SENSITIVITY ANALYSIS FOR DECISION-MAKING: FROM DECISION TREES TO ARTIFICIAL INTELLIGENCE

Emanuele Borgonovo
Bocconi University and Massachusetts Institute of Technology

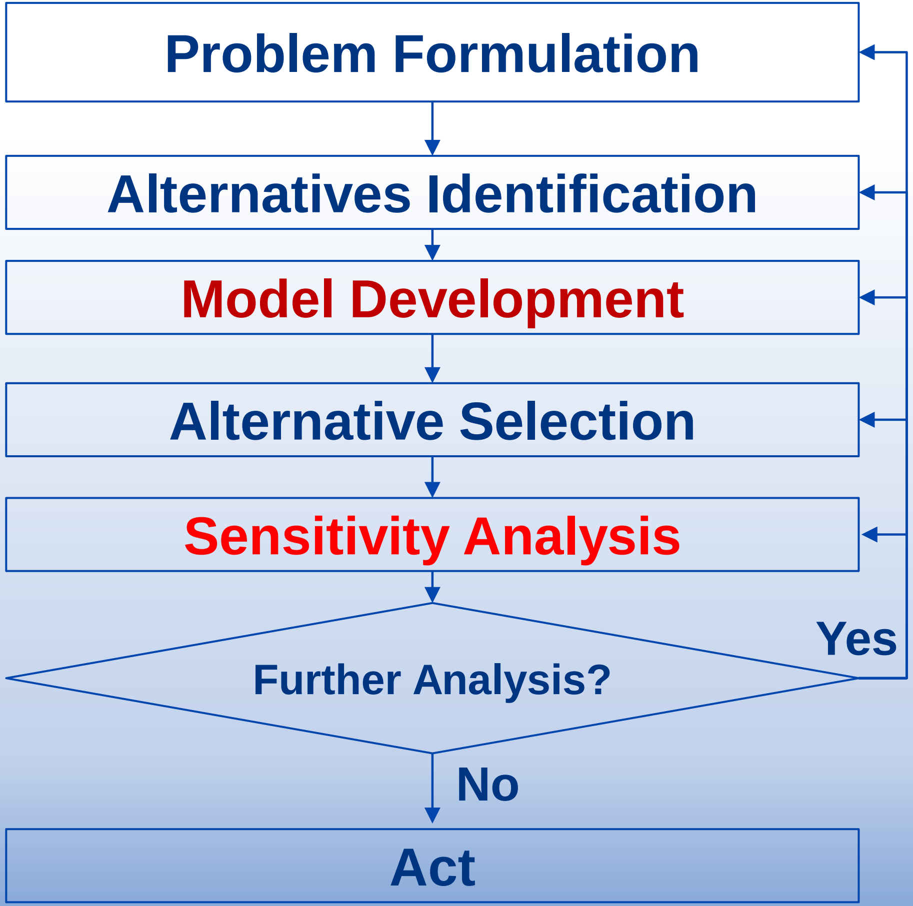
Università Bocconi
MILANO

Nuclear Science and Engineering

Premise

We increasingly rely on models to make decisions

The Decision Analysis Cycle



Clemen, 1997
Figure 1.1

Back to Howard 1968

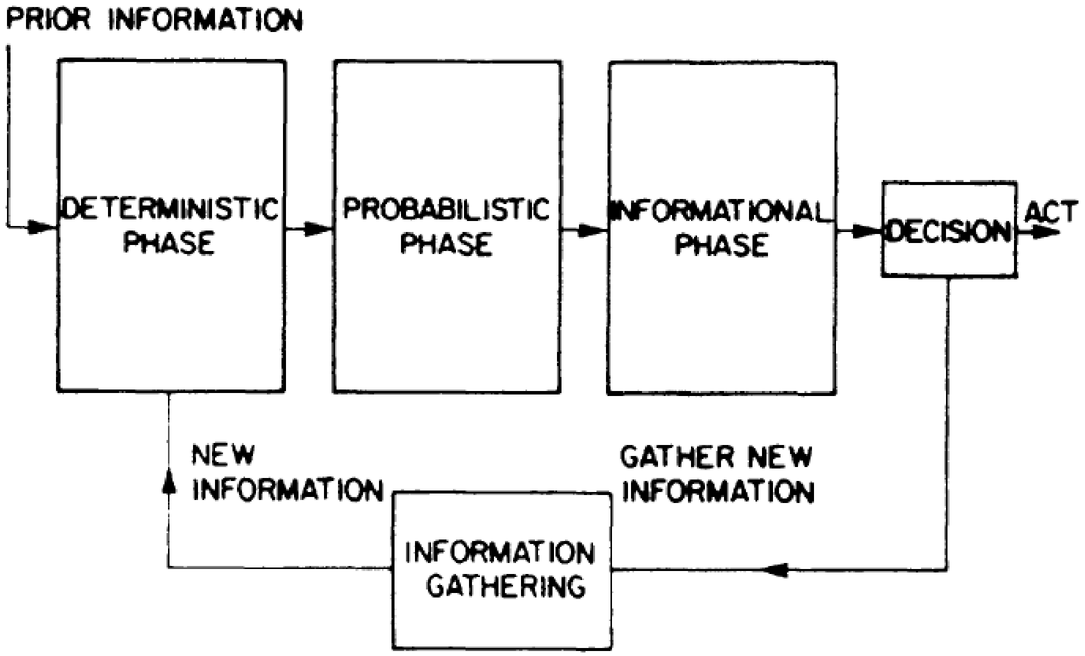


Fig. 3. Decision analysis cycle.

Sensitivity in Howard (1968)

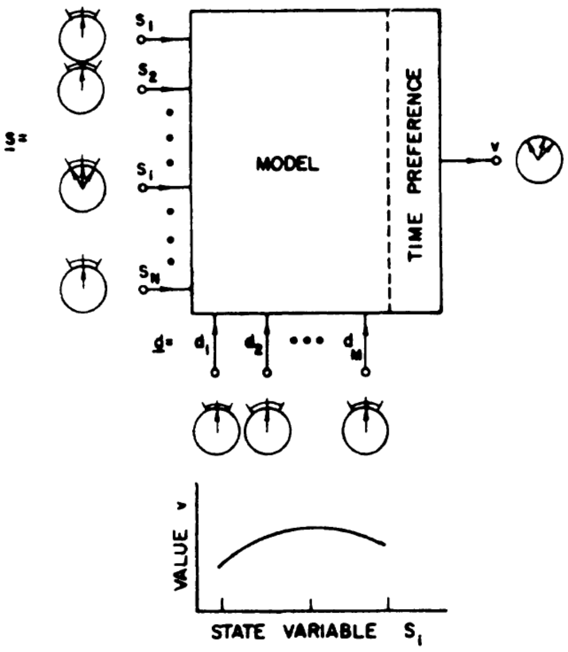


Fig. 5. Deterministic sensitivity.

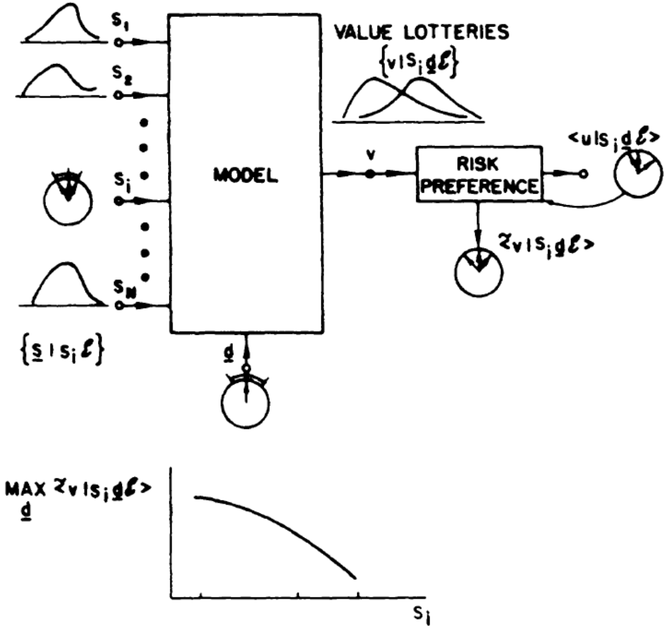


Fig. 7. Stochastic sensitivity.

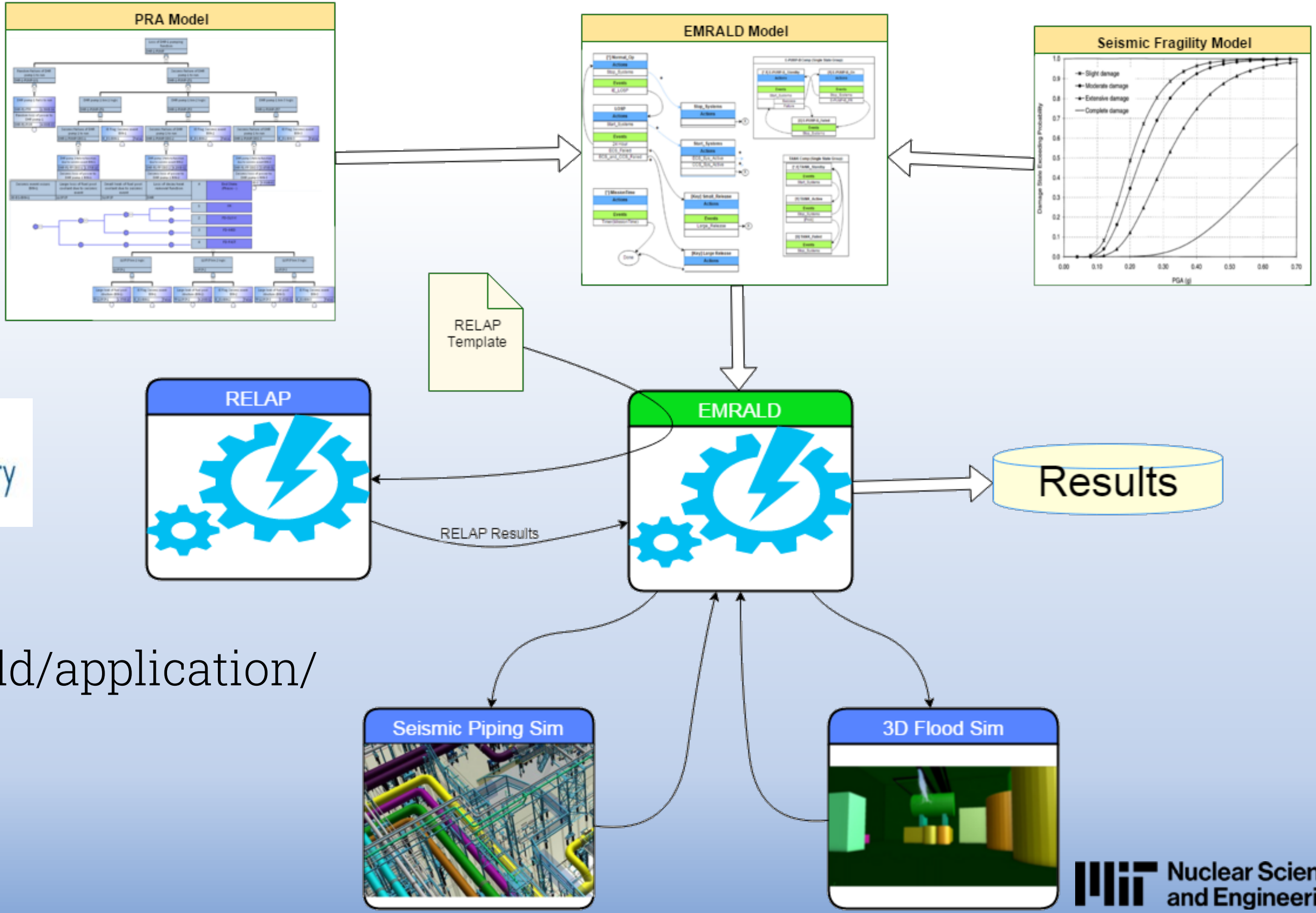
Models

Probabilistic Risk Assessment	Decision Trees/Influence Diagrams	Discrete Event Simulators	Agent Based Simulators
System Dynamics	Epidemiological Models	Hybrid Simulators	Digital Twins

Models of Phenomena

Friggs and Hartmann, 2020

A Complex Simulator Prototype



Idaho National Laboratory

https://inl.gov/emrald/application/

Università Bocconi MILANO

Nuclear Science and Engineering

The Setup

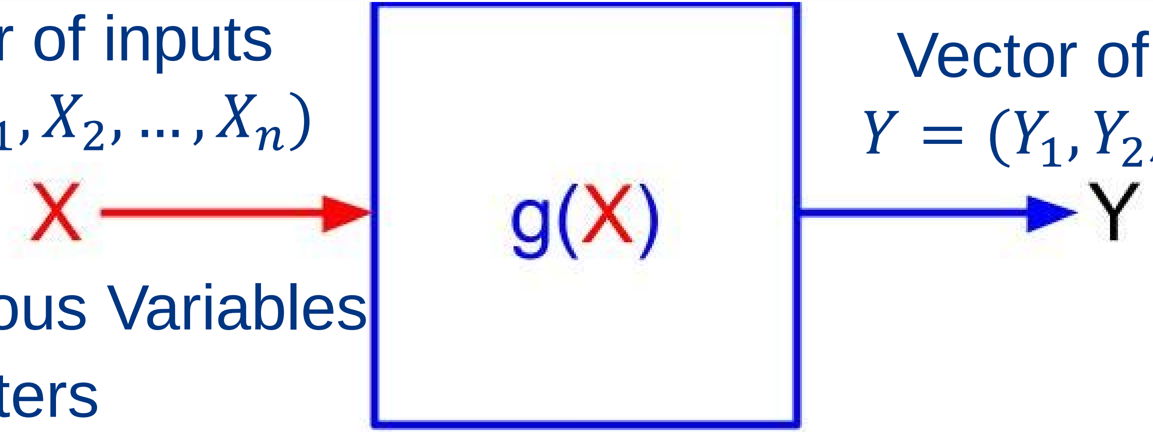
Input-output mapping

Vector of inputs

$\boldsymbol{X} = (X_1, X_2, \ldots, X_n)$

Vector of Outputs

$Y = (Y_1, Y_2, \ldots, Y_n)$



Exogenous Variables

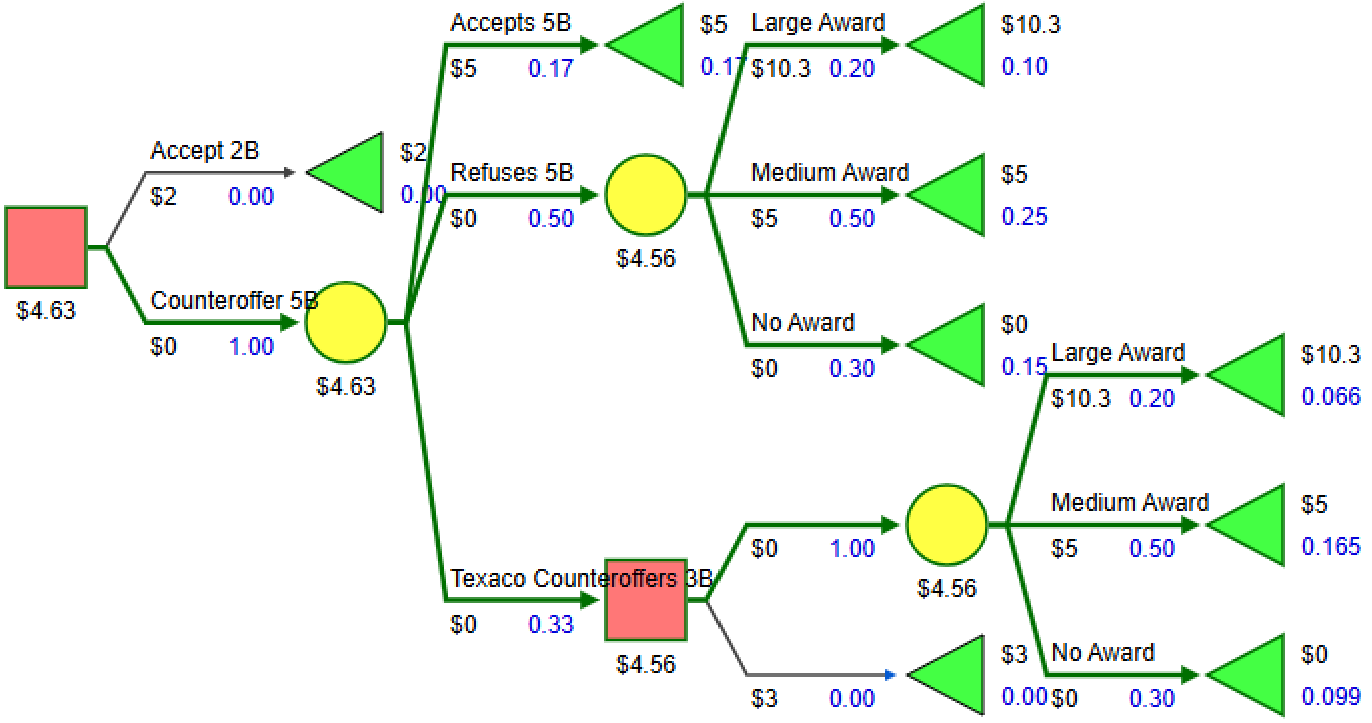
Parameters

Features in Machine Learning

Pennzoil vs Texaco

- 1984: Pennzoil and Getty Oil agreed to merge, but Getty backed out after Texaco offered a higher price.
- Pennzoil sued Texaco for interference and won an $11.1B judgment (later reduced to $2B, but interest pushed it to $10.3B).
- Texaco's CEO warned of bankruptcy if Pennzoil pursued liens and vowed to appeal up to the U.S. Supreme Court.
- In 1987, Texaco offered Pennzoil $2B to settle before liens were filed.
- Pennzoil's chairman faced a decision: accept $2B or counter with $5B as advised.

A toy Model: Pennzoil



SPECIFIC VS GENERAL GOALS

Specific Goals

- They answer a tailored sensitivity analysis question
- The answer can be problem and model-dependent
- E.g., Sensitivity of temperature on carbon emissions
- Methods can be ad-hoc

General Goals

- Broad questions, with broad classes of methods

Variable Importance

- Quantify predictive contribution: On which variable does the model rely?
- Fairness: Is the model relying on variables that can create implicit bias and discrimination?
- Model Improvement: What are the areas in which further modelling efforts are needed?
- Eschenbach 1992: Factors on which to focus managerial attention during implementation
- Various Stakeholders: Inputs to prioritize in practice

Direction of Impact

- How does the output of a model respond on the input variation?
- Is the variation causing an increase or decrease?
- Is the model monotonic, convex or periodic in the input?

Stability

- How uncertain are the model forecasts?
- Is the preferred alternative robust to the input variations?

Tornado Diagram Setup

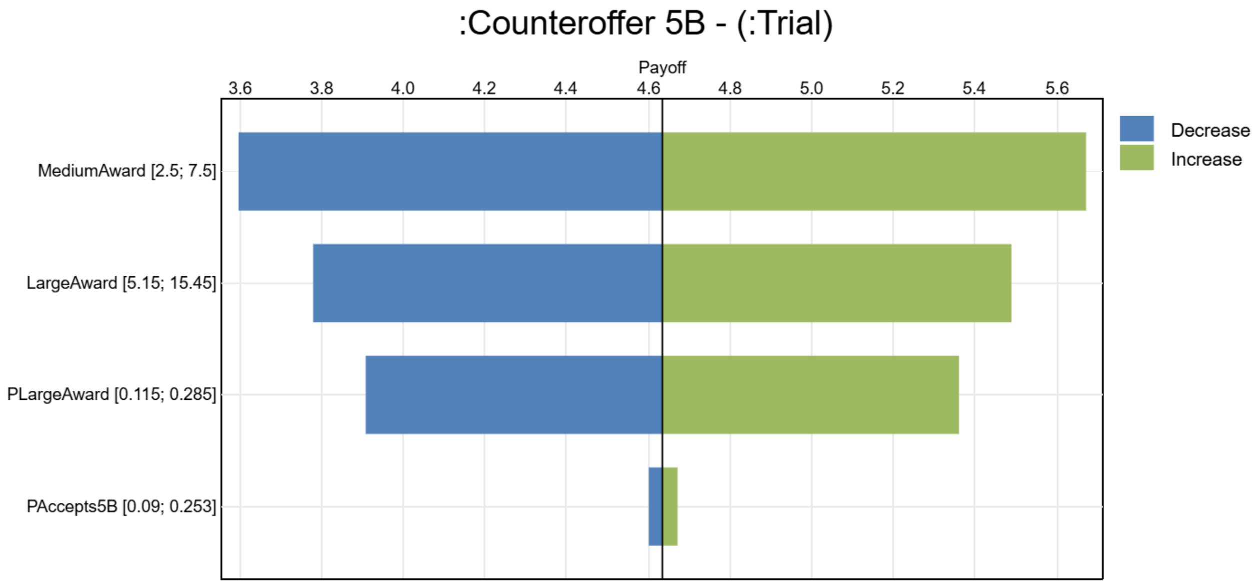
Tornado diagram

Variables

Name	Min	Max	Length	Default value
PLargeAward	0.115	0.285	2	0.2
PAccepts5B	0.09	0.253	2	0.17
LargeAward	5.15	15.45	2	10.3
MediumAward	2.5	7.5	2	5

Pennzoil Tornado Diagram

Sensitivity analysis

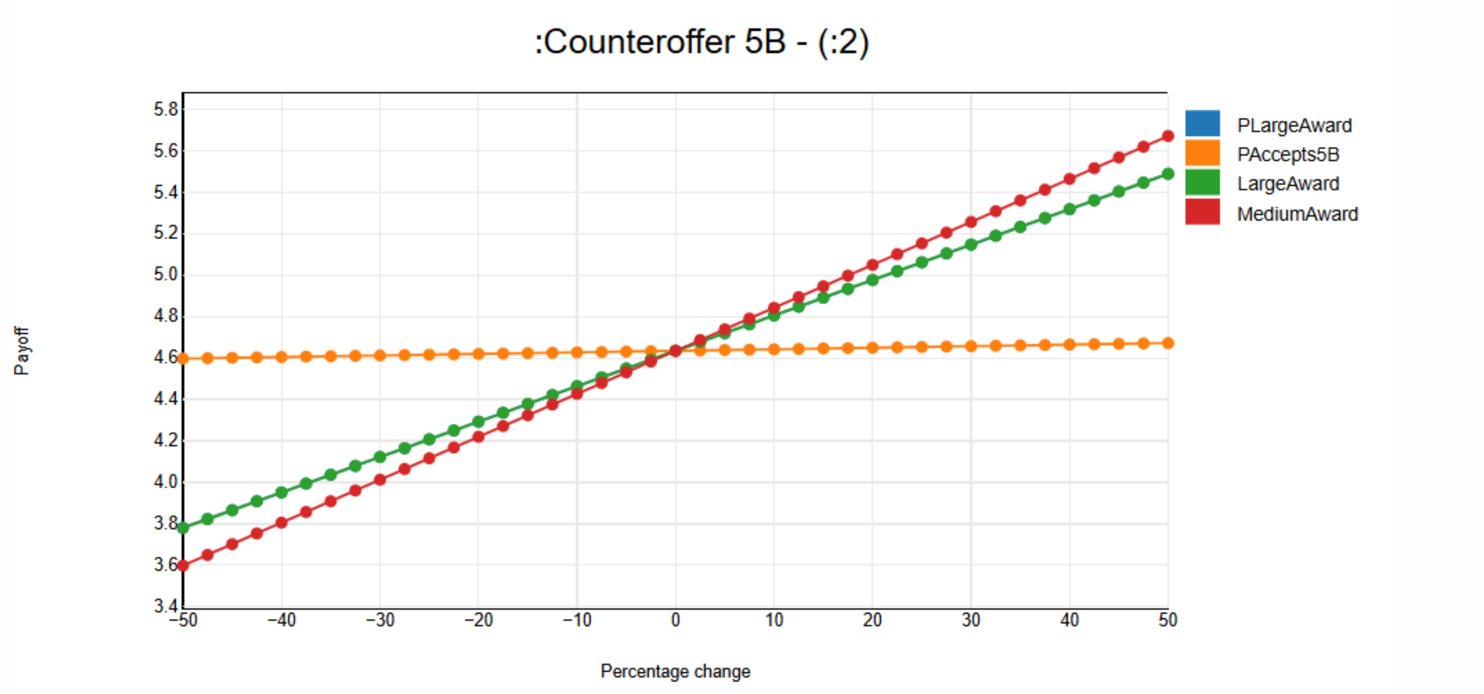


Howard, 1988, MS

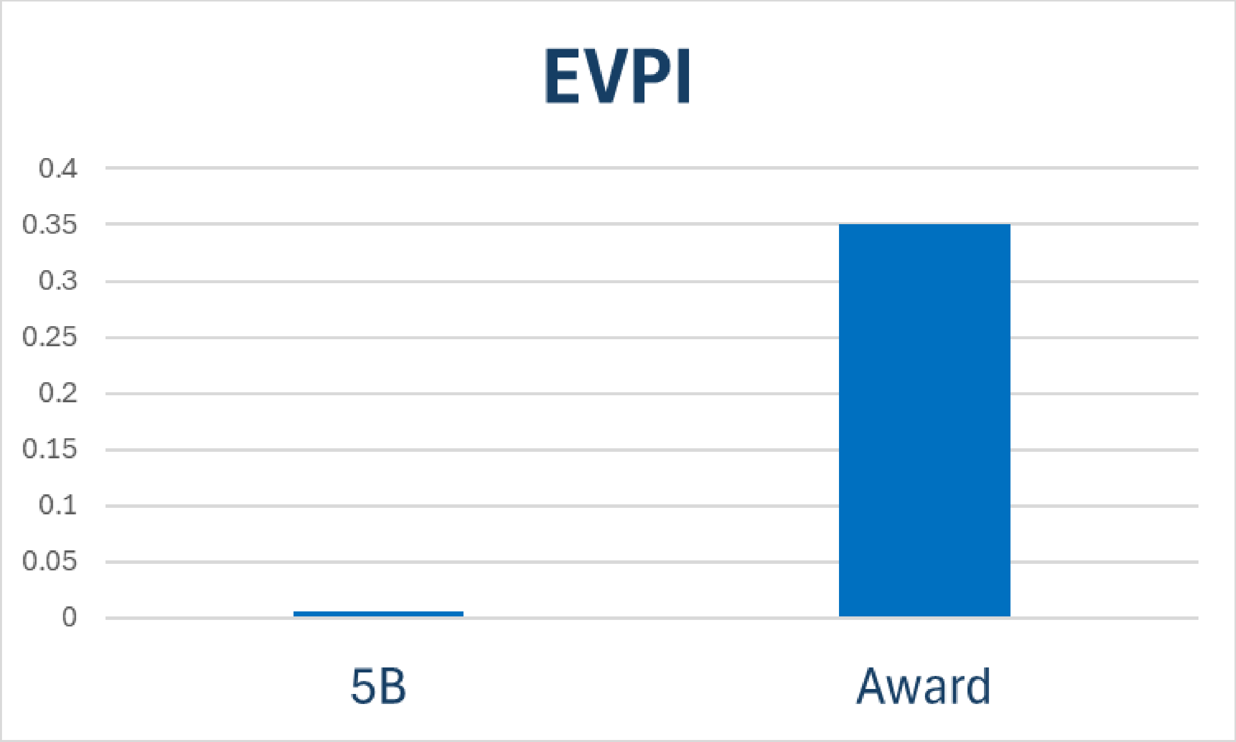
Spiderplot

Eschenbach 1992, Interfaces

Sensitivity analysis



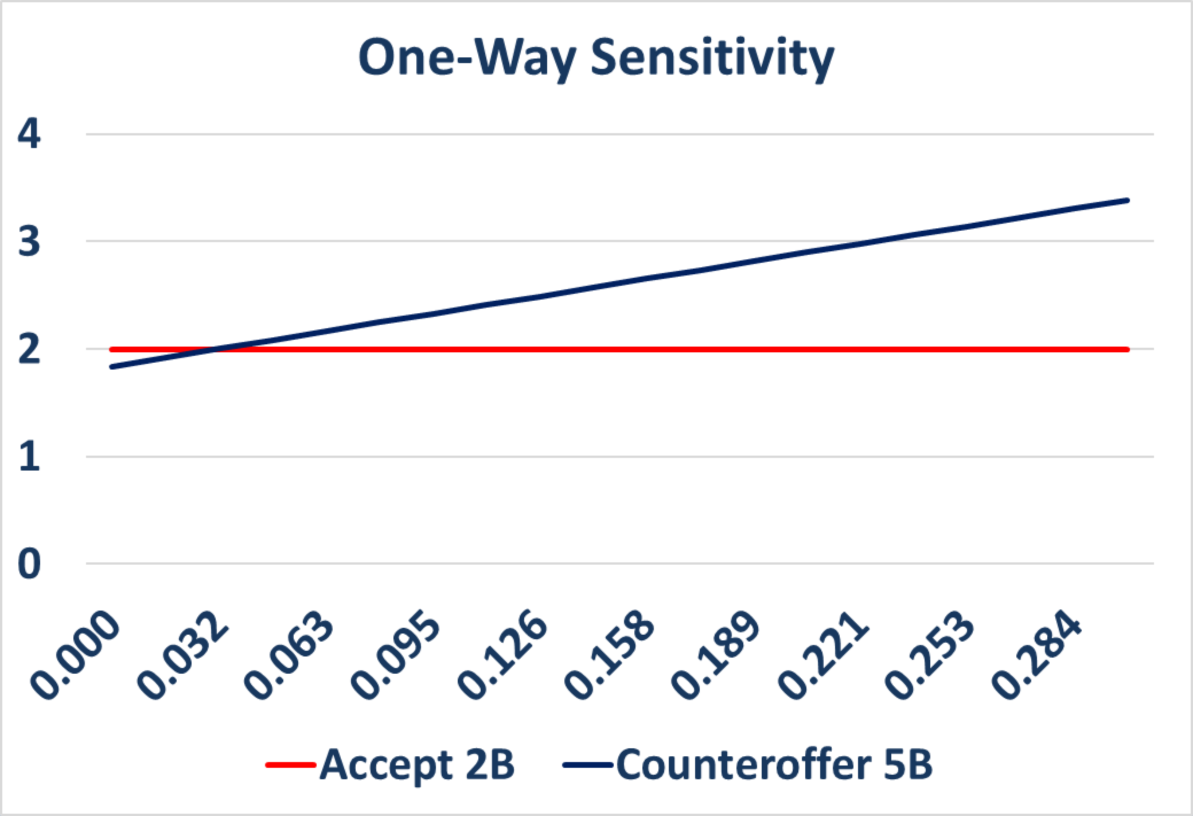
Expected Value of Perfect Information



Value vs Decision Sensitivity (Hazen and Felli, 1998, MDM)

Howard, 1966, Raiffa-Schlaifer, 1961

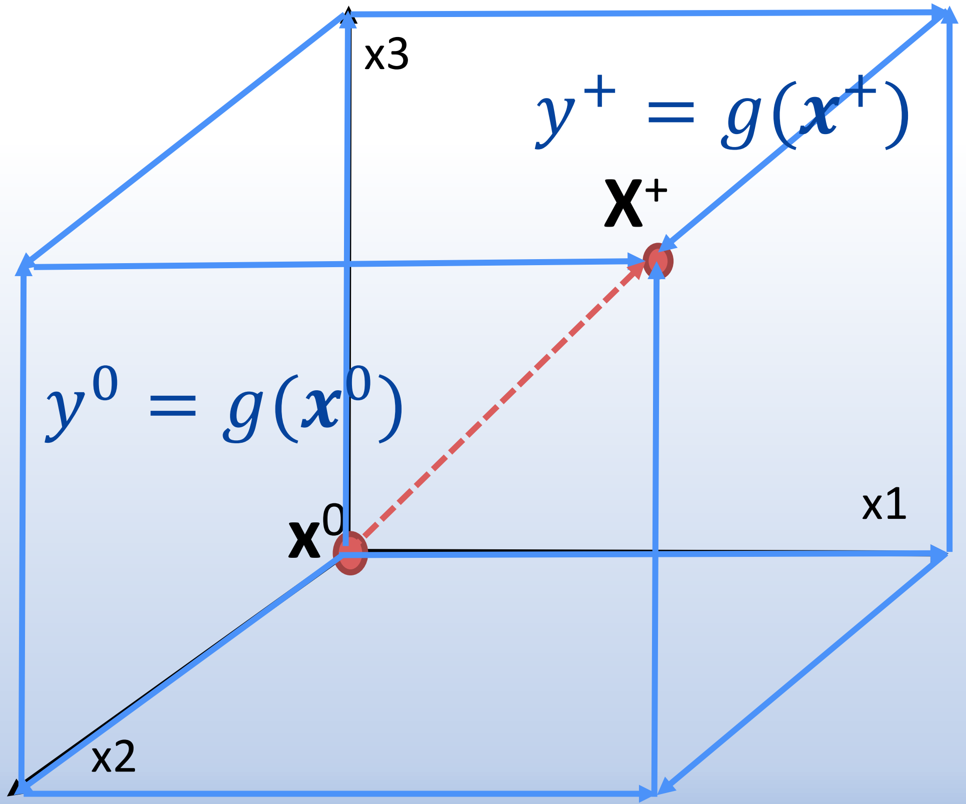
Pennzoil Stability on Pr(Accepts5B)



Structural Discovery

- Is the model response the sum of the responses to the individual input variations or interactions emerge?
- How relevant are interactions?
- Are interactions synergistic (complementarity) or antagonistic (substitutability)?

Deterministic Sensitivity



Typically, we define ranges, $[x_i^0, x_i^+]$

Model output change:

$$\Delta y = g(\mathbf{x}^+) - g(\mathbf{x}^0)$$

Model output change varying only x_i

$$\Delta_i y = g(x_i^+, \mathbf{x}_{-i}^0) - g(\mathbf{x}^0)$$

Bar of Tornado Diagram

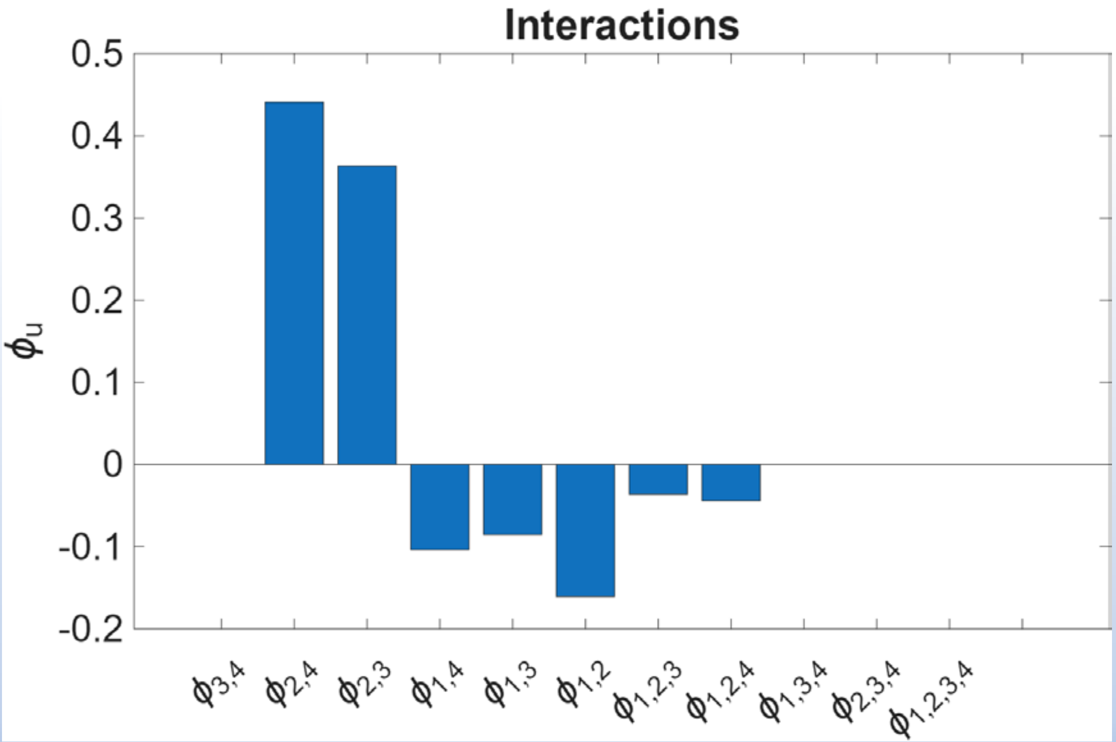
Finite-Difference Interactions

- We have: $$\Delta y = \sum_{i=1}^{n} \phi_i + \sum_{i<j} \phi_{i,j} + \ldots + \phi_{1,2,\ldots,n}$$

- Where:

$$\begin{cases} \phi_i = g(x_i^+; \mathbf{x}_{\sim i}^0) - g(\mathbf{x}^0) \longleftarrow \text{First Order} \\ \phi_{i,j} = g(x_i^+, x_j^+; \mathbf{x}_{\sim i,j}^0) - \phi_i - \phi_j - g(\mathbf{x}^0) \longleftarrow \text{Second Order} \\ \phi_{i,j,k} = g(x_i^+, x_j^+, x_k^+; \mathbf{x}_{\sim i,j,k}^0) - \phi_{i,j} - \phi_{i,k} - \phi_{j,k} - \phi_i - \phi_j - \phi_k - g(\mathbf{x}^0) \\ \ldots \end{cases}$$

Pennzoil Interactions

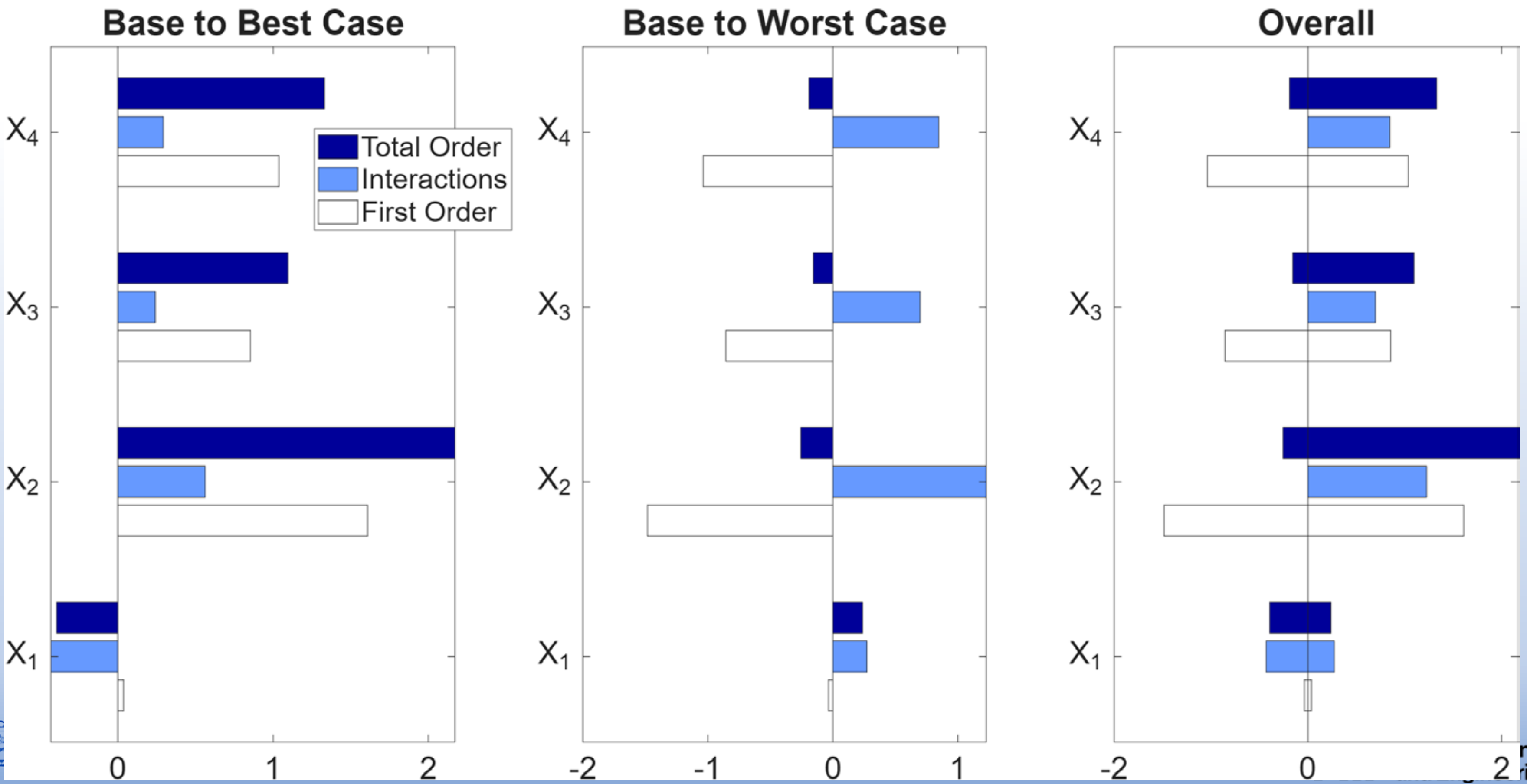


Generalized Tornado Diagrams

- Include Information on the Relevance of Interactions for each input
- Individual Effect: ϕ_i
- Total Effect: $\phi_i^T = \sum_{u:i\in u} \phi_u$
- Interaction Effect: $\phi_i^I = \phi_i^T - \phi_i$
- Computational saving result by B. and Smith (2011, OR) abates the cost from 2^d to $2(d+1)$.
- Just double the cost of a Tornado

Università Bocconi
MILANO

Generalized Tornado Pennzoil



PROBABILISTIC SENSITIVITY

Monte Carlo Simulation

- Assign Inputs Probability Distribution
  - Delicate Step
- Propagate uncertainty through the model
  - Typically via Monte Carlo or Quasi Monte Carlo
- Evaluate Results
- Post-process Results

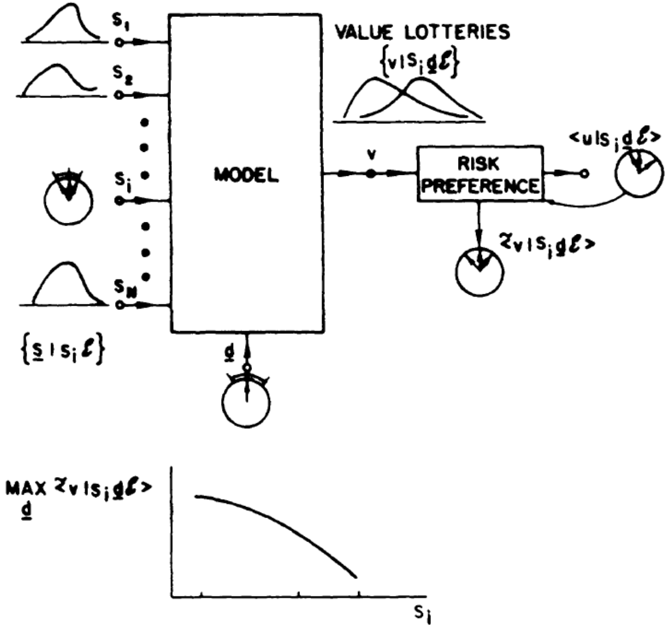
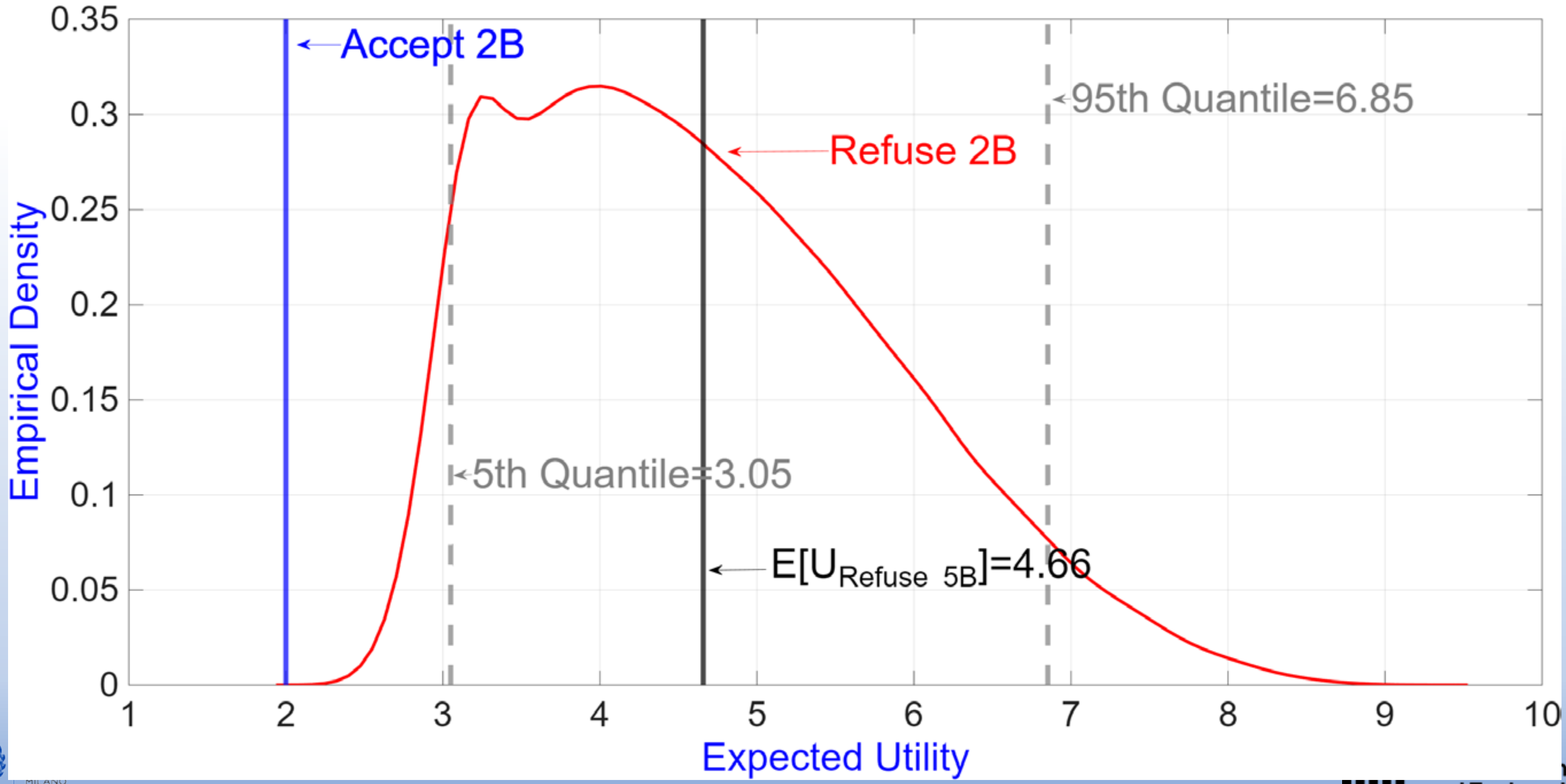


Fig. 7. Stochastic sensitivity.

Probabilistic Sensitivity for Pennzoil



Insights

- Information on the expected value, confidence intervals
- Given the assigned ranges, the preferred alternative is “stable” (Stochastic Dominance)

Global Importance

- When measuring importance of variables in a probabilistic sensitivity analysis setting, we can resort to several alternatives.

- One important class: measures of statistical association

A Common Rationale

- Several importance measures can be written in the following form:

$$\xi_i = \mathbb{E}_i[d(\mathbb{P}_Y, \mathbb{P}_{Y|X_i})]$$

with d() a generic operator between the marginal and conditional distributions of Y.

- The above equation encompasses variance-based, kernel-based, scoring function-based global sensitivity measures, as well as value of information (B., Hazen et al 2016, RA).

Measures and Properties

- First order variance-Based sensitivity measures η_i^2 do not possess the zero-independence property.

$$\eta^2(Y,X) = \frac{\mathbb{V}[\mathbb{E}[Y|X]]}{\mathbb{V}[Y]}$$

The following do possess zero-independence (B. 2017, book) :

$$\delta_i^{KL} = \mathbb{E}[\int_{\mathcal{Y}} f_Y(y) \ln \frac{f_Y(y)}{f_{Y|X_i}(y)} dy]$$

based on the Kullback-Leibler divergence (Rahman, 2016, SIAM/ASA-JUQ)

$$\delta_i = \frac{1}{2}\mathbb{E}_i\left[\int_{\mathcal{Y}} | f_Y(y) - f_{Y|X_i}(y;x_i) | dy\right]$$

based on the L1-norm between densities (B., 2001, RESS)

$$\beta_i^{Ku} = \mathbb{E}_i[\int_{\mathcal{Y}} \left(F_Y(y) - F_{Y|X_i}(y)\right)^2 dF_Y(y)]$$

Kuiper-based (Baucells and B., 2013, MS)

$$\beta_i^{CvM} = \mathbb{E}_i[\int_{\mathcal{Y}} \left(F_Y(y) - F_{Y|X_i}(y)\right)^2 dF_Y(y)]$$

Chatterjee's new correlation coefficient (Chatterjee, 2021, JASA)

Desirable Properties

- Zero-Independence
  - Renyi (1959) postulate D: a measure of statistical dependence between Y and X should be equal to zero if and only if Y is independent of X.

- Max-Functionality
  - Renyi (1959) postulate E: The sensitivity measure should be maximal if Y=g(X)

New Indices based on Optimal Transport Theory

$$i^{W_2^2}(\mathbf{Y}, X_i) := \frac{\mathbb{E}\left[W_2^2(\mathbb{P}_{\mathbf{Y}}, \mathbb{P}_{\mathbf{Y}|X_i})\right]}{2\mathbb{V}[\%\mathbf{Y}]}$$

where W2 is the 2-squared Wasserstein distance [Wiesel, 2022, B. et al., 2025, MS].

$$0 \leq i^{W_2^2}(\boldsymbol{Y}, X_i) \leq 1$$

The index is normalized between 0 and 1, with 0 indicating independence and 1 indicating full functional dependence.

Moreover

Residual error term

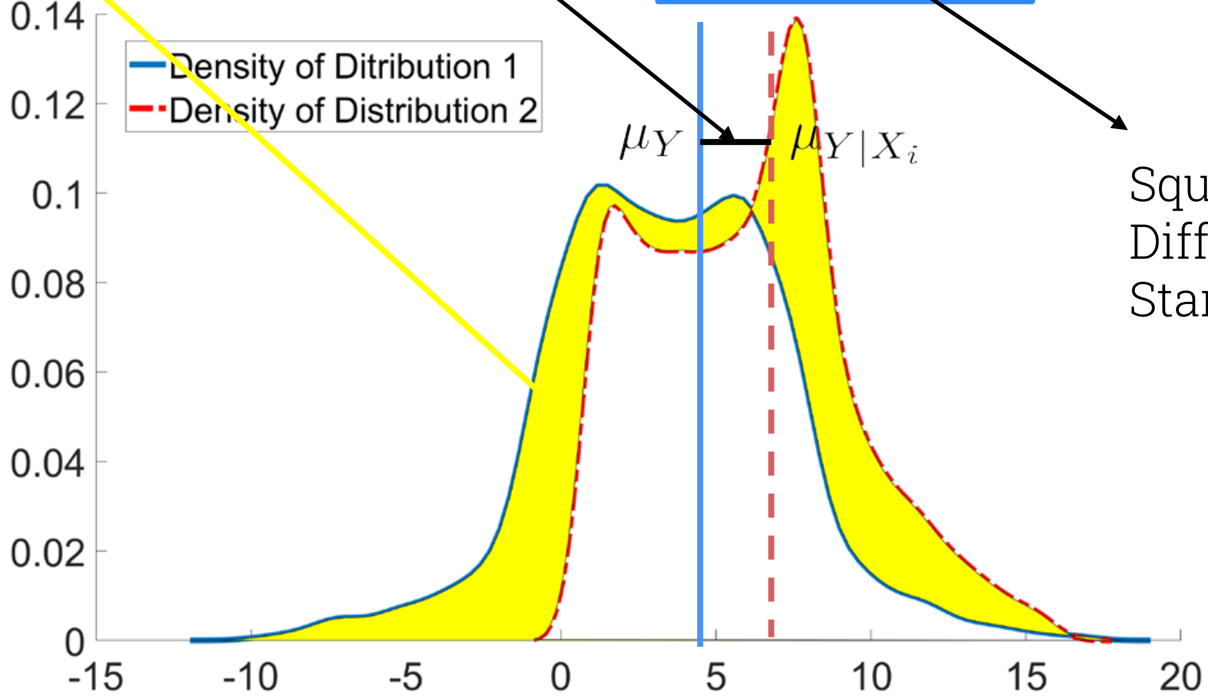
$$i^{W_2^2}(\mathbf{Y}, X_i) = i^{VB}(\mathbf{Y}, X_i) + i^{\Sigma}(\mathbf{Y}, X_i) + i^{\Gamma}(\mathbf{Y}, X_i).$$

Variance-Based

Impact on second moments

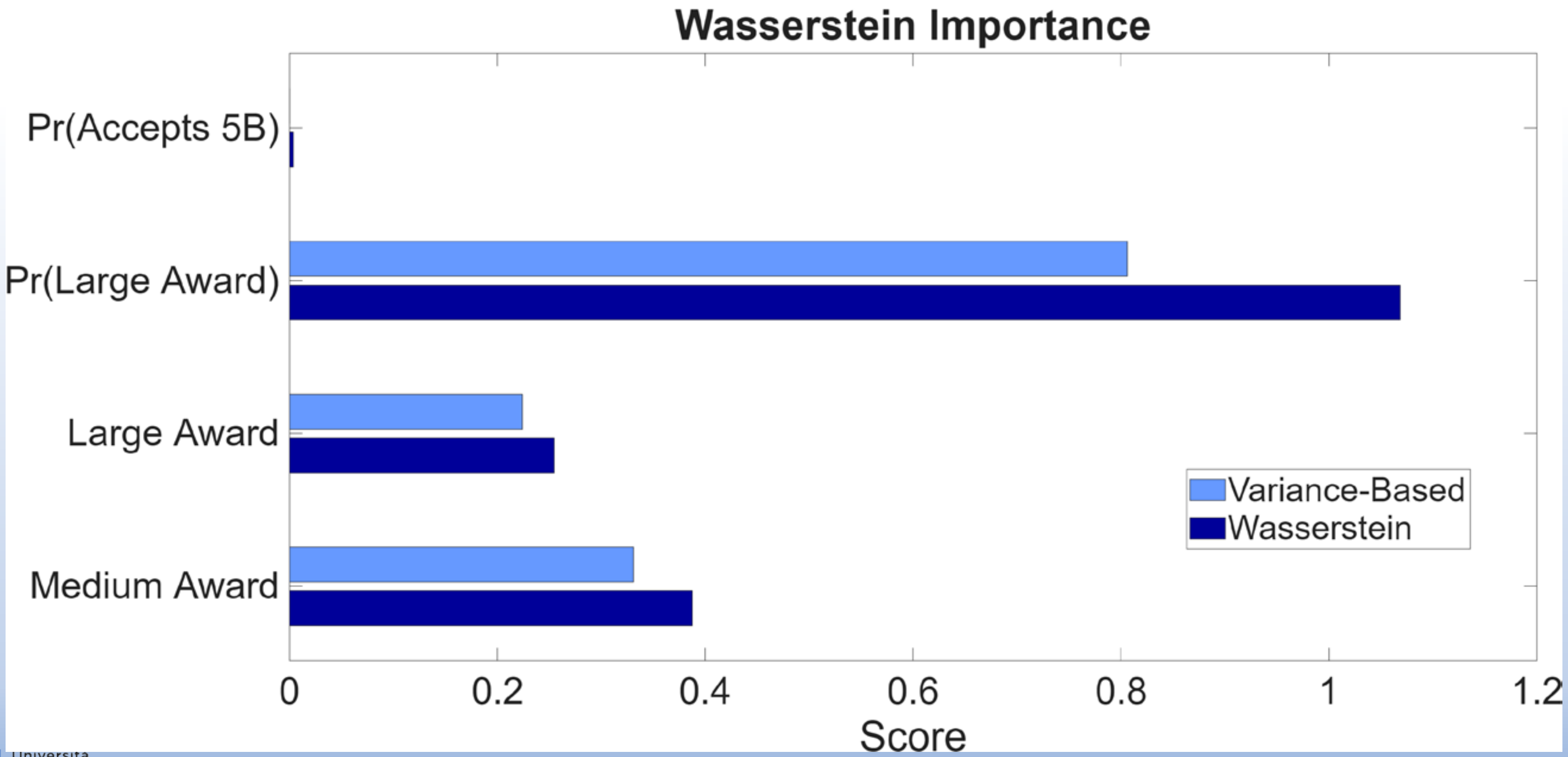
Interpretation (cont).

$$\iota^{W_2^2}(Y, X_i) = \boxed{\frac{\mathbb{E}\left[\left(\mu_Y - \mu_{YX_i}\right)^2\right]}{2\mathbb{V}[\mathbf{Y}]}} + \boxed{\frac{\mathbb{E}\left[\left(\sigma_Y - \sigma_{YX_i}\right)^2\right]}{2\mathbb{V}[\mathbf{Y}]}} + \iota^{\Gamma}(Y, X_i)$$



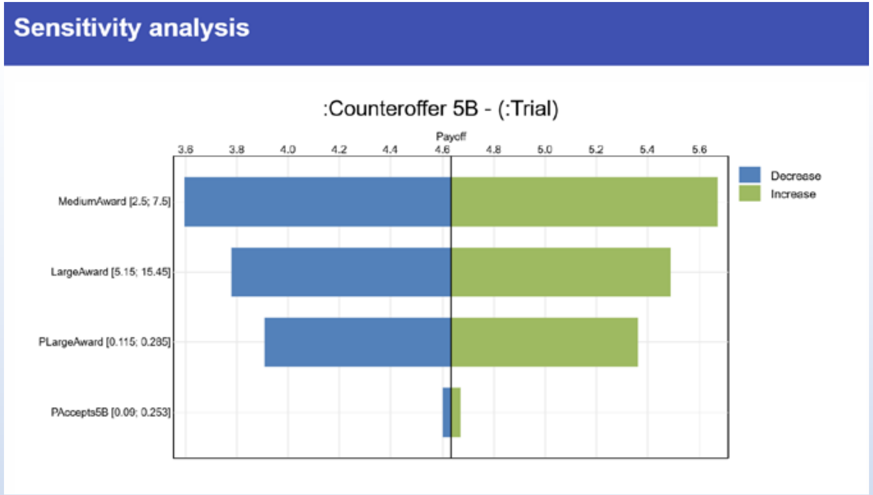
Square of the Difference in Standard deviations

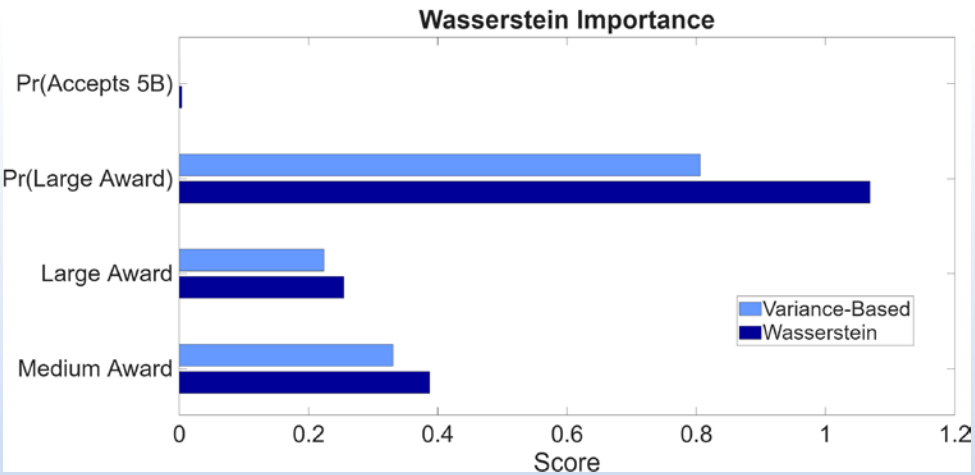
Pennzoil Global Sensitivity



Tornado Diagrams

Global Sensitivity Indices





Deterministic

Probabilistic

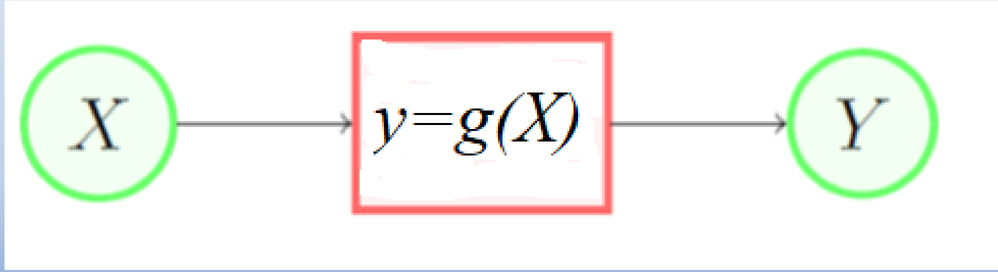
Ongoing Applications

- Multivariate output large climate models (emission patterns (Chiani et al., 2025, RA)
- NASA Mars Return Program, in cooperation with NASA and Jet Propulsion Lab (Cataldo et al, 2025, RA)
- Energy Systems: in cooperation with University of North Carolina and Politecnico of Turn, (Nicoli et al., Energy, 2025)
- Modeling of Small Modular Reactors with Politecnico of Milan (Marchetti et al., Nuclear Engineering and Technology, 2025)

MACHINE LEARNING SETUP: MODELS OF DATA

Models of Data

Machine Learning tools try to quantify the relationship under "nature" creating an input-output mapping:



The Machine Learning Problem

- Empirical Machine Learning Problem:

$$\min_{\theta\in\Theta} \mathbb{E}[\mathcal{L}(Y, \hat{g}(\mathbf{X};\theta))]$$

- Where $\mathcal{L}(\cdot,\cdot)$ is a loss function, $\widehat{g}(\mathbf{X},\theta)$ a family of parameterized machine learning models

Models

Generalized Linear Models

Regression Trees

Neural Networks

Deep Learning

Support Vector Machines

Random Forests

XGBoost

Gaussian Process Regression

Self Orienting Maps

Models of Data

Friggs and Hartmann, 2020

The Black Box Menace

- (Begoli, Bhattacharya, and Kusnezov 2019, p. 21) underline the issue that the "Absence of Theory" makes a proper uncertainty quantification an essential ingredient in the use of black-box machine learning algorithms.
- (Guidotti et al., 2018) suggests that interpretability and explainability issues impact the use of artificial intelligent methods.
- (Saltelli, 2019) underlines that modelling quality and validity is at stake, especially as model complexity increases

Against the Black-Box Menace, we need Interpretability and Explainability

Interpretability (I)

- (Miller, 2019) offers an insightful review on artificial intelligence explanation, without a strong distinction between intepretability and explainability
- Conversely, Rudin (2019) starts marking the distinction sharply
- (Guidotti et al., 2018): A survey of methods to explain black box models
- (Murdoch et al., 2019): Interpretable Machine Learning
- (Bertsimas & Kallus, 2020) : From predictive to Prescriptive Analytics
- (Rudin et al., 2022): 10 Grand Challenges in Interpretable ML
- (De Boeck et al, 2023): Explainable AI in OR

The Prejudice

	Model-based interpretability	Post hoc interpretability
Predictive Accuracy	Generally unchanged or decrease (data-dependent)	No Effect
Descriptive Accuracy	Increase	Increase

We believe that greater accuracy implies more complex models: this is **NOT** necessarily true

Explainability

- (Guidotti et al., 2018), Saltelli (2019), suggest post-hoc explanations as essential to:
- Increase transparency
- Enhance interpretation and communication
- Increase awareness of the model behavior
- Obtain Crucial Insights

All goals remain valid

- Variable importance
- Direction of Impact
  - ALE Plots, Partial Dependence Plots, developed in machine learning
- Structural Discovery
- Stability Analysis

Type of “Explanations”

- Local explanations
  - Prediction by prediction
- Dataset level explanations
  - One Single Number
- Data generating process
- Model predictions

Some Local Explanation Methods

- LIME (Ribeiro et al, 2016)
- Shapley-value based:
  - SHAPs (Lundberg and Lee, 2017)
  - Baseline Shapley Values (Sundararajan, 2020, B. and Rabitti, 2023)
  - Cohort Shapley Values (Mase and Owen, 2022)

Permutation Feature Importance

- Breiman's feature importance of X_j is defined as:

$$\widehat{\nu}_j = \underbrace{\frac{1}{N}\sum_{n=1}^{N}\mathcal{L}\left(\mathbf{y}^n, \widehat{g}(\mathbf{x}^n_{j,\pi};\theta^*)\right)}_{\text{Loss after } X_j \text{ permuted}} - \underbrace{\frac{1}{N}\sum_{n=1}^{N}\mathcal{L}\left(\mathbf{y}^n, \widehat{g}(\mathbf{x}^n;\theta^*)\right)}_{\text{Original Loss}}$$

Total Indices

- Total indices (Homma and Saltelli, 1996):

$$\tau_j = \mathbb{V}_{-j}[\mathbb{E}_j[Y|\mathbf{X}_{-j}]] = \mathbb{V}[Y] - \mathbb{E}_{-j}[\mathbb{V}_j[Y|\mathbf{X}_{-j}]],$$

- Expected residual variance after we fix all features by X_j
- Jansen (1999) shows that they can be written as

$$\tau_j = \frac{1}{2}\left(\mathbb{E}\left[\left(g(X'_j, \mathbf{X}_{-j}) - g\left(\mathbf{X}\right)\right)^2\right]\right),$$

- Which is equivalent to:

Within Distribution

$$\tau_j = \frac{1}{2}\int_{\mathcal{X}}\int_{\mathcal{X}_j}\left(g\left(x'_j, \mathbf{x}_{-j}\right) - g\left(\mathbf{x}\right)\right)^2 dF_{X_j|\mathbf{X}_{-j}}(x'_j|\mathbf{x}_{-j})dF_{\mathbf{X}}(\mathbf{x}).$$

- If we ignore the correlation between X_j and $\mathbf{X}_{-j}$, we get the Verdinelli and Wasserman index

Out of Distribution

$$\tau'_j = \frac{1}{2}\int_{\mathcal{X}}\int_{\mathcal{X}_j}\left(g\left(x'_j, \mathbf{x}_{-j}\right) - g\left(\mathbf{x}\right)\right)^2 dF_{X_j}(x'_j)dF_{\mathbf{X}}(\mathbf{x}).$$

Proposition

If $\mathcal{L}(\cdot,\cdot)$ is the quadratic loss and if the model is a perfect predictor, then

1)

$$\nu_j = \mathbb{E}\left[\left(\widehat{g}(\mathbf{X};\theta^*) - \widehat{g}(X_j^{'},\mathbf{X}_{-j};\theta^*)\right)^2\right]. \tag{1}$$

2) If $X_j^{'}$ is sampled independently of $\mathbf{X}_{-j}$, then

$$\nu_j = 2\tau_j^{'}, \quad \text{Non Classical total index} \tag{2}$$

where $\tau_j^{'}$ is the Ψ_{DLoco} total index in Verdinelli and Wassermann (2023).

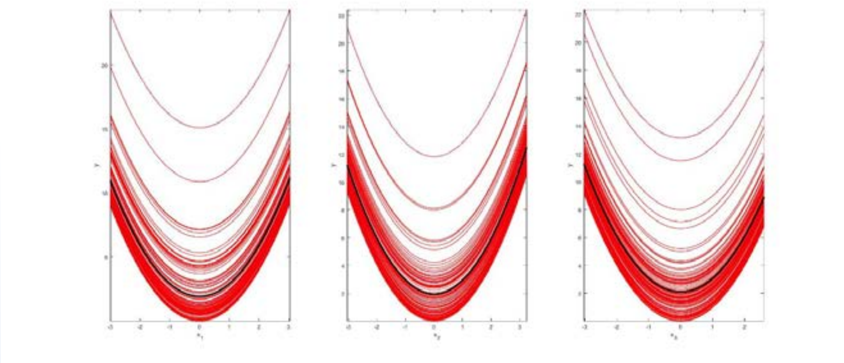
3) If $X_j^{'}$ is sampled conditionally on $\mathbf{X}_{-j}$, then

$$\nu_j = 2\tau_j, \quad \text{Classical total index} \tag{3}$$

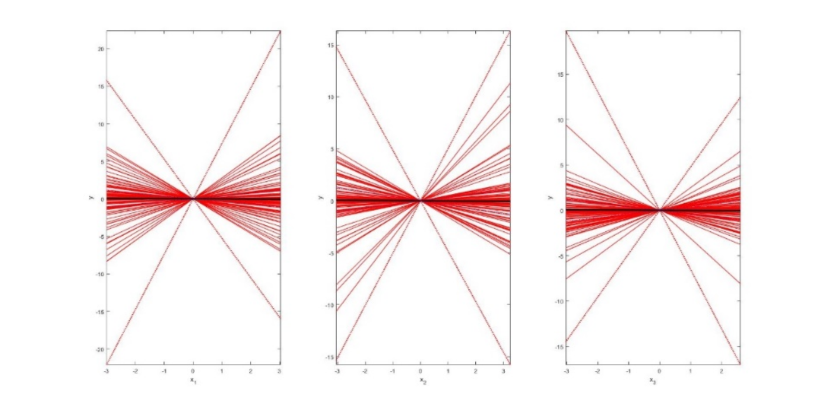
where τ_j is the classical total index

Direction of Change

DOES Y BEHAVE LIKE THIS?



OR LIKE THIS?



Global trend indicators

Conditional Regression Curves (Wooldridge 2013)

$$r_i(x_i) = \mathbb{E}[g(\mathbf{X})|X_i = x_i] = \int_{\mathcal{X}_{\sim z}} g((\mathbf{x}_{-i}; x_i))dF_{\mathbf{X}|X_i}(\mathbf{x}_{-i}; x_i)$$

Partial dependence functions (Friedman 2001)

$$h_i(x_i) = \int_{\mathcal{X}_{\sim i}} g(x_i; \mathbf{x}_{\sim i})dF_{\mathbf{X}_{\sim i}}(\mathbf{x}_{\sim i})$$

Global trend indicators (II)

ALE plots (Apley and Zhu, 2021, JRSSB)

$$\text{ALE}_i(x_i) = \int_{x_{i,\min}}^{x_i} \mathbb{E}[g_i'(\mathbf{X})|X_i = z_i]dz_i$$

Consistent Indicators

- Determine whether the indicators correctly report properties of the original input-output mapping
- Monotonicity Consistency
- Convexity Consistency
- Lipschitz consistency

A Summary

Indicator	Symbol	Monotonicity consistency	Lipschitz consistency	Concavity consistency	Discrete inputs	Handles distributions
Gradients/Hessians	$g_i'(\mathbf{X})$	Yes	Yes	N/A	Yes	No
Tornado	$\Delta g_i^+, \Delta g_i^-$	Yes	Yes	N/A	Yes	No
One-way function	$g(x_i; \mathbf{x}_{-i}^0)$	Yes	Yes	Yes	Yes	No
Correlation coefficient	ρ_{Y,X_i}	Independent	N/A	N/A	Yes	Yes
Conditional expectation	$r_i(x_i)$	Independent	No	Independent	Yes	Yes
Partial dependence function	$\mathrm{PD}_i(x_i)$	Yes	Yes	Yes	Yes	Yes
ALE function	$\mathrm{ALE}_i(x_i)$	Independent	No	Independent	Yes	Yes

From B., Baucells et al., 2025, Risk Analysis

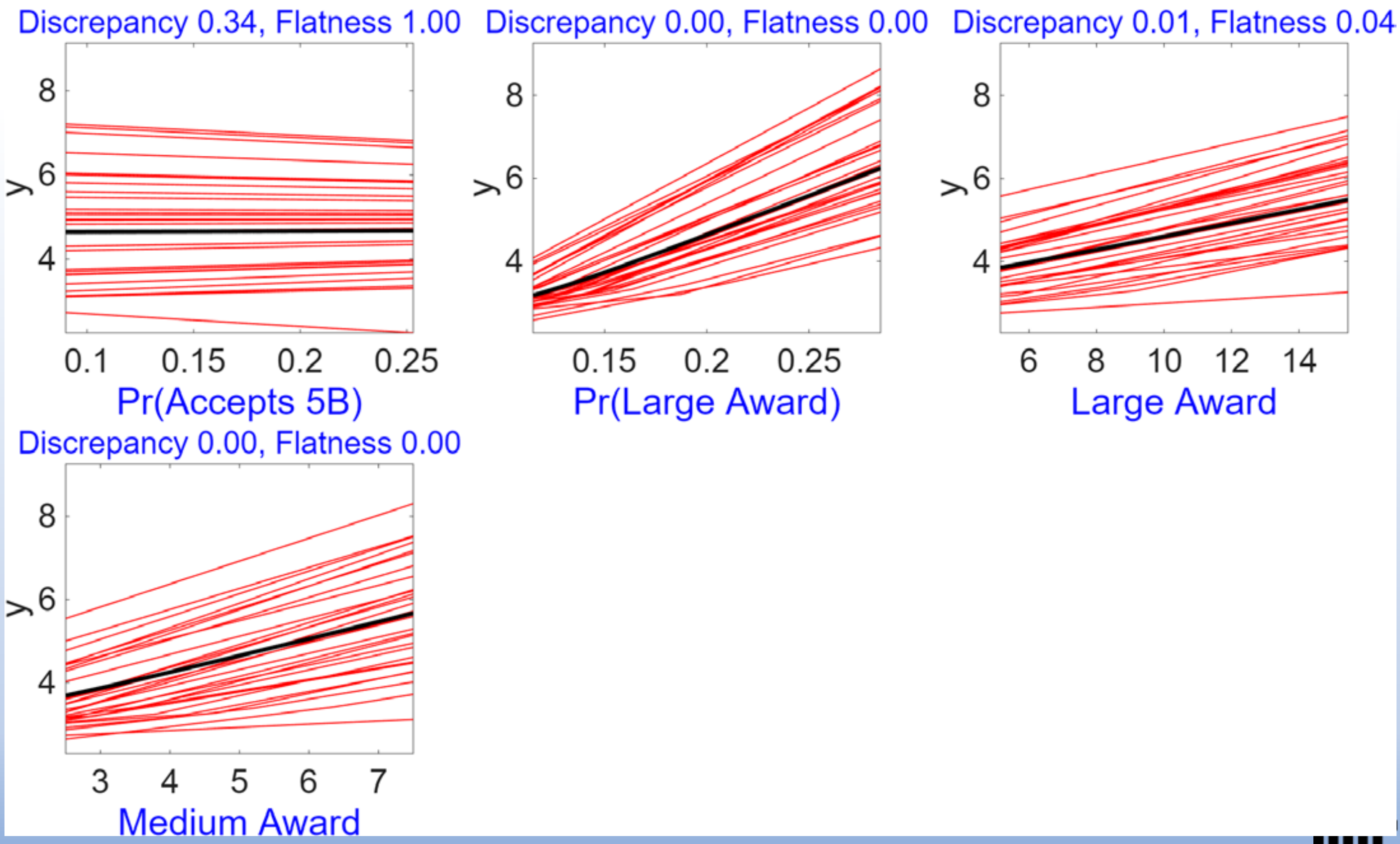
PD-One-way Plots

- ICE plots overlap “individual conditional expectations” and partial dependence functions
- It turns out that partial dependence functions are averages of one-way sensitivity functions of spiderplots
- … and individual conditional expectations are one-way sensitivity functiond

Two indices

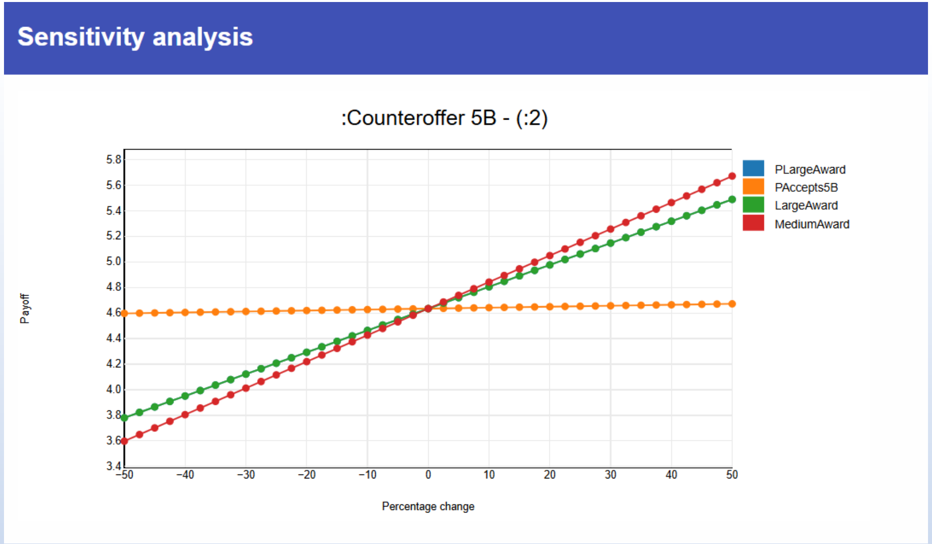
- Flatness Index: A partial dependence function can be flat even if Y depends on X_i
- Discrepancy index: Does the average hide the individual behavior in each of the one-way sensitivity functions?
- Insights on additivity

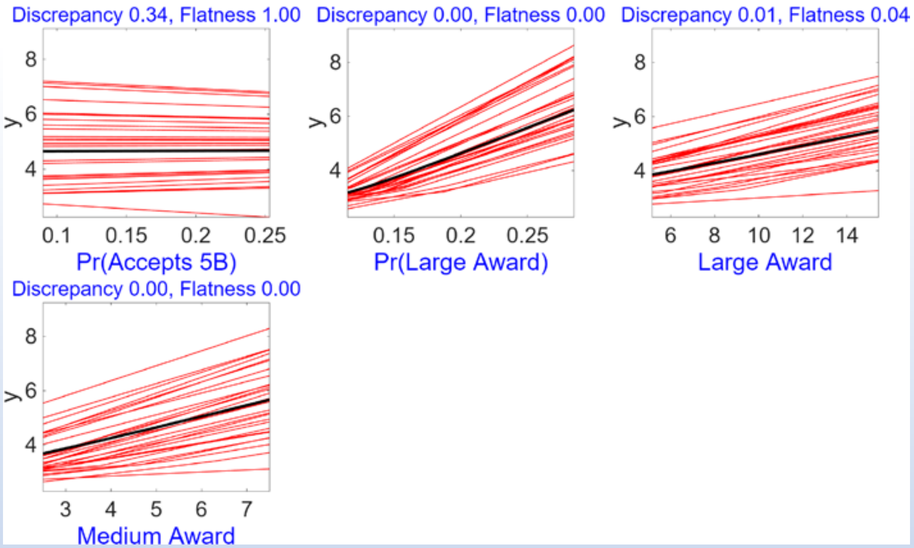
PD-One-way Plots for Pennzoil



Spiderplots

PD-Oneway Plots





Deterministic

Probabilistic

Some Quotes

- The judicious application of sensitivity analysis techniques appears to be the key ingredient needed to draw out the maximum capabilities of mathematical modeling (Rabitz, 1989), p. 221.

- Sensitivity Analysis for Modelers: Would you go to an orthopedist who did not use X-Ray? (Fuerbringer, 1996).

- In order for the analysis to be useful it must provide information concerning the way in which our equilibrium quantities will change as a result of changes in the parameters (Samuelson, 1941).

Bibliography

- Apley, D. W., & Zhu, J. (2020). Visualizing the Effects of Predictor Variables in Black Box Supervised Learning Models. Journal of the Royal Statistical Society. Series B: Statistical Methodology, 82(4), 1059–1086.
- Baucells, M., & Borgonovo, E. (2013). Invariant Probabilistic Sensitivity Analysis. Management Science, 59(11), 2536–2549.
- Begoli, E., Bhattacharya, T., & Kusnezov, D. (2019). The Need for Uncertainty Quantification in Machine-Assisted Medical Decision Making. Nature Machine Intelligence, 1(1), 20–23.
- Bertsimas, D., & Kallus, N. (2020). From Predictive to Prescriptive Analytics. Management Science, 66(3), 1025–1044.
- Borgonovo, E., & Plischke, E. (2016). Sensitivity Analysis: A Review of Recent Advances. European Journal of Operational Research, 3(1), 869–887.

Bibliograpy (ii)

- Borgonovo, E., V. R. R. Jose, M. Knowlton, R. Schachter, J. U. Siebert, and C. Ulu. 2025. “Fifty Years of Decision Analysis in Operational Research: A Review.” European Journal of Operational Research, April 2025.
- Breiman, L. (2001). Random forests. Machine Learning, 45(1), 5–32.
- Cataldo, R., et al. 2025. “Risk Analysis for NASA Mars Sample Return Program.” Risk Analysis forthcoming
- Chatterjee, S. (2021). A New Coefficient of Correlation. Journal of the American Statistical Association, 116(536), 2009–2022.
- Chiani, M., et al. 2025. “Global Sensitivity Analysis for Climate Models.” Risk Analysis, forthcoming.
- Clemen, R. T. (1997). Making Hard Decisions: An Introduction to Decision Analysis. Duxbury Press, CA, USA.

Bibliography (iii)

- de Bock, K. W., Coussement, K., et al. (2023). Explainable AI for Operational Research: A defining framework, methods, applications, and a research agenda. European Journal of Operational Research, 317(2), 249–272.
- Eschenbach, T. G. (1992). Spiderplots versus Tornado Diagrams for Sensitivity Analysis. Interfaces, 22(6), 40–46.
- Friedman, J. H. (2001). Greedy Function Approximation: a Gradient Boosting Machine. The Annals of Statistics, 29(5), 1189–1232.
- Fuerbringer, J. M. (1996). Sensitivity Analysis for Modellers. Air Infiltr. Rev, 17(4), 8–10.
- Guidotti, R., Monreale, A., et al. (2018). A Survey of Methods for Explaining Black Box Models. ACM Comp. Surveys, 51(5), 1–42.
- Felli, J. C., & Hazen, G. B. (1998). Sensitivity Analysis and the Expected Value of Perfect Information. Medical Decision Making, 18, 95–109.

Bibliography (iv)

- Homma, T., & Saltelli, A. (1996). Importance Measures in Global Sensitivity Analysis of Nonlinear Models. Reliability Engineering and System Safety, 52(1), 1–17.
- Howard, R. A. (1966). Information value theory. IEEE Transactions on Systems Science and Cybernetics, 2(1), 22–26.
- Howard, R. A. (1968). The foundations of decision analysis. IEEE Transactions on Systems Science and Cybernetics, 4(3), 211–219.
- Howard, R. A. (1988). Decision Analysis: Practice and Promise. Management Science, 34(6), 679–695.
- Jansen, M. J. W. (1999). Analysis of variance designs for model output. Computer Physics Communications, 117(1–2), 35–43.
- Lundberg, S. M., & Lee, S.-I. (2017). A unified approach to interpreting model predictions. Advances in Neural Information Processing Systems, 2017-Decem, 4766–4775.

Bibliography (v)

- Mase, M., Owen, A. B., & Selier, B. B. (2022). Variable Importance Without Impossible Data. ArXiv, 2205.(15750), 1–36.
- Miller, T. (2019). Explanation in Artificial Intelligence: Insights from the Social Sciences. Artificial Intelligence, 267, 1–38.
- Murdoch, W. J., Singh, C., Kumbier, K., Abbasi-Asl, R., & Yu, B. (2019). Definitions, Methods and Applications in Interpretable Machine Learning. Proceedings of the National Academy of Sciences, 116(44), 22071–22080.
- Raiffa, H., & Schlaifer, R. (1961). Applied Statistical Decision Theory. Harvard University Press.

Monograph

- Borgonovo, E. 2017. Sensitivity Analysis: An Introduction for the Management Scientist. New York: Springer.

Github

https://github.com/emanueleborgonovo/SensitivityDecisionMaking/tree/main

Github

Thank You for Your Attention!