex system dynamics (SD) models. We present a suite of visualization-assisted ensemble-based techniques for hypothesis generation and testing, and for sensitivity analysis. By employing HPC to provide parallel, on-demand simulation of SD models, one can “steer” an ensemble of simulated scenarios in real time as one first formulates and then informally tests those hypotheses: this provides rapid feedback for analysts to refine their understanding of the causal relationships emergent from a model. Such understandings can be followed and augmented by rigorous application of statistical methods, namely global variance-based sensitivity analysis, Monte-Carlo filtering, adaptive regional sensitivity analysis, and self-organized maps: here timely computation relies on HPC, while effective presentation emphasizes high-dimensional multivariate data visualization. Immersive visualization in virtual 3D environments provides an excellent adjunct to the traditional 2D graphics typically used for SD models, as it generates an embodied understanding of model behavior and facilitates an active, collaborative critique of model structure and output. Finally, we summarize prospects for HPC-enabled visual analytics applied to SD modeling.

Overview

Outline



This presentation focuses on techniques that provide information to accelerate the refinement of mental models.

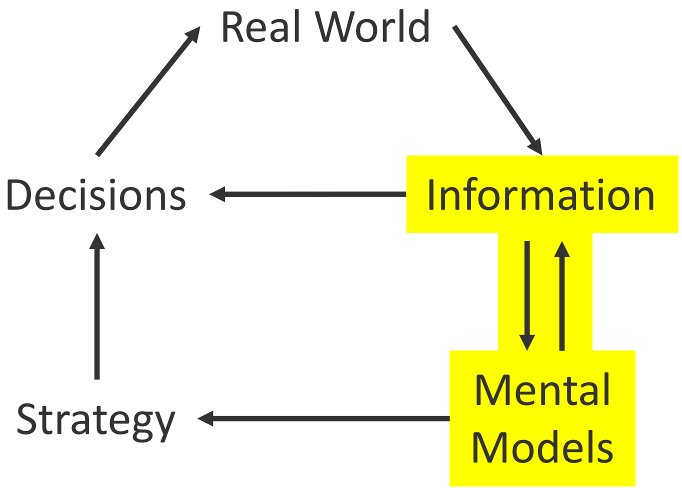


Diagram adapted from Figure 1-11 of double-loop learning in Sterman's *Business Dynamics* (2000).

System-dynamics models mentioned in this presentation:

- Biomass Scenario Learning Model (BSLM)
- Biomass Scenario Model (BSM)
- Bioproduct Transition Dynamics (BTD)
- Lithium Ion Battery Resource Analysis Model (LIBRA)
- Waste-to-Energy Systems Simulation (WESyS)

Motivation and Objective

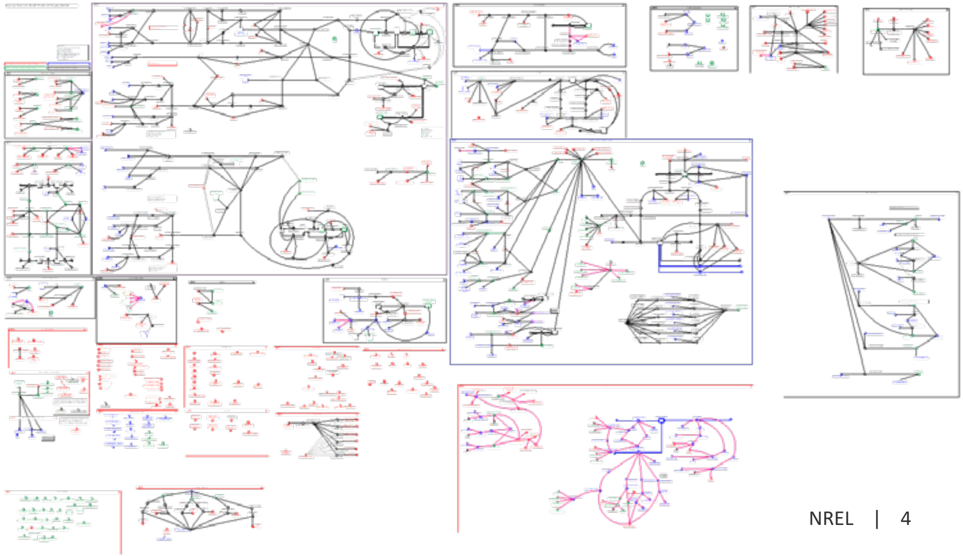
Motivation

- Complex system-dynamics models often have a high degree of dimensionality:
  - *Independent variables:* perhaps thousands of input variables and parameters.
  - *Dependent variables:* perhaps dozens of output variables and metrics.
- Developing insights requires understanding
  - Causal influences.
  - Rich feedback structures.
  - Correlations among multiple variables.
- Higher dimensionality requires computing and examining more simulations: tens of thousands of “scenarios”, “cases”, or “runs”.

Objective

- Augment analysts’ and stakeholders’ resources to rapidly and confidently explore, understand, develop, and test hypotheses and insights in complex models.

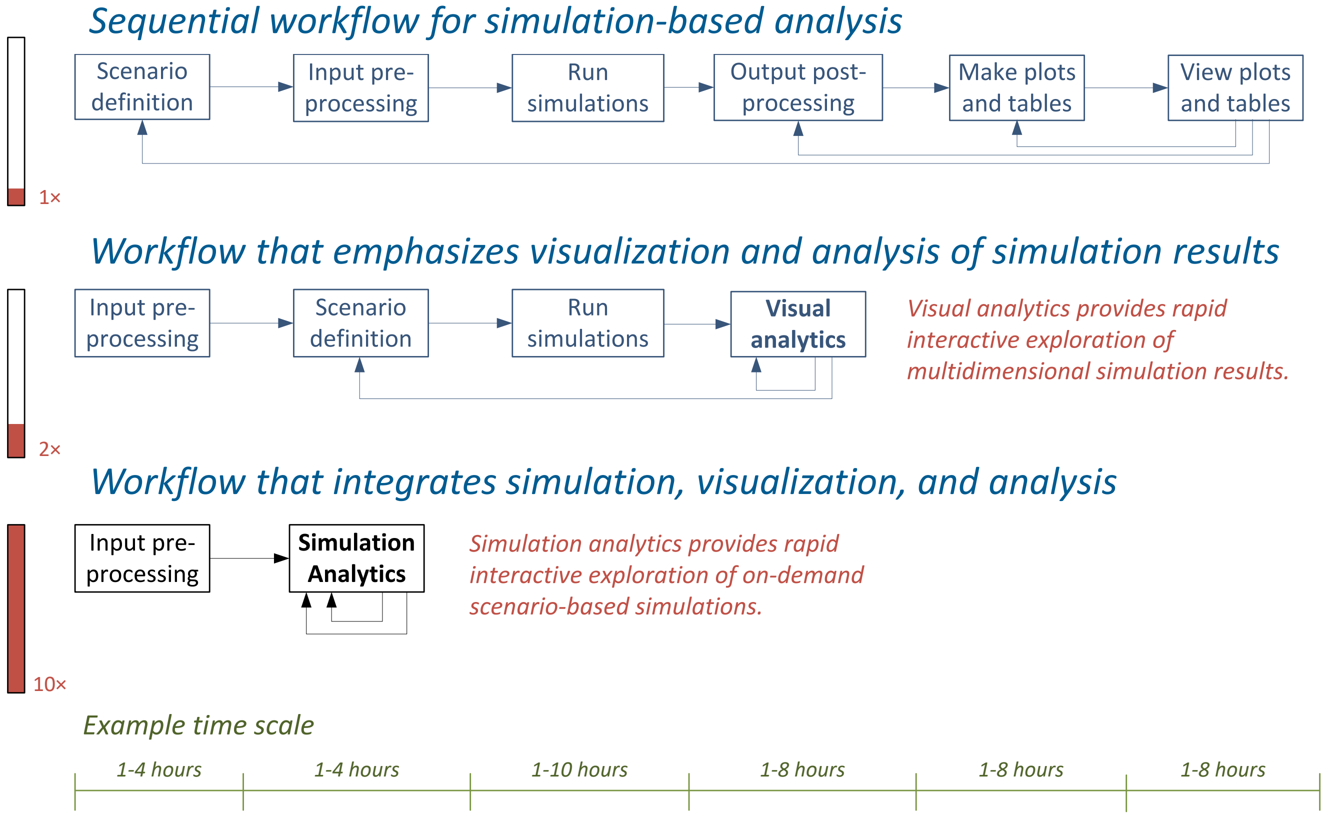
Feedstock module of the Biomass Scenario Model (BSM)



Workflows for Developing Insights into SD Models

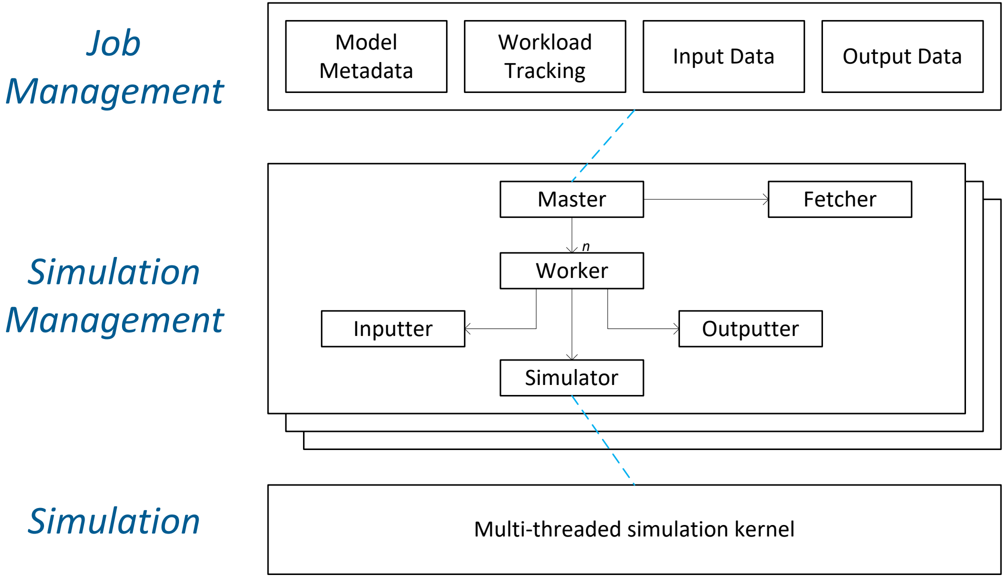
More efficient workflows:

- Shorten time between conceiving a potential insight and testing it.
- Highlight causal patterns and feedbacks involving many variables.
- Explore numerous scenarios using *ensemble steering*.
- Ease cognitive burdens on digesting and remembering patterns in simulation results.
- Facilitate collaboration.

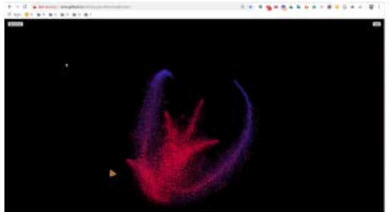


High Performance Computing

- Both commercial SD simulators and custom-built ones can readily be run in high-performance computing (HPC) and cloud-computing environments.
- This allows simultaneous, timely, and responsive computation of *ensembles* of dozens, hundreds, or thousands of SD scenarios.

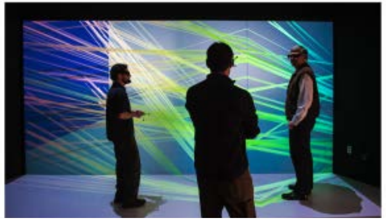


Extended Reality (XR)



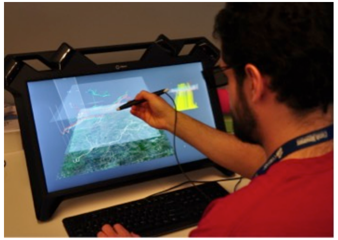
Web Browser

Virtual Reality Headset

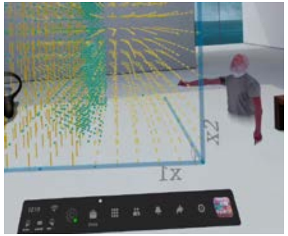


Immersive Vis. Environment

Augmented Reality Headset



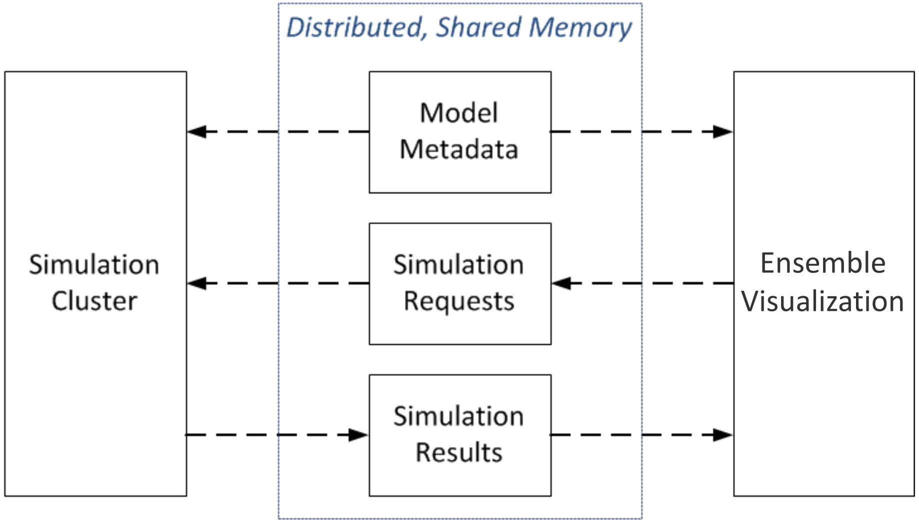
Virtual Reality Desktop



Virtual Reality Cloud

Connecting HPC to Ensemble Visualization

Architectural Principles



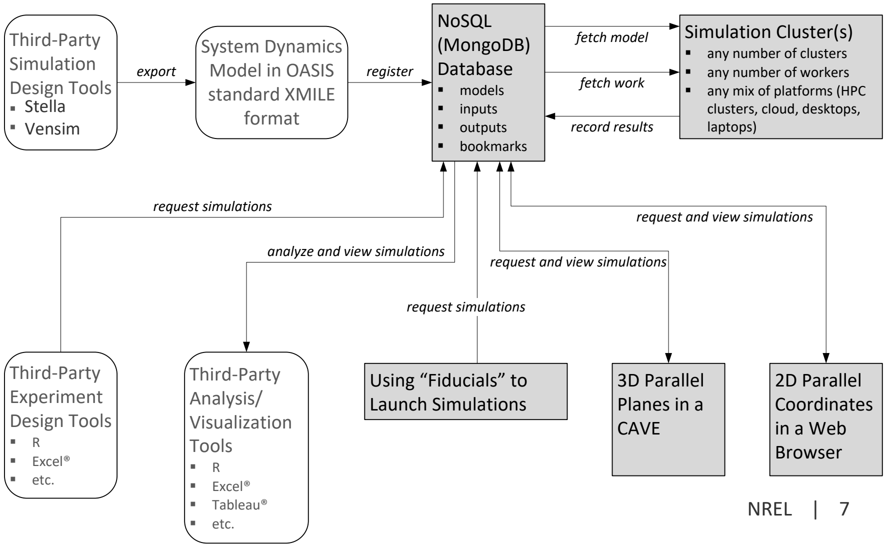
Employing a distributed, shared memory approach allows for loose coupling between the SD simulations and their visualization.

- Simplifies interfaces between components.
- Streamlines error-handling, restarts, and reruns.
- Provides archive of simulation inputs and outputs.

Open-source tools like Python and MongoDB can orchestrate on-demand simulations of ensembles of SD scenarios in HPC, cloud, or desktop environments.

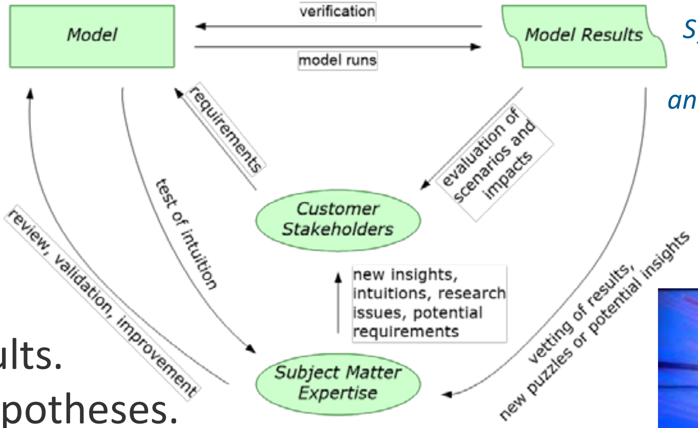
1. Commercial tools like Stella™ and Vensim™ work in such environments but require sufficient runtime licenses.
2. Custom-built simulators can ingest XMILE-format SD models and run large numbers of simulations quickly.

Example Implementation



Applications and Outcomes

- Collaborative extended reality
  - "Hands on" interaction.
  - Filtering and highlighting.
  - Viewing many dimensions.
- Applications
  - Debugging models.
  - Designing scenarios.
  - Identifying anomalous results.
  - Formulating and testing hypotheses.
  - Understanding large sensitivity studies.
- Outcomes
  - Experiential learning.
  - Rapidly understanding.
  - Collaborative design.
  - Information exchange.

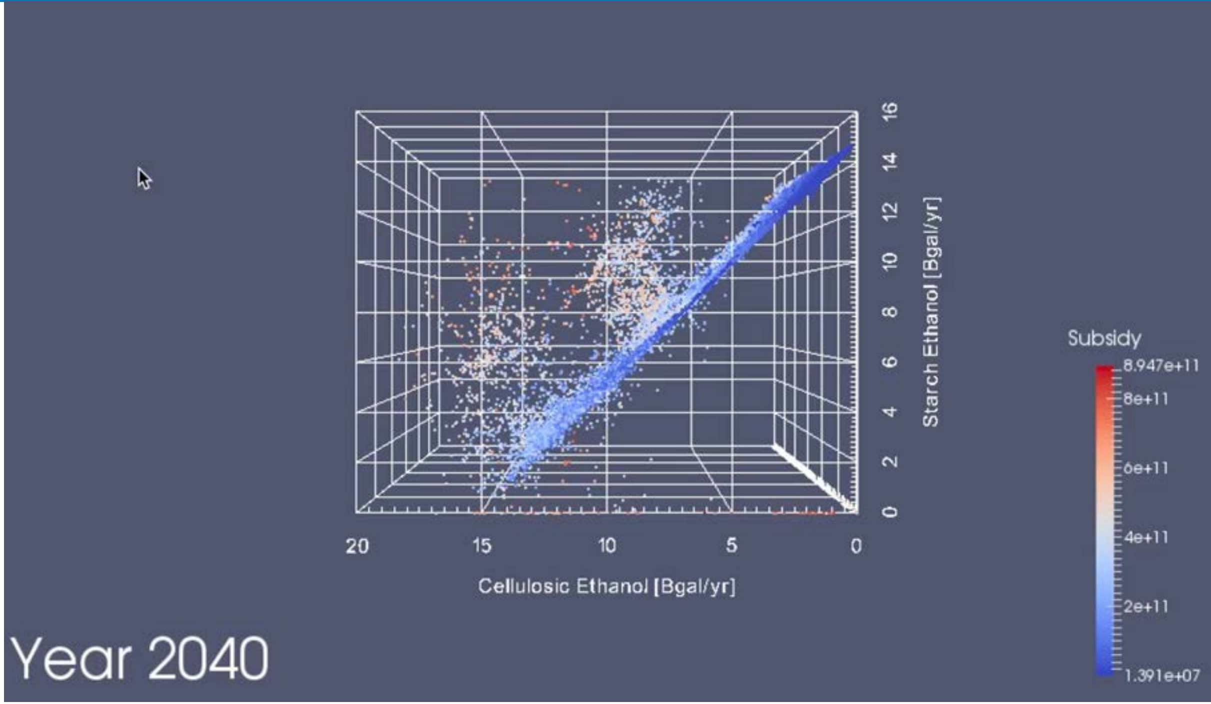


System-dynamics modeling and analysis as a system

Discussion among researchers while interacting with a dynamic, high-dimensional visualization of a system-dynamics model in NREL's Immersive Visualization Environment.

3D Scatterplots of SD Simulation Ensembles

- Biofuel production from three competing technology pathways are on the coordinate axes:
  - Starch ethanol
  - Cellulosic ethanol
  - Hydrocarbon
- Points are colored by the cumulative amount of the government subsidy in the scenario.
- Patterns in the visualization correspond to insights into the biofuel system:
  - The ethanol "blend wall" appears as a diagonal plane.
  - The nationwide availability of land for biomass feedstock production appears as a triangular "cut" on that plane.
  - Synergies between policies appear as points that "escape" the plane.



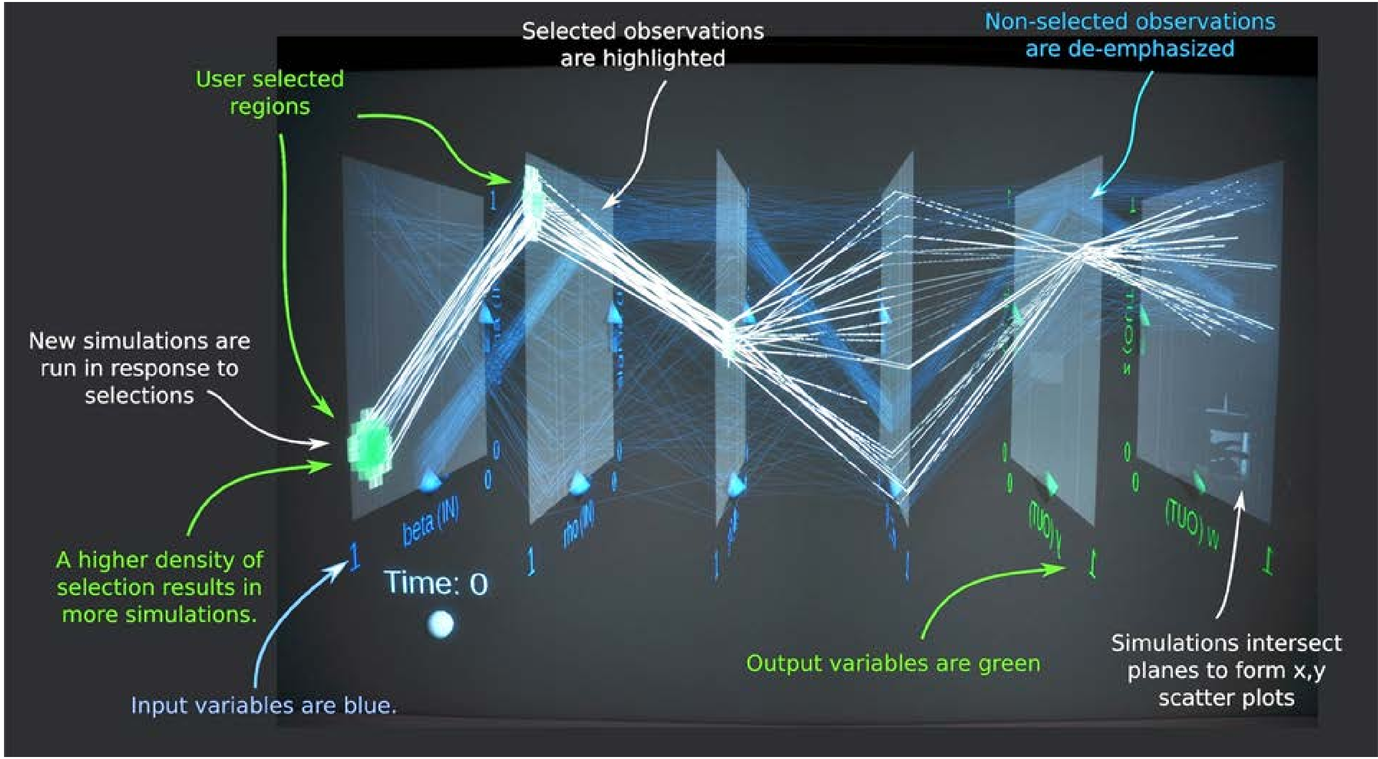
- Ensemble simulation results for the Biomass System Model (BSM).
- Each point represents the result of a system-dynamics simulation.

Parallel Planes in Immersive Virtual Reality

Immersive, interactive 3D *steering of* ensemble visualizations enables collaborative conception and testing of systems hypotheses.

Researchers interactively steering an ensemble of simulations of the Waste-to-Energy Model in NREL's Immersive Visualization Environment.

Parallel-planes visualizations encode high-dimensional ensemble inputs and outputs.



Self-Organizing Maps for Ensembles of SD Simulations

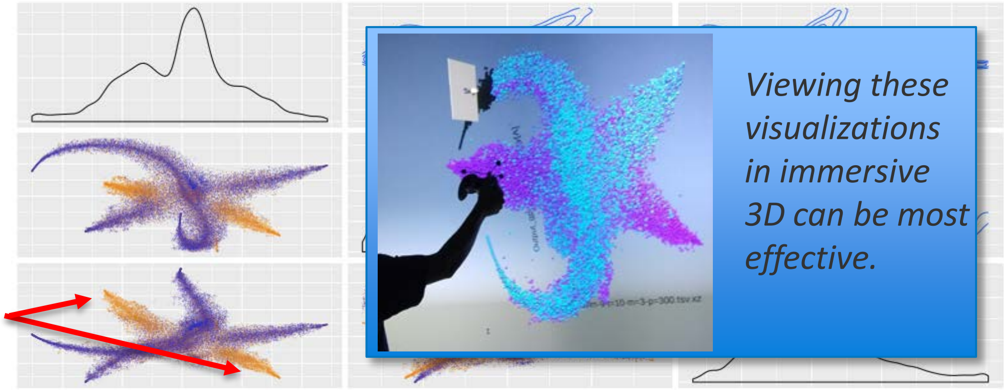
Each point in the display corresponds to the 50-dimensional outcome of a simulation of the Biomass Scenario Learning Model (BSLM).

Scenario outcomes with similar success of *fast pyrolysis pathways* group together naturally.

Scenario outcomes with similar success of *Fischer-Tropsch pathways* group together naturally.

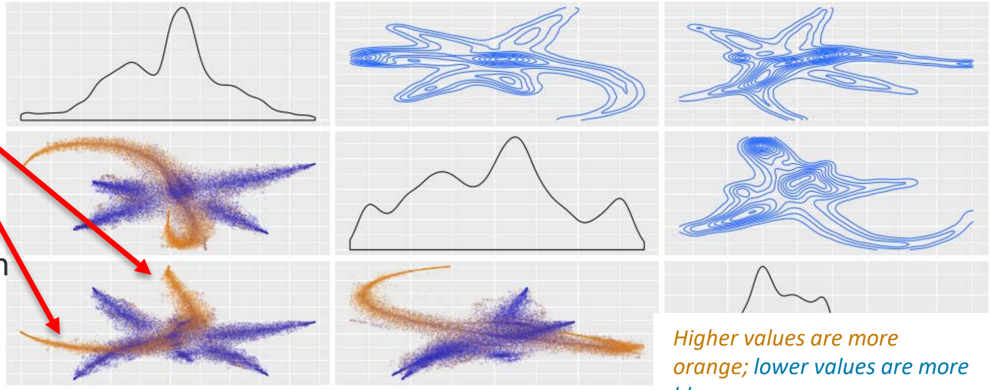
In general, this visualization simultaneously groups scenario outcomes by their similarity *in all dimensions.*

Self-Organized Map colored on biofuel production *via fast pyrolysis.*



Viewing these visualizations in immersive 3D can be most effective.

Self-Organized Map colored on biofuel production *via Fischer-Tropsch.*



Higher values are more orange; lower values are more blue.

- This self-organizing map (SOM) displays an ensemble of Biomass Scenario Learning Model (BSLM) SD simulation results so that scenarios are positioned in 2D or 3D according to the similarity in their output metrics timeseries.
- Exploring such visualizations aids analysts in discovering themes, trends, niches, and extremes.

Workflow: Global Sensitivity

Global Sensitivity Analysis of large SD Models:

- Systematic approach to identifying highly influential model factors
- Rapid model vetting and intuition
- **Factor fixing (FF)**- what factors may be fixed at any value in their range without affecting the model output?
- **Factor prioritization (FP)** - what factor(s) contribute the most to the variance in model output?

For large SD models we use a two-stage approach:

1. Elementary Effects (EE) screening on **all model parameters** to select potentially influential factors (input parameters).

- EE uses a sparse sample space and provides qualitative results of potentially influential factors.

2. Sobol variance decomposition on all **potentially influential** factors.

- Much larger study design
- Calculate the Total effects (FF), First-order (FP), and second-order (FP) indices

Expected value of the conditional variance = the average conditional variance for $X_{i\,(1,\,N)}$

Conditional variance in model output for measure A; factor X_i is fixed at a constant value.

$$S_{Ti} = 1 - \frac{V[E(Y|X_{\sim i})]}{Var\ (Y)}$$

Variance in model output for measure A; all factors varied.

Total Effects (S_{Ti})

Global Sensitivity Analysis for System-Dynamics Models

Identify and characterize specific conditions of interest

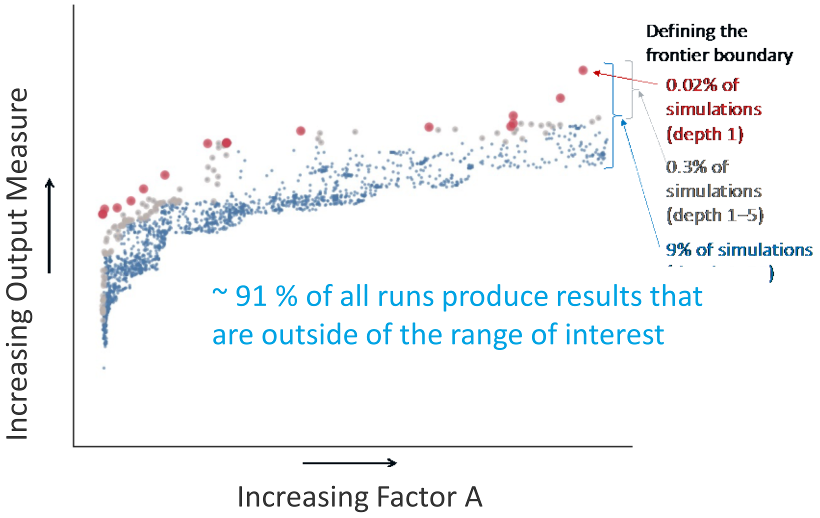
- Parameter space
- System feedbacks
- Likelihood of results of interest given the parameter space

Rank the influence of input parameters on output metrics:

- Singly influential inputs.
- Synergistic combinations of inputs.

The ensemble results from global sensitivity analysis have additional valuable uses:

- Map input-output relationships.
- Select representative scenarios from clusters of qualitatively similar ones.
- Build reduced-form models, using machine learning.
- Estimate Pareto frontiers.



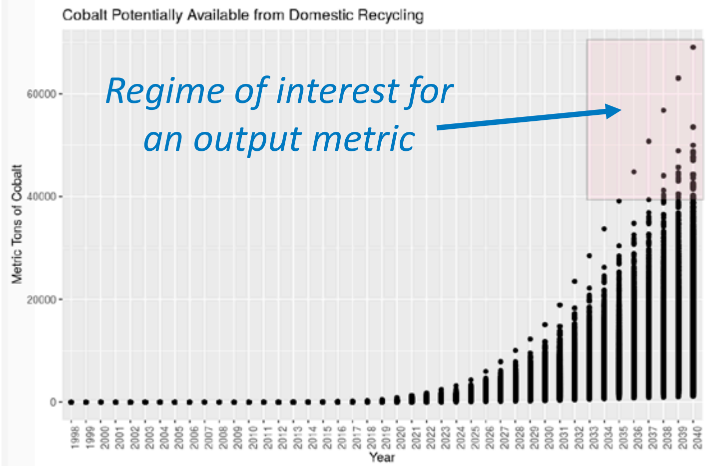
Pareto frontier of investment efficiency for the Biomass Scenario Model (BSM).

Interactive Monte-Carlo Filtering

Interactive Monte-Carlo filtering (MCF)

- Factor Mapping
- What factors conditions are sufficient for a given result(s)
- Statistically significant (bootstrapping)
- Immediate feedback
- Real-time hypothesis testing
- Builds system intuition

Selection/Filtering in Web Browser (Shiny App)

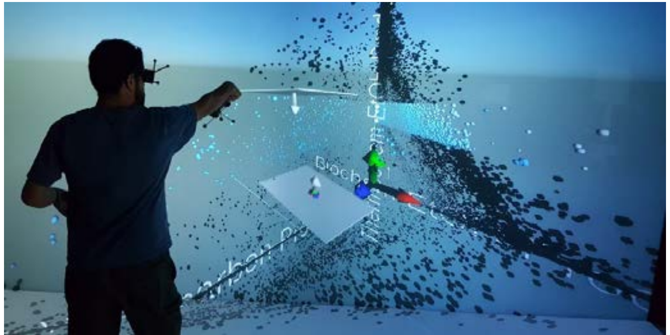


Monte Carlo Filtering

factor	p.value	B.mean	B.bar.mean
EV.battery.recovery.fraction	0.00	0.83	0.50
US.EV.multiplier	0.00	1.86	1.25
static.EV.battery.share.BEV..LCO.	0.00	0.80	0.50
static.EV.battery.share.BEV..NMC811.	0.00	0.24	0.50
static.EV.battery.share.PHEV..NMC622.	0.00	0.46	0.50
avg.time.to.recycle.Co.in.batteries	0.00	0.88	0.62

Ranking of variables predicting that the SD model output will be in the regime of interest

Selection/Filtering in Immersive VR

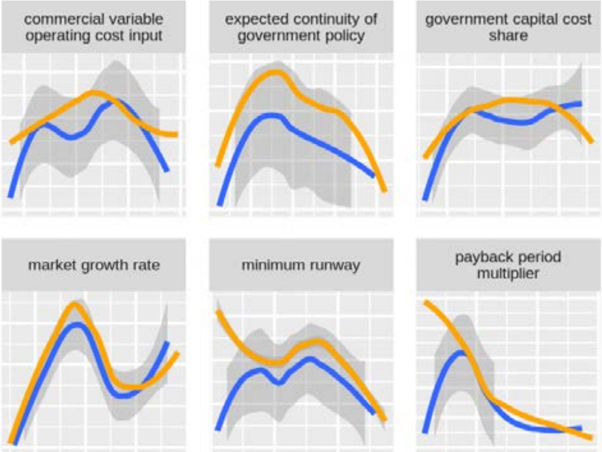


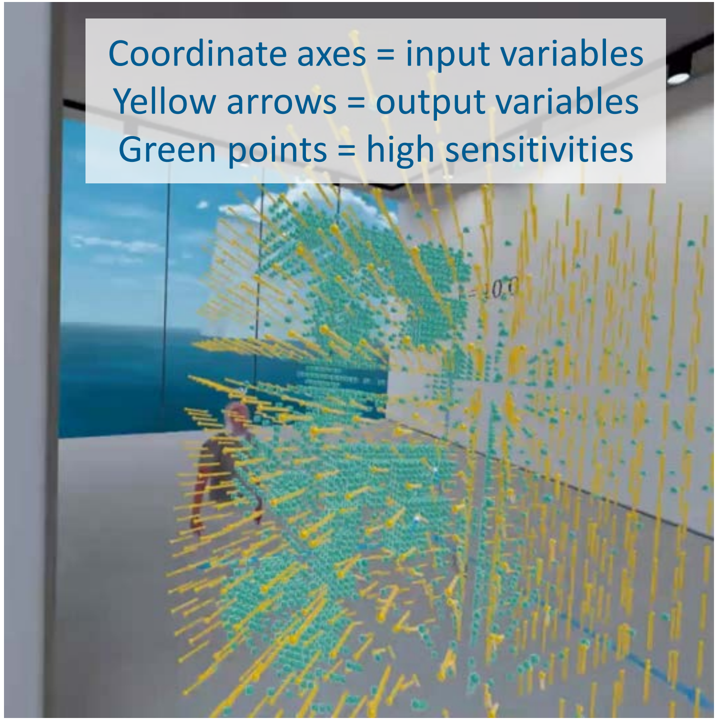
Decomposition of Variance-Based Sensitivity Indices

Method

1. Partition variance according to values of input parameters.
2. Compute local contributions to total sensitivity indices.
3. Summarize as a "density of sensitivity".

Below: *Density of sensitivity with respect to six input parameters in the Bioproducts Transition Dynamics (BTD) SD model: blue lines and gray bands show a small-ensemble estimate, while the orange line shows a large-ensemble result.*





Above: *Virtual reality exploration of variance-based sensitivity analysis of a system-dynamics model via distributed remote collaboration, with participants in Colorado and New Mexico.*

Applications

- *Adaptive sampling:* focus computational resources on regions of highest sensitivity.
- *Visualization:* develop insight into differing regimes of input-output relationships in the SD model.

Summary

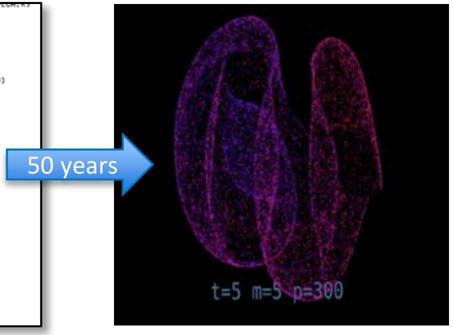
- HPC, cloud, and desktop computing readily enable simulation of large ensembles of system-dynamics scenarios.
- Novel visualization techniques and devices assist analysts in translating large volumes of simulation output into robust insights about system behavior.
- Interactively visual steering of SD ensembles accelerates hypothesis generation, particularly in collaborative environments.
- Statistical techniques such as variance-based sensitivity analysis, self-organizing maps, and Monte-Carlo filtering test hypotheses in real time.

Prospects

- Multimodal, asynchronous, distributed collaboration:
  - Bookmarking, snapshotting, pruning, and sharing ensemble visualizations.
  - Synchronized collaboration among multiple sites, using different visualizations.
  - Stronger links to statistics and desktop tools.
  - Ad-hoc construction of customized visualizations.
- Bundling of similar simulations to reduce cognitive load.
- Attention-triggered proactive computation of new additions to simulation ensembles.
- User-experience testing.

```
31.1   C     POLN=1
31.2   C     POLN1=1
31.3   C     SWT6=1970
32     A     POLCM.K=TABHL(POLCMT,CIR.K,0,5,1)
32.1   T     POLCMT=.05/1/3/5.4/7.4/8
33     R     POLA.KL=POL.K/POLAT.K
34     A     POLAT.K=TABLE(POLATT,POLR.K,0,60,10)
34.1   T     POLATT=.6/2.5/5/8/11.5/15.5/20
35     L     CIAF.K=CIAF.J+(DT/CIAFT)(CFIFR.J*CIQR.J-CIAF.J)
35.1   N     CIAF=CIAFI
35.2   C     CIAFI=.2
35.3   C     CIAFT=15
36     A     CFIFR.K=TABHL(CFIFRT,FR.K,0,2,.5)
36.1   T     CFIFRT=1/.6/.3/.15/.1
37     S     QL.K=(QLS)(QLM.K)(QLC.K)(QLF.K)(QLP.K)
37.1   C     QLS=1
38     A     QLM.K=TABHL(QLMT,MSL.K,0,5,1)
38.1   T     QLMT=.2/1/1.7/2.3/2.7/2.9
39     A     QLC.K=TABLE(QLCT,CR.K,0,5,.5)
39.1   T     QLCT=2/1.3/1/.75/.55/.45/.38/.3/.25/.22/.2
40     A     QLF.K=TABHL(QLFT,FR.K,0,4,1)
40.1   T     QLFT=0/1/1.8/2.4/2.7
41     A     QLP.K=TABLE(QLPT,POLR.K,0,60,10)
41.1   T     QLPT=1.04/.85/.6/.3/.15/.05/.02
       NOTE        EQUATION 42 LOCATED BETWEEN EQ. 4 AND 9.
42     A     NRMM.K=TABHL(NRMMT,MSL.K,0,10,1)
42.1   T     NRMMT=0/1/1.8/2.4/2.9/3.3/3.6/3.8/3.9/3.95/4
       NOTE        INPUT FROM EQN. 38 AND 40 TO EQN. 35
43     A     CIQR.K=TABHL(CIQRT,QLM.K/QLF.K,0,2,.5)
43.1   T     CIQRT=.7/.8/1/1.5/2
       NOTE
       NOTE  CONTROL CARDS
       NOTE
43.5   C     DT=.2
43.6   C     LENGTH=2100
43.7   N     TIME=1900
44     A     PRTPER.K=CLIP(PRTP1,PRTP2,PRSWT,TIME.K)
```

World Dynamics Model in DYNAMO



3D SOM Visualization of World Dynamics Model

Resources

- Publications
  - “High Performance Interactive System Dynamics Visualization.” 2016. https://www.osti.gov/servlets/purl/1393382
  - “Simulation Exploration through Immersive Parallel Planes.” 2017. https://doi.org/10.1109/IMMERSIVE.2016.7932377
  - “Application of Variance-Based Sensitivity Analysis to a Large System Dynamics Model.” 2018. https://arxiv.org/abs/1803.10722
  - “Coupling Visualization, Simulation, and Deep Learning for Ensemble Steering of Complex Energy Models.” 2018. https://doi.org/10.1109/DSIA.2017.8339087
  - “Application of a Variance-Based Sensitivity Analysis Method to the Biomass Scenario Learning Model.” 2018. https://doi.org/10.1002/sdr.1594
  - “Enabling Immersive Engagement in Energy System Models with Deep Learning.” 2019. https://doi.org/10.1002/sam.11419
  - “Collaborative Exploration of Scientific Datasets Using Immersive and Statistical Visualization.” 2021. https://www.osti.gov/servlets/purl/1805190
- Models
  - Biomass Scenario Learning Model: https://github.com/NREL/bsm-learning
  - Biomass Scenario Model: https://github.com/NREL/bsm-public
  - Connected and Automated Vehicles Scenario Generation: https://github.com/NREL/CSG-public
  - Global Land Use Change: https://github.com/NREL/bioluc
  - Waste-to-Energy System Simulation: https://github.com/NREL/WESyS-Model
- Software
  - https://dinman.shinyapps.io/CoModel/
  - https://github.com/NREL/infovis-parallel

Thank You

brian.bush@nrel.gov

daniel.inman@nrel.gov

www.nrel.gov

NREL/PR-6A20-80539

This work was authored by the National Renewable Energy Laboratory, operated by Alliance for Sustainable Energy, LLC, for the U.S. Department of Energy (DOE) under Contract No. DE-AC36-08GO28308. Funding provided by the U.S. Department of Energy Office of Energy Efficiency and Renewable Energy Bioenergy Technologies Office. Part of this work was supported by the Laboratory Directed Research and Development (LDRD) Program at NREL. Part of this research was performed using computational resources sponsored by the Department of Energy's Office of Energy Efficiency and Renewable Energy and located at the National Renewable Energy Laboratory. The views expressed in the article do not necessarily represent the views of the DOE or the U.S. Government.

NREL Transforming ENERGY