BSM Sensitivity Analysis and Meta-Modeling Next Steps

Daniel Inman and Dylan Hettinger

FY 23 Q4 Annual Milestone Briefing

2023.08.08

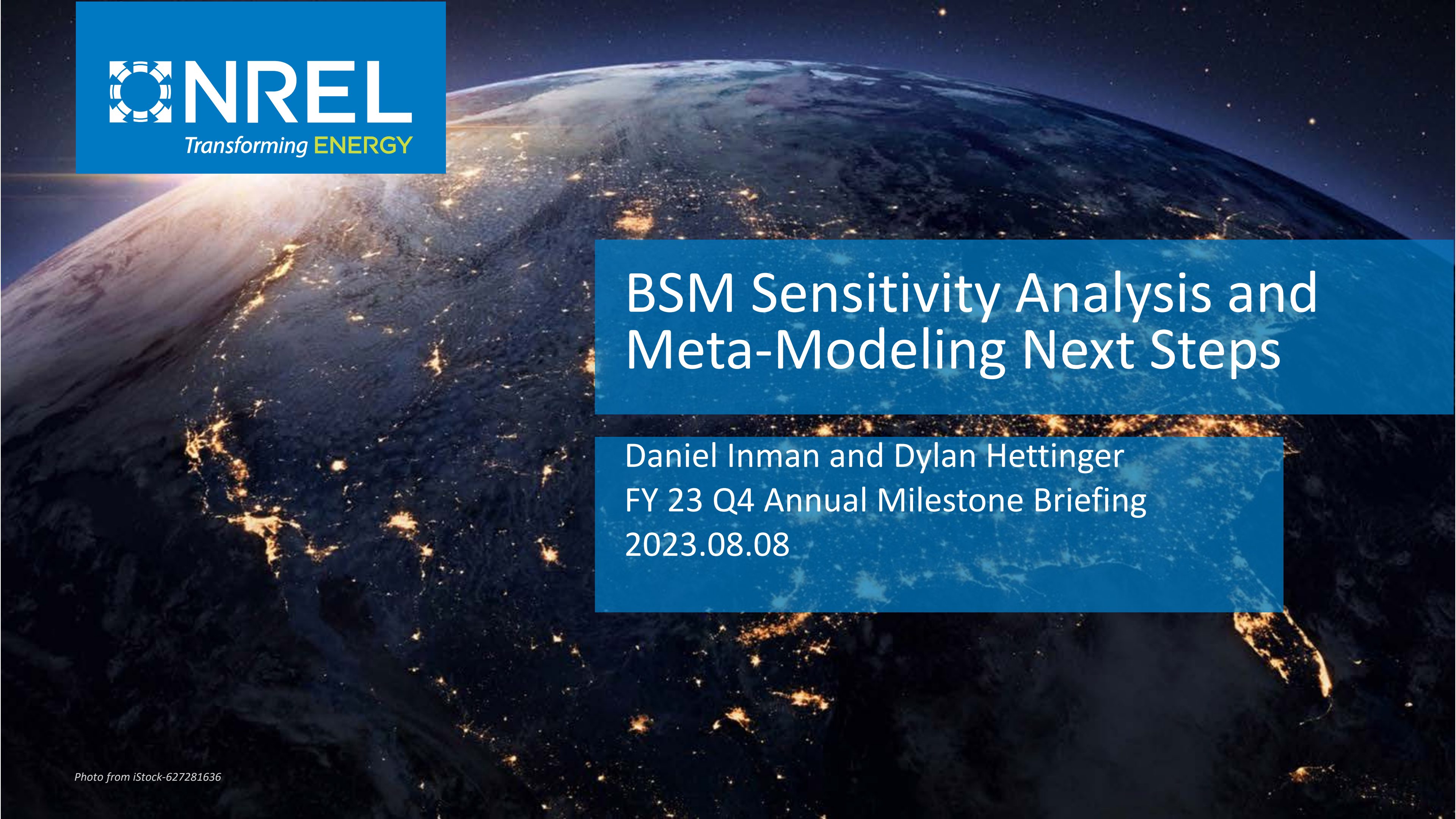


Photo from iStock-627281636

Definitions

Bioenergy Scenario Model

- **Uncertainty analysis**: Focuses on just quantifying the uncertainty in model output
- **Sensitivity analysis**: The study of the relative importance of different input factors on the model output
- **Global sensitivity analysis**: "The study of how the uncertainty in the output of a model (numerical or otherwise) can be apportioned to different sources of uncertainty in the model input"
- **Factor prioritization**: Used to identify a factor (or group of factors) which, when fixed to its true value, leads to the greatest reduction in the variance in model output – this identifies the factor that accounts for the greatest amount of variance in model output. *First-order, Nth-order indices*
- **Factor fixing:** Used to identify factors in the model, when varied over their uncertainty range, make no significant contribution to the variance in model output. *Total Effects*

Model Factors and Output

Bioenergy Scenario Model

Sustainable aviation fuel (SAF)

- Cellulose to hydrocarbons factors
- Oils to hydrocarbons factors
- Wet waste factors
- Starch to jet factors

Policies, TEA, feedstock parameters, learning, technology readiness level

$Y = f(X_1, X_2, \ldots, X_k) \Rightarrow \text{SAF (gal/yr)} = f(CHC_{1\text{-}n}, OHC_{1\text{-}n}, WW_{1\text{-}n}, SJ_{1\text{-}n})$

Elementary Effects Sensitivity Analysis

Bioenergy Scenario Model

Elementary Effects (EE)

- First stage for large studies
- Relative contribution to output variance
- Exact calculation for first-order effects
- Estimates higher-order interactions
- Sufficient surrogate for total effects
- **Useful for screening non-influential factors**

- Elementary Effects(EE): $\boldsymbol{EE_i^j}$ for each input $\boldsymbol{i_1,\ldots,i_k}$ and trajectory j are calculated by varying a single input a fixed amount Δ.
- Input values are determined using random trajectories $\boldsymbol{j_1, \ldots, j_n}$ through the input space $\boldsymbol{x_1, \ldots, x_M}$.
- The standard deviation $\boldsymbol{\sigma_i}$ is interpretated as the intensity of higher order interactions.
- The mean μ_i over each $\boldsymbol{EE_i}$ is taken as an estimate of total effects.
- $\boldsymbol{g_i^*}$ is calculated as the $\boldsymbol{\ell^2}$ norm of $\boldsymbol{\mu^*}$ and σ to represent total effects augmented by interactions.

Sensitivity Index Convergence

- A primary challenge with large studies is knowing how many runs are required to show stable results.
- To show convergence of each sensitivity index and calculate S_i, which is simply each g_i^* normalized by the largest g_i^* in the study.
- We then calculate $Stat_{indices}$, the maximum range of the 95% confidence interval of S_i across all inputs.
- When $Stat_{indices}$ is less than a given threshold, we say the EE_i has converged to a stable estimate.

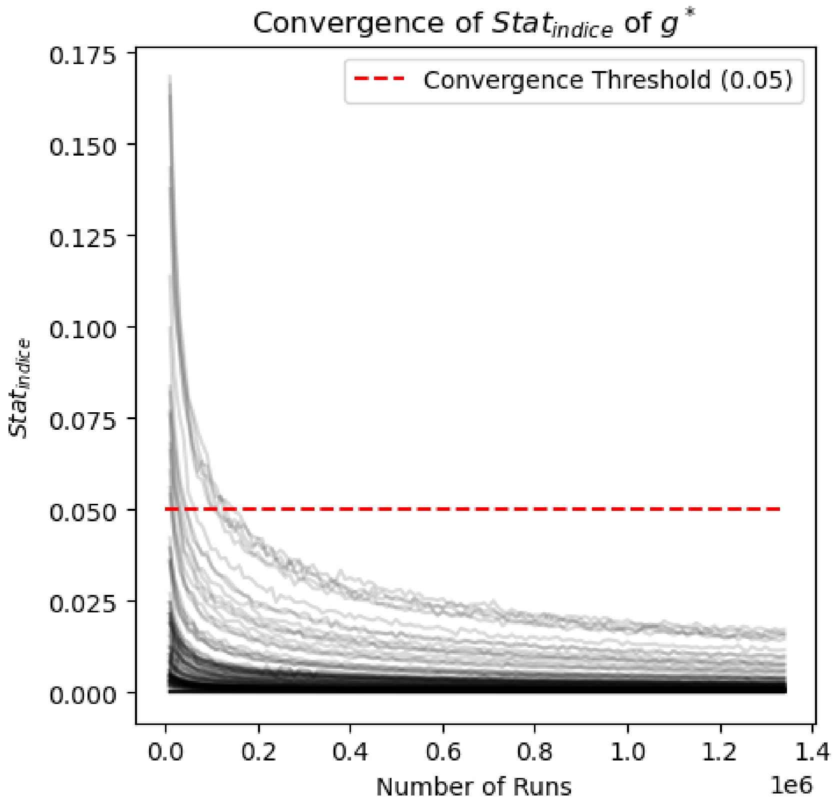
$$\mathrm{S}_i = \frac{g_i^*}{\max_k g_i^*}$$

$$Stat_{indices} = \max_{i=1..M} (S_i^{ub} - S_i^{lb})$$

Elementary Effects Analysis Results

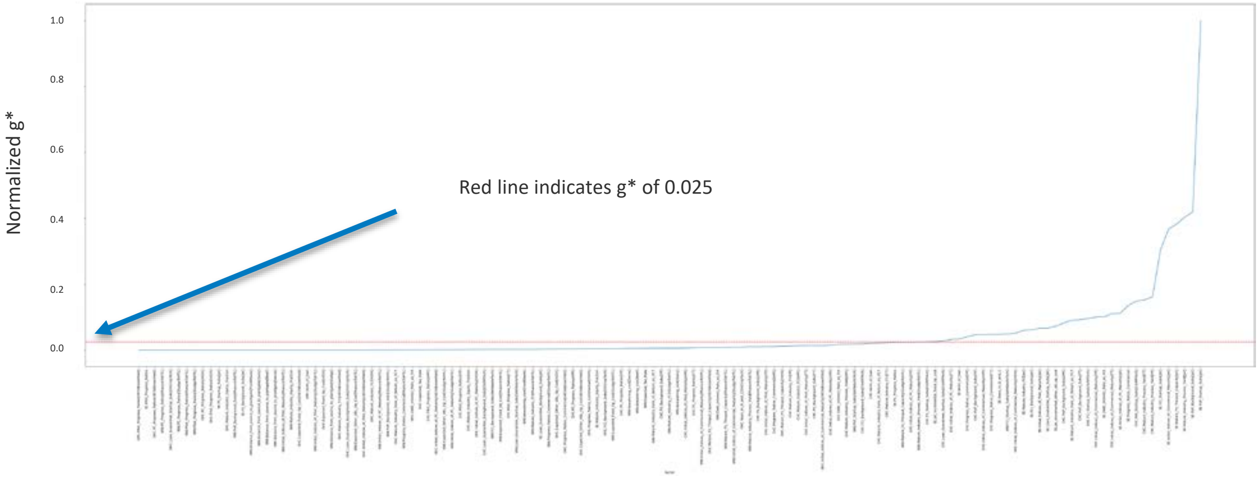
EE: Sensitivity Index Convergence

- The initial Elementary Effects study consisted of about 4.5 million runs.
- Using a literature-recommended threshold of 0.05, we see reasonable convergence by about 750k runs, and clear convergence with the full set of runs.



EE: Influential Factor Screening

- By plotting the sorted, normalized *g** values, we see an increased slope starting at a normalized *g** value of about 0.025.
- Any factor with a normalized *g** value less than 0.025 is considered non-influential.



EE: Factor Screening Validation (1 of 2)

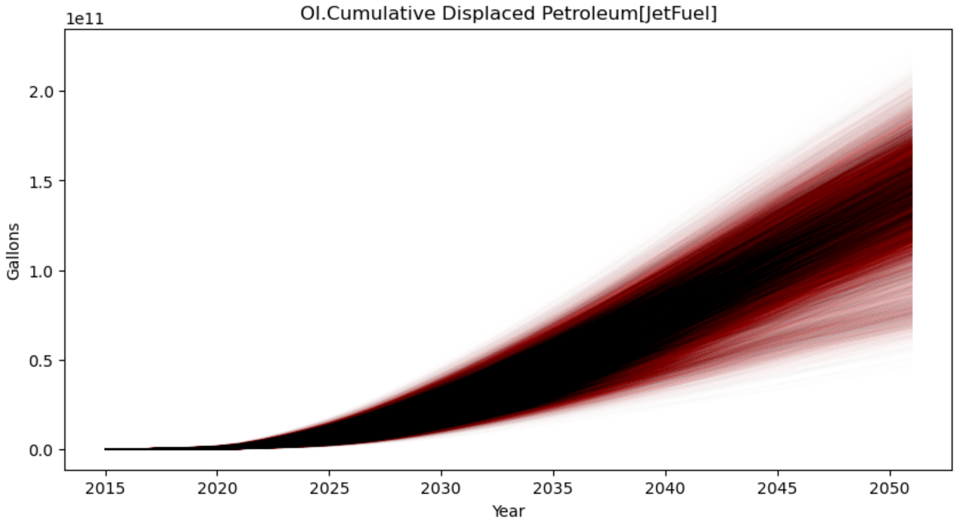
Bioenergy Scenario Model

To validate the threshold two new studies are conducted:

1. All factors are randomly varied (black).
2. Only the influential factors are varied and any factor with a g* below the threshold is left at its default value (red).

The output distribution should remain consistent between both studies if the factors deemed non-influential indeed have no effect on the output.

A two-sided Kolmogorov-Smirnov test provides statistical evidence that non-influential factors have been correctly identified.



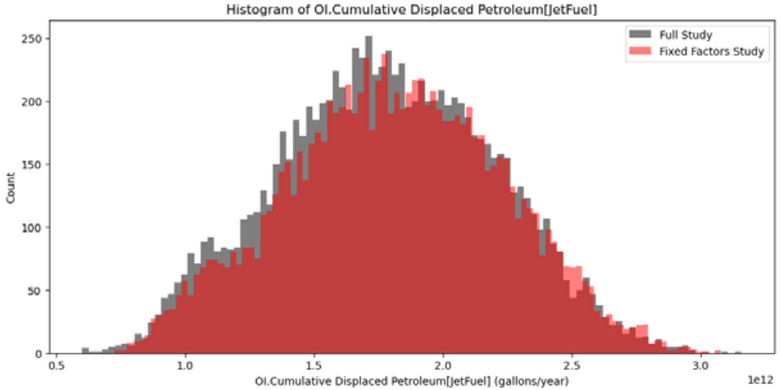
Sustainable aviation fuel (SAF) production (gallons of petroleum fuel displaced by SAF) from 2015 – 2050. Red lines represent individual model runs where only influential factors were varied while the black lines represent individual model runs where all factors were varied.

EE: Factor Screening Validation (2 of 2)

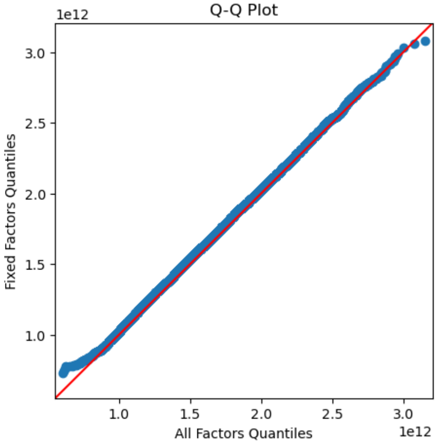
Bioenergy Scenario Model

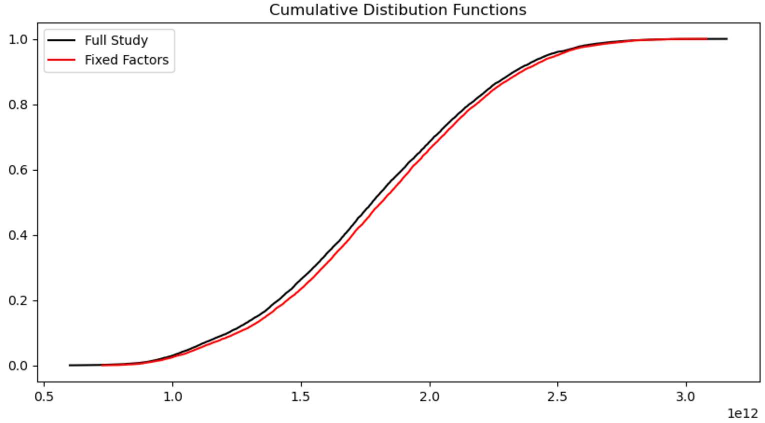
The output distributions are largely consistent, with some minor differences evidenced by the histograms, cumulative distribution functions (CDF), and a Q-Q plot.

As seen in the histogram and Q-Q plot, the full study has slightly more high and low values. The CDF suggests the full model is also slightly more likely to produce a higher value in the output than the fixed factors study.



A two-sided Kolmogorov-Smirnov test provides further evidence that non-influential factors have been correctly identified (bootstrapped p-value of 0.44).

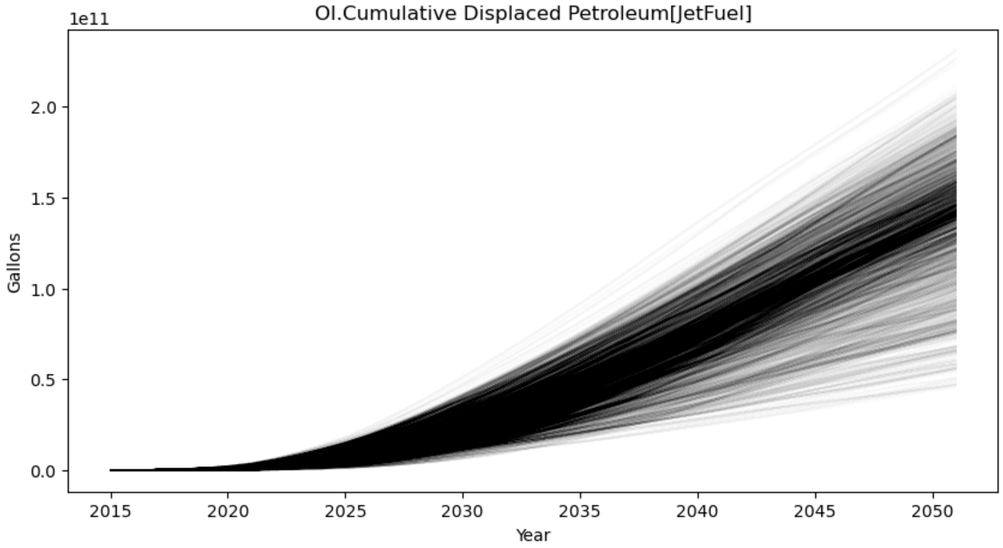




EE: Total SAF Production (2015 – 2050)

Bioenergy Scenario Model

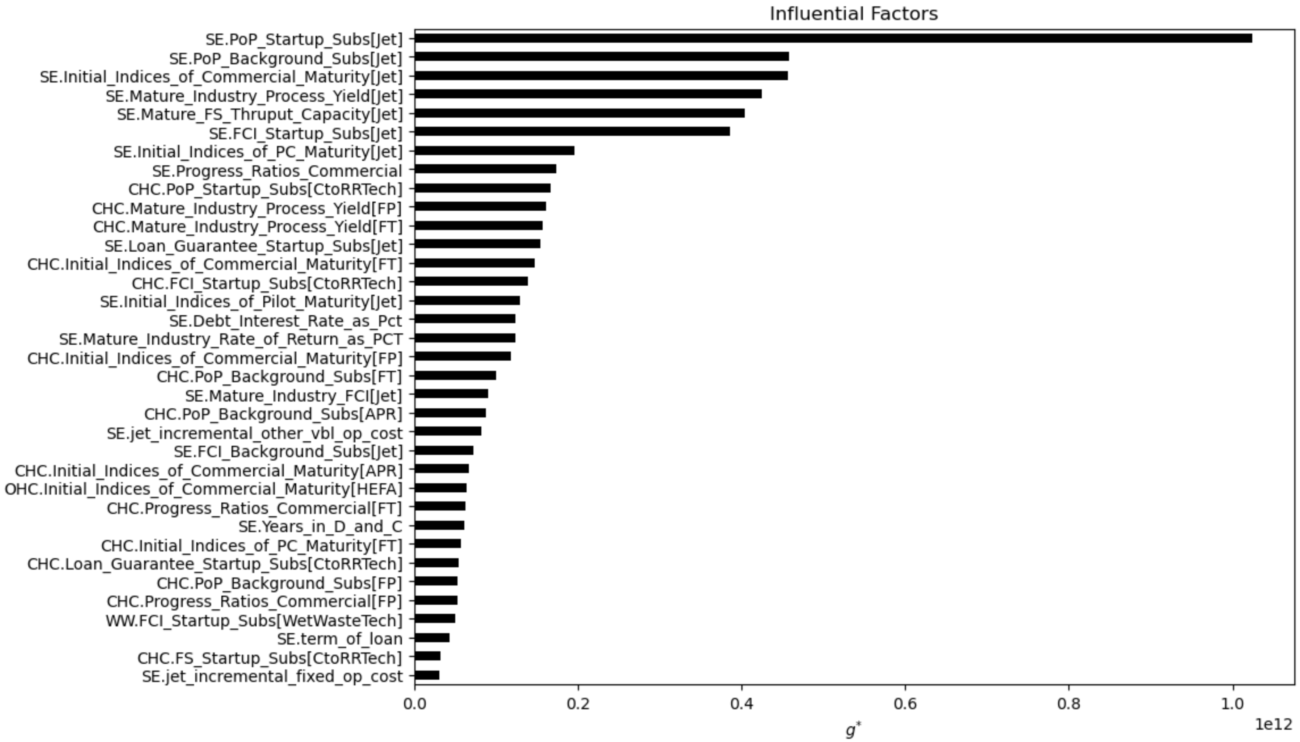
- Studying the influence of input parameters on cumulative petroleum displaced by SAF
- Identified 220 factors across 4 modules (CHC, OHC, SE, and WW)
- Across all model runs, total SAF production in 2050 ranged from 50 to over 200 billion gallons



Sustainable aviation fuel (SAF) production (gallons of petroleum fuel displaced by SAF) from 2015 – 2050. Lines represent individual model runs from the elementary effects study.

EE: Influential Factors

Bioenergy Scenario Model

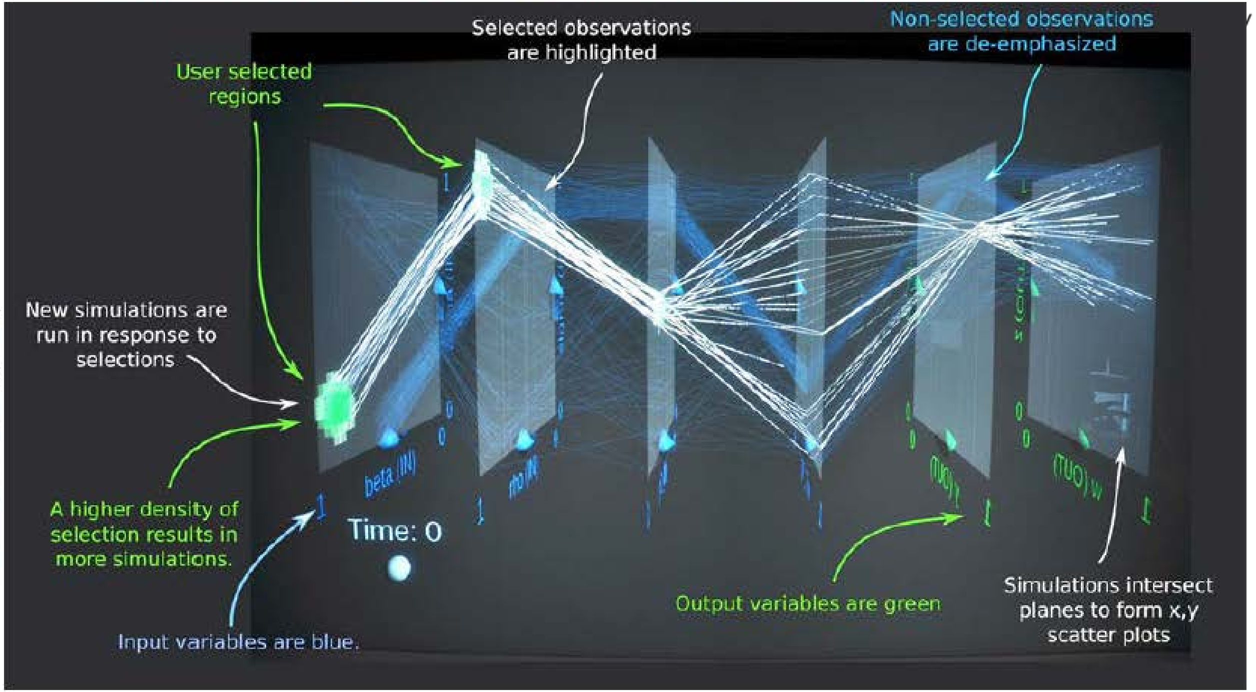


35 model factors were identified through EE analysis as being potentially influential. Model factors are listed in order of influence along the Y axis. The normalized g* score is listed on the X axis. Note that the factor names are in their model input format.

Next Steps: Data Exploration and Meta Modeling

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

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



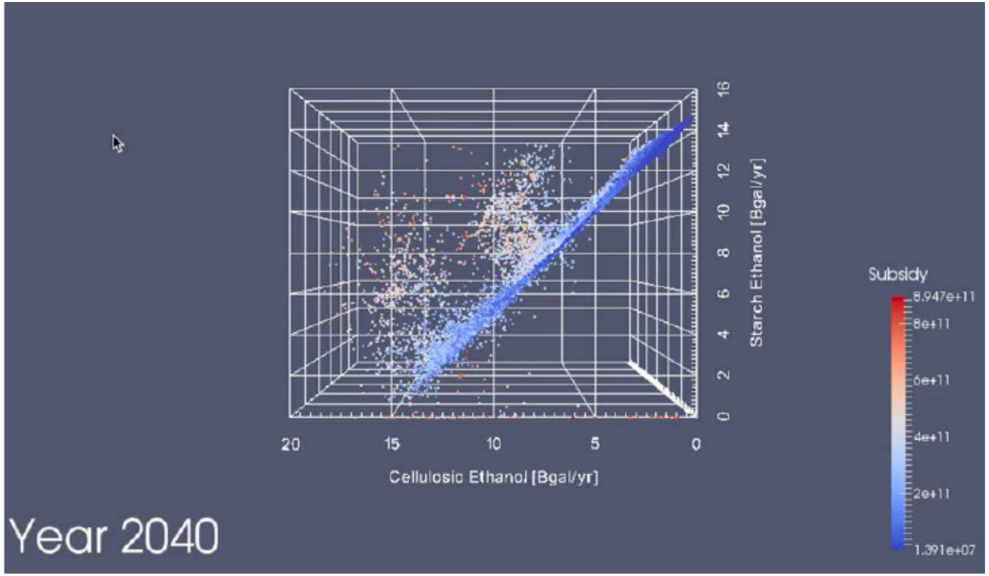
Researchers interactively steering an ensemble of simulations of the Waste-to-Energy Model in NREL's Immersive Visualization Environment.

Advanced Computing and System Dynamics Integration

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.

- 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.

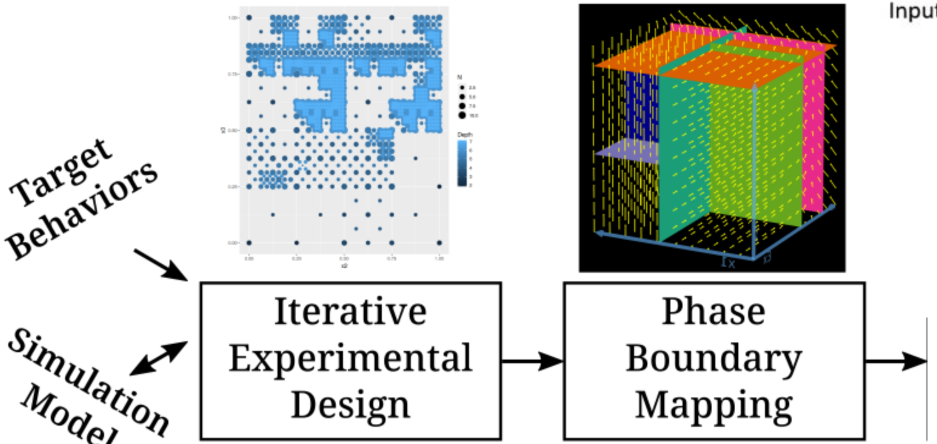


- Ensemble simulation results for the Biomass System Model (BSM).
- Each point represents the result of a system-dynamics simulation.

Bioenergy Scenario Model

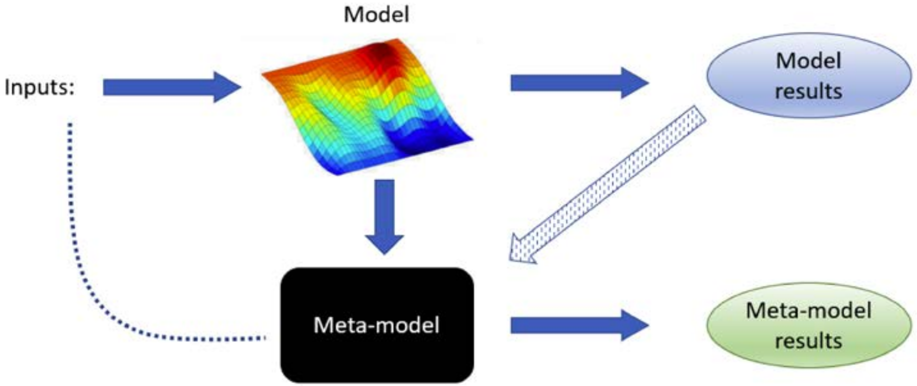
Advanced Computing and System Dynamics Integration

Visualizing thousands of simulations in 3-D allows us to identify insightful results that leads to a deeper system understanding.



1. Reinforcement learning uncovers target behaviors via nested, adaptive design of experiments.

2. Edge detection, Monte Carlo filtering, and regional sensitivity analysis delimit borders of active subspaces.



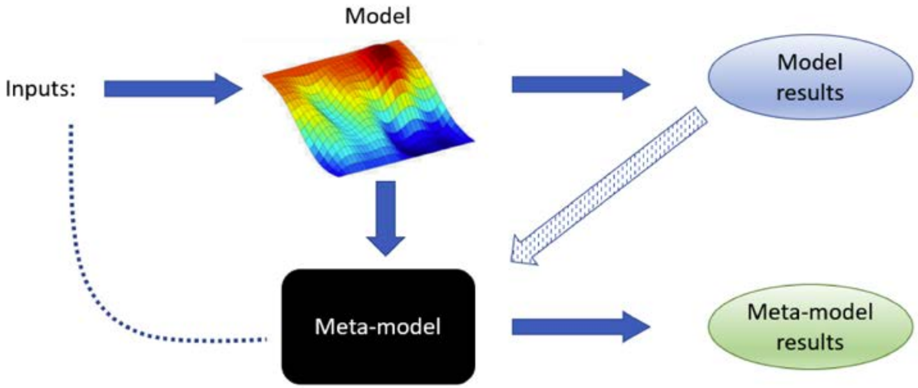
$$Y = f(X_1, \ldots, Xk)$$

Integration of System Dynamics Models

Capturing the dynamic representation in system dynamics model through meta-modeling and model reduction approaches allows these models to be integrated into larger modeling efforts.

FY24 Plans

- Factors identified through EE will be further evaluated using variance-based sensitivity analyses.
- Further sensitivities will be focused on SAF production from HEFA and individual pathways.
- The reduced form model will be rigorously tested and validated using the extensive data generated during both the EE and VBSA performed in FY23.



What impactful analysis is yet to be done?

Bioenergy Scenario Model

Reduced-form BSM

- The BSM is a mature model with lots of thought put into it by DOE, modelers, and analysts.
- In its current form, the BSM publicly available, but not an accessible platform (https://github.com/NREL/bsm-public).
- A reduced-form model coded in an open-source language will increase usage.
- This will help to inform questions with multi-sector interest by being able to combine the BSM with other models.

Important Questions

- How do renewable diesel and sustainable aviation fuel compete and complement each other?
- What impact can DOE have in an electrified future?
- What synergies exist with e-fuels and biofuels?

Thank You

www.nrel.gov

NREL/PR-6A20-88045

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. The views expressed in the article do not necessarily represent the views of the DOE or the U.S. Government.

Photo from iStock-627281636

NREL
Transforming ENERGY