Managing Uncertainty with Intelligent Scenarios:
Intelligent Scenarios for What?

Mort Webster

Pennsylvania State University

EPRI’s 43rd Annual Seminar on Resource Planning for Electric Power Systems

October 29, 2024

PennState

Overview of This Presentation

- Problem Framing
- A Simple Example
- Critical Assumptions: Timing, Risk, and Constraints
- Overview of Scenario Reduction Methods
- Discussion

Framing the Problem

So, you want to use scenarios…

- **What is your question?**

1. What are different futures that could occur?
2. How can I compare the risks between Plan A and Plan B?
3. What is a plan that, on average, is least cost?
4. What is a plan that protects me from the "worst case"?
5. Are there strategic near-term opportunities to hedge against uncertainty?

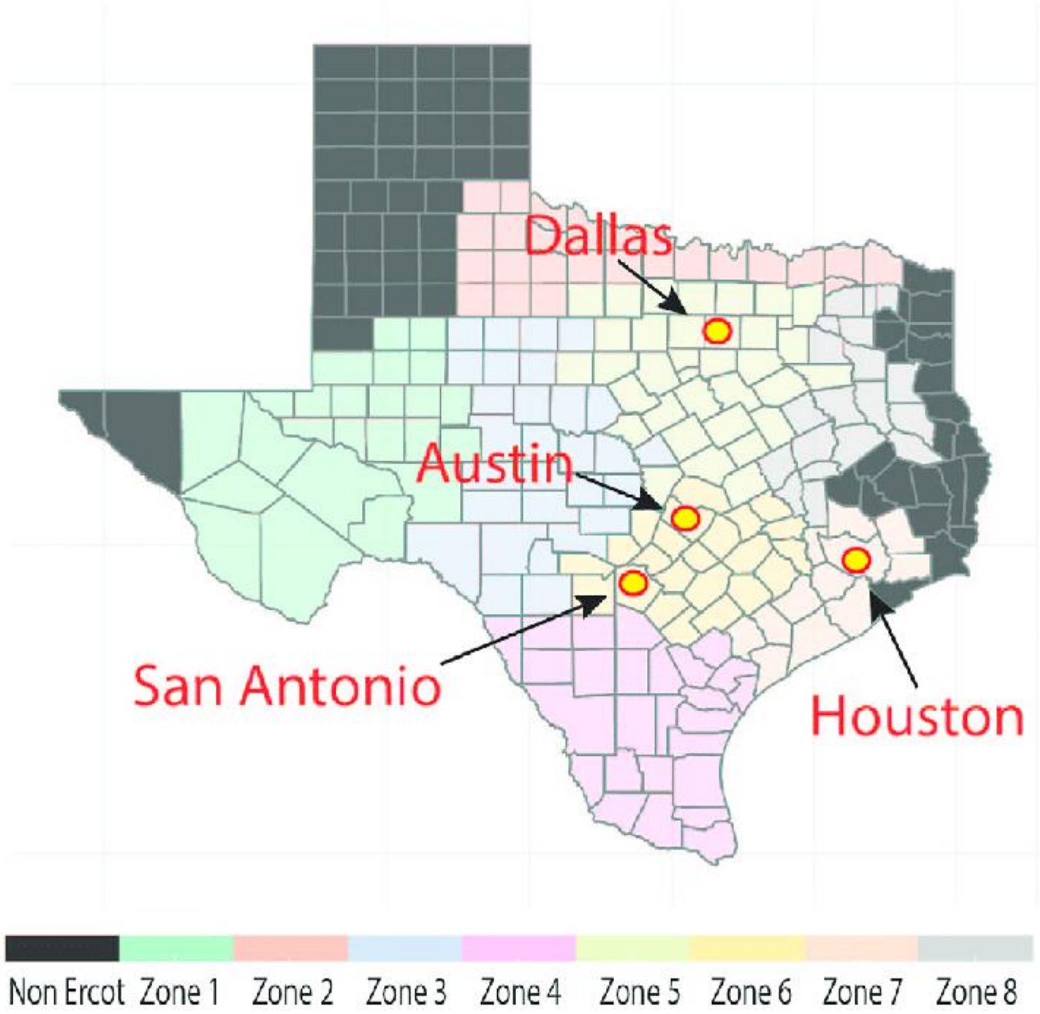
Framing the Problem

Each question requires:

- A different set of scenarios
- A different analysis/solution method

Illustrative Example: ERCOT

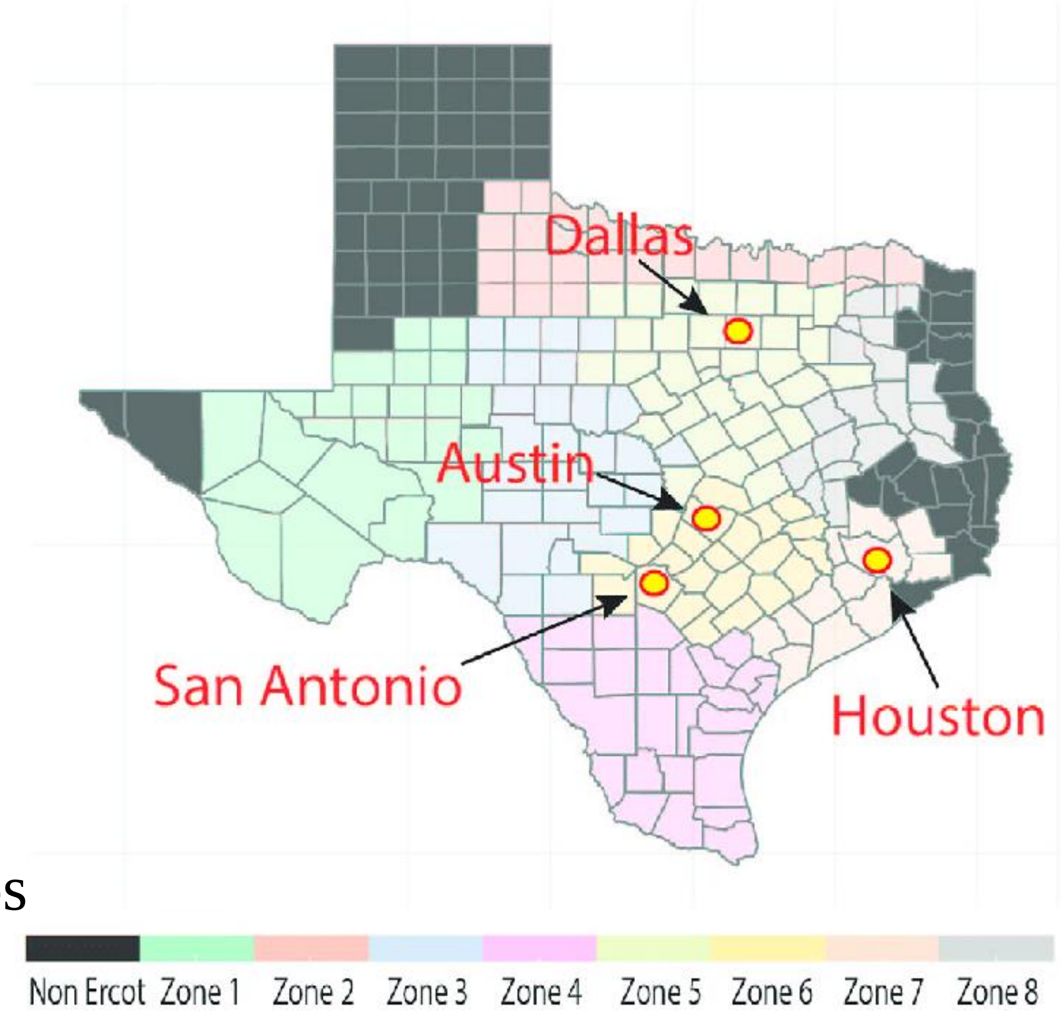
- Based on ERCOT
  - 2018 existing generation mix
  - Omit zonal/transmission constraints
- 15-year planning horizon
  - Focus on two periods: 2030 and 2040
- Candidate Technologies
  - Natural Gas Combined Cycle
  - Natural Gas Combustion Turbine
  - Nuclear
  - Solar
  - Wind



Lee et al. (2022). iScience. 25. 103723.

Problem Formulation

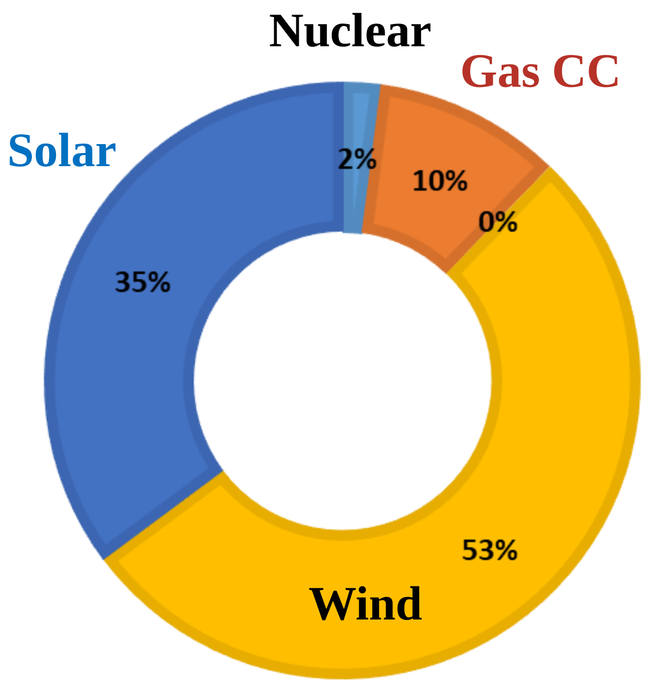
- Constraint:
  - Meet a cumulative CO_2 emissions limit
- Uncertainties (only in 2040)
  - Natural gas price
  - Load growth
  - Emissions limit quantity
  - Cost factor for nuclear capital costs
- "Full" Uncertainty: 50 Scenarios
  - Minimize expected total costs
  - Meet emissions constraint in all scenarios
  - Allow violations with fixed penalty



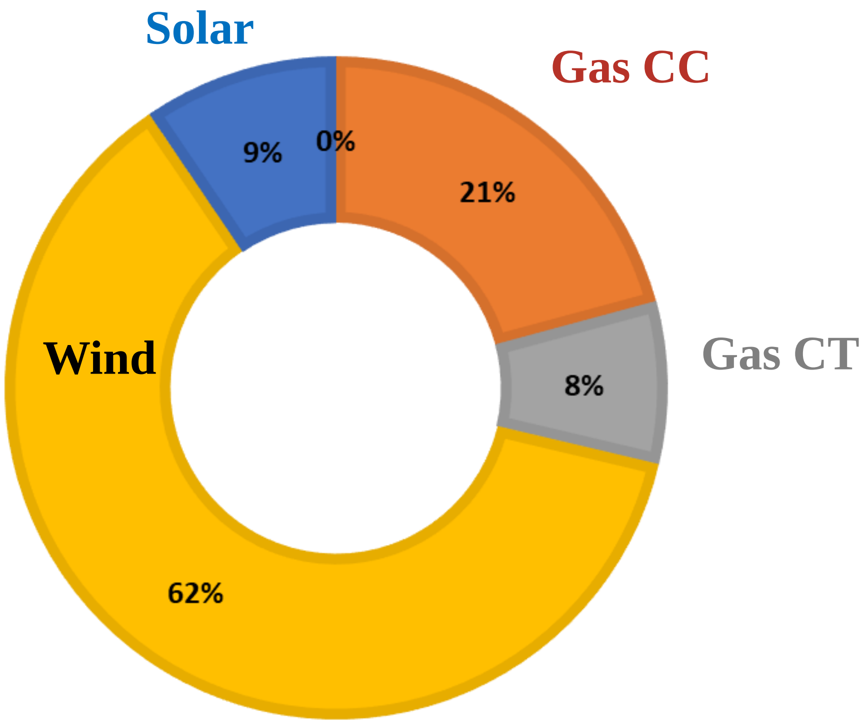
Lee et al. (2022). iScience. 25. 103723.

Three Conceptual Scenarios

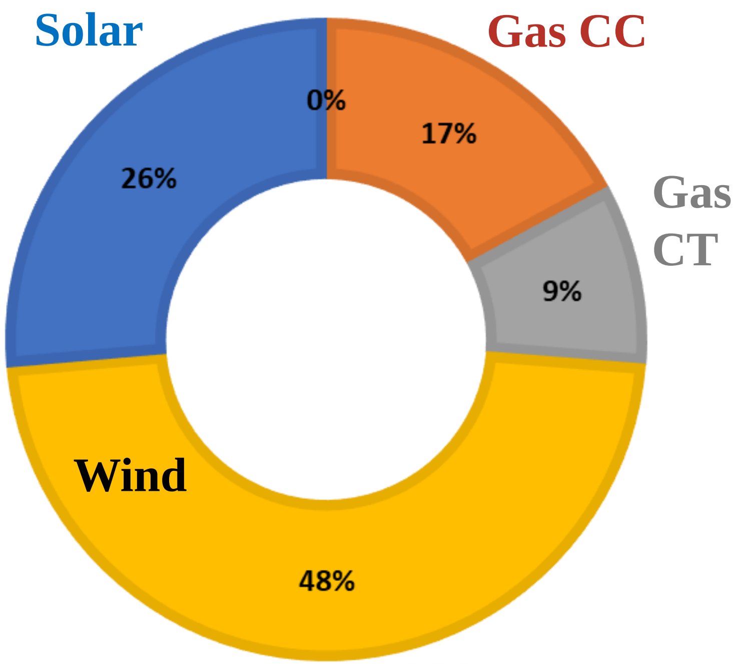
Scenario C:
High Natural Gas Price
Aggressive Emissions Target



Scenario A:
Low Natural Gas Price
Minimal Emissions Target

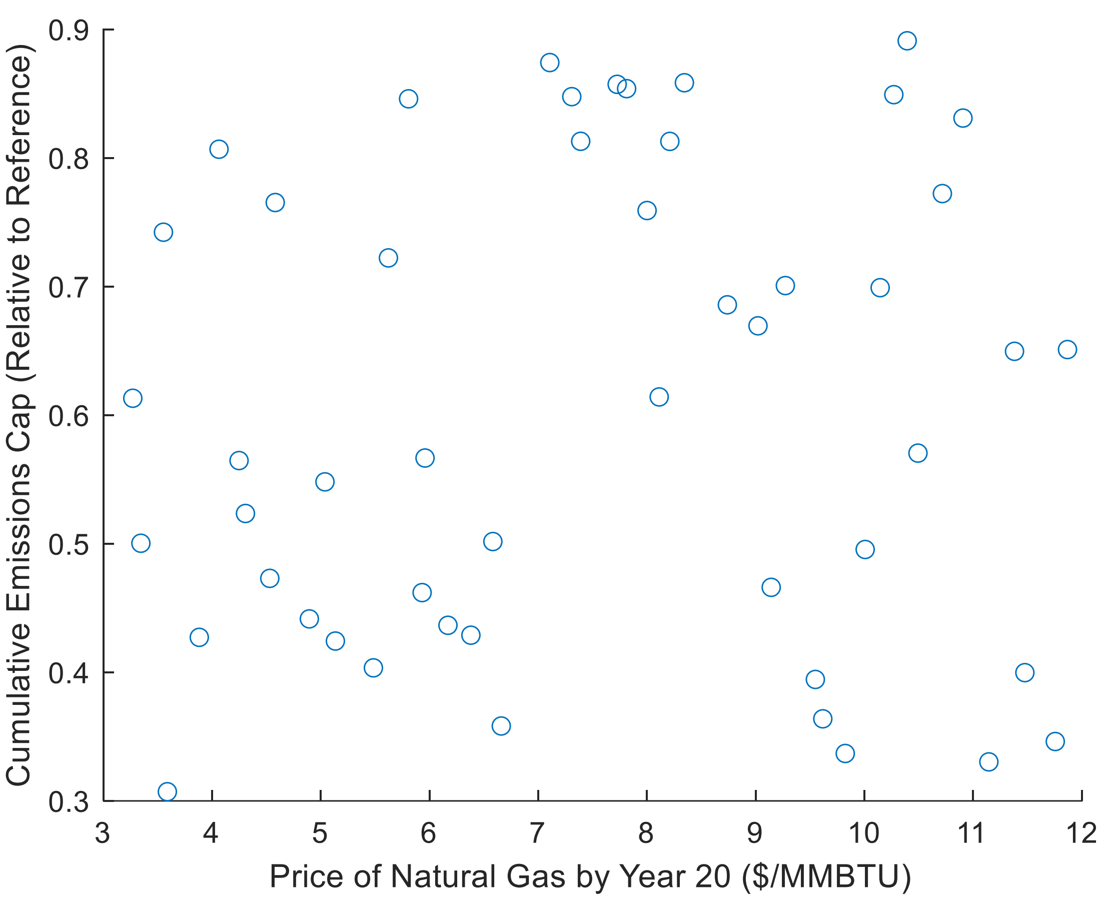


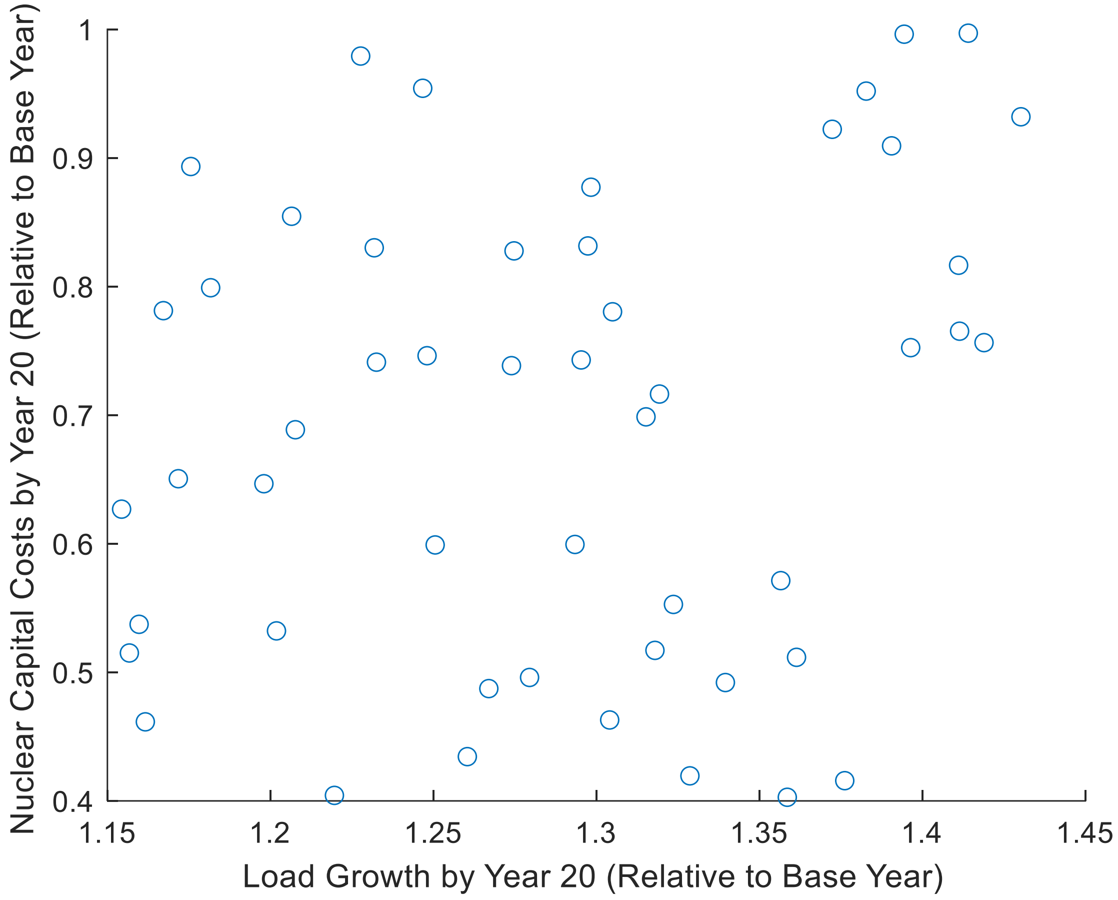
Scenario B:
High Natural Gas Price
Minimal Emissions Target



Share of Cumulative New Capacity (%)

Scenario Set for Illustrative Example

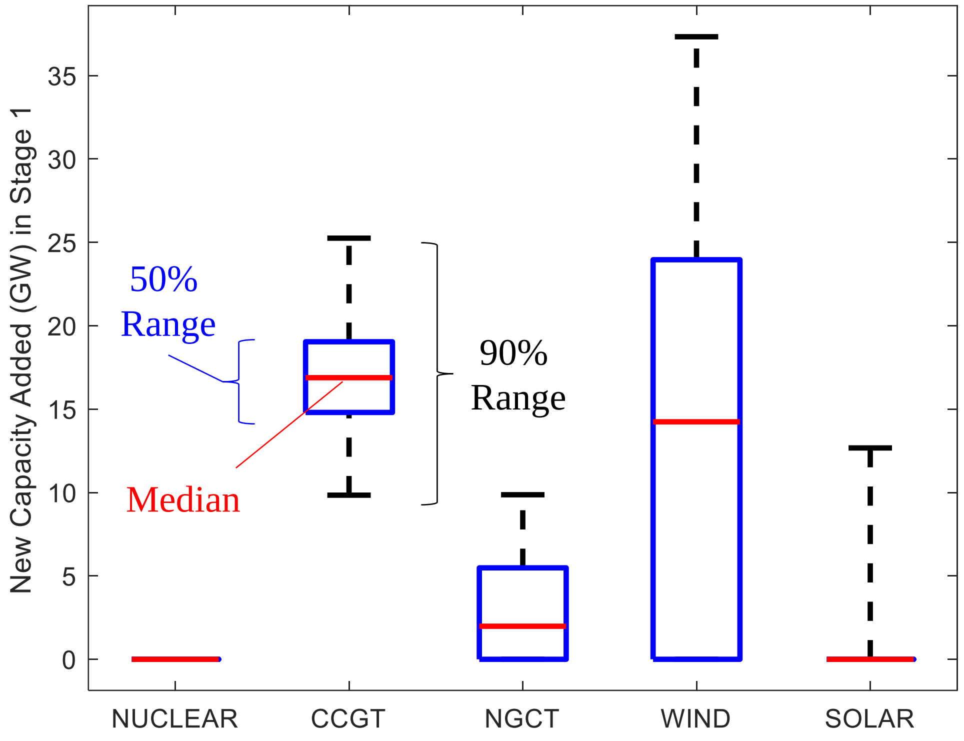




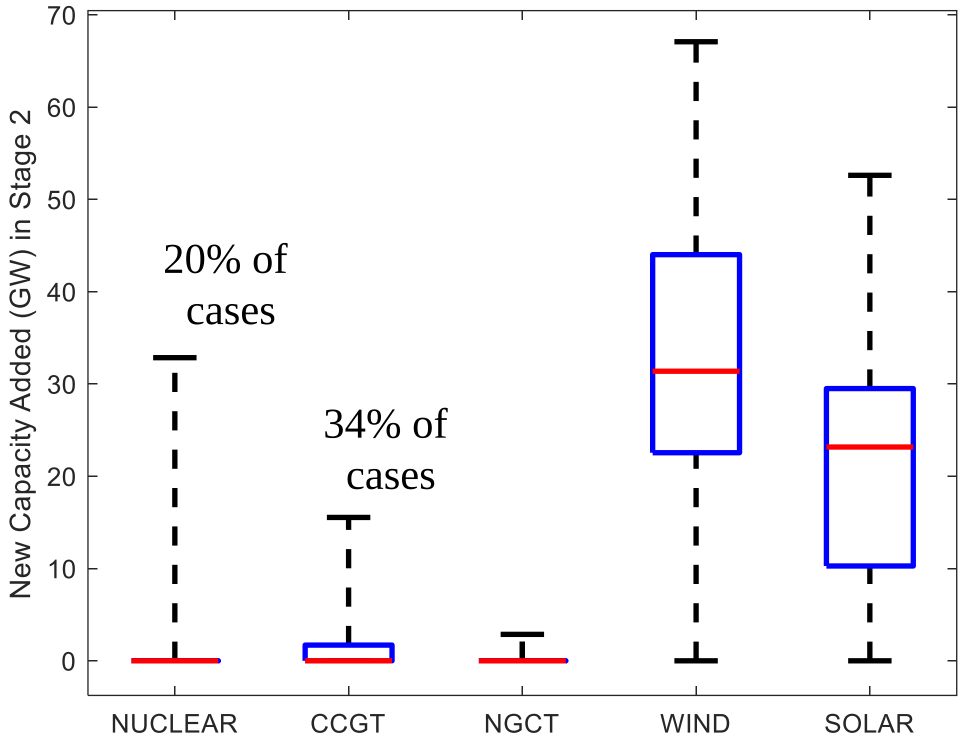
50 Scenarios: Randomly Sampled (Sobol sampling)

Monte Carlo Simulation: 50 Optimal Investment Plans

Stage 1 Investments



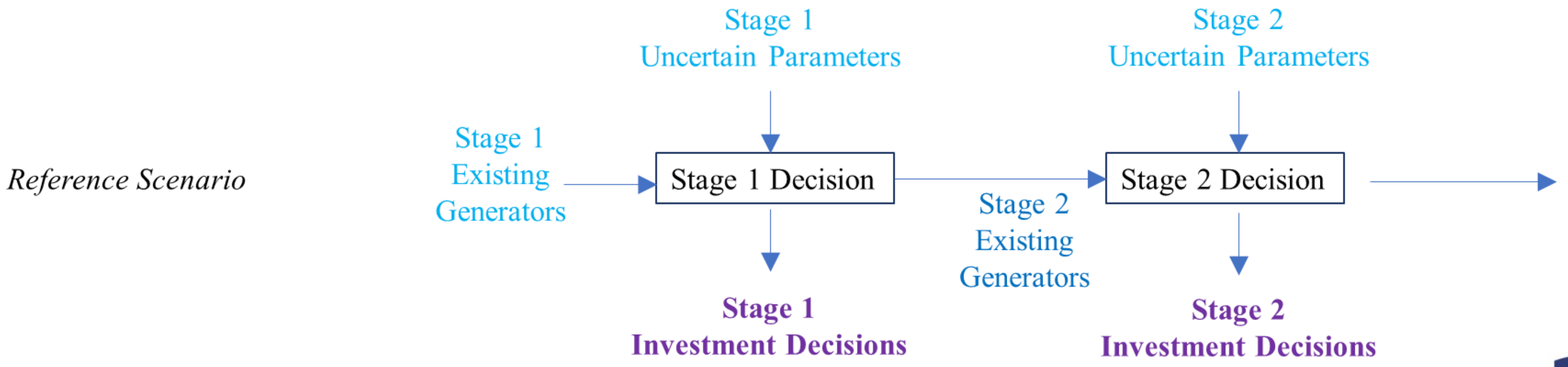
Stage 2 Investments



Each plan is optimal for one scenario; assumes “perfect information”

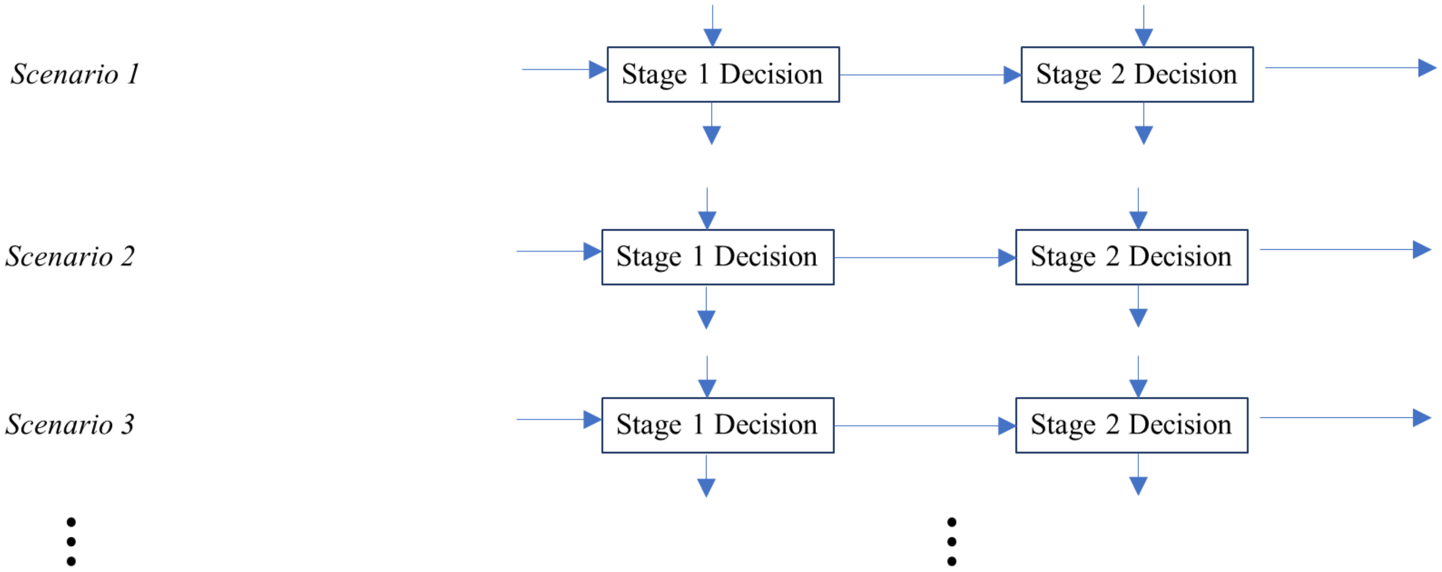
Analysis Frameworks

1. **Scenario Analysis (Three scenarios)**
   - **Three investment plans**
2. Monte Carlo Simulation (50 scenarios)
   - 50 investment plans
3. Two-Stage Stochastic
   - 2030 investments common across scenarios



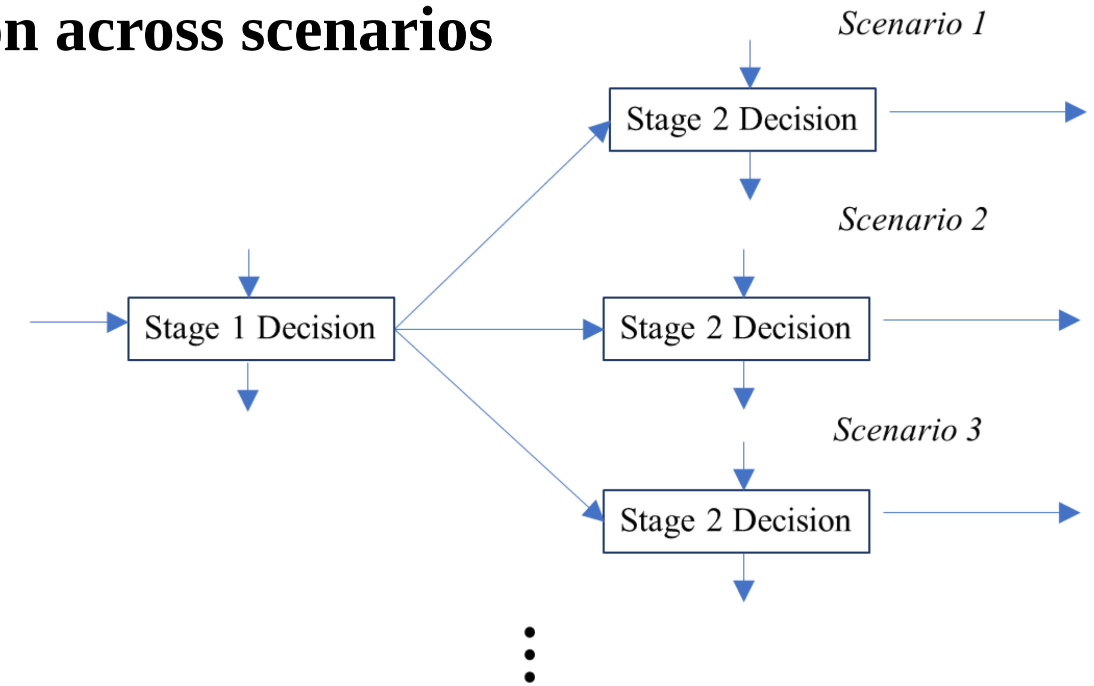
Analysis Frameworks

1. Scenario Analysis (Three scenarios)
   - Three investment plans
2. **Monte Carlo Simulation (50 scenarios)**
   - **50 investment plans**
3. Two-Stage Stochastic
   - 2030 investments common across scenarios

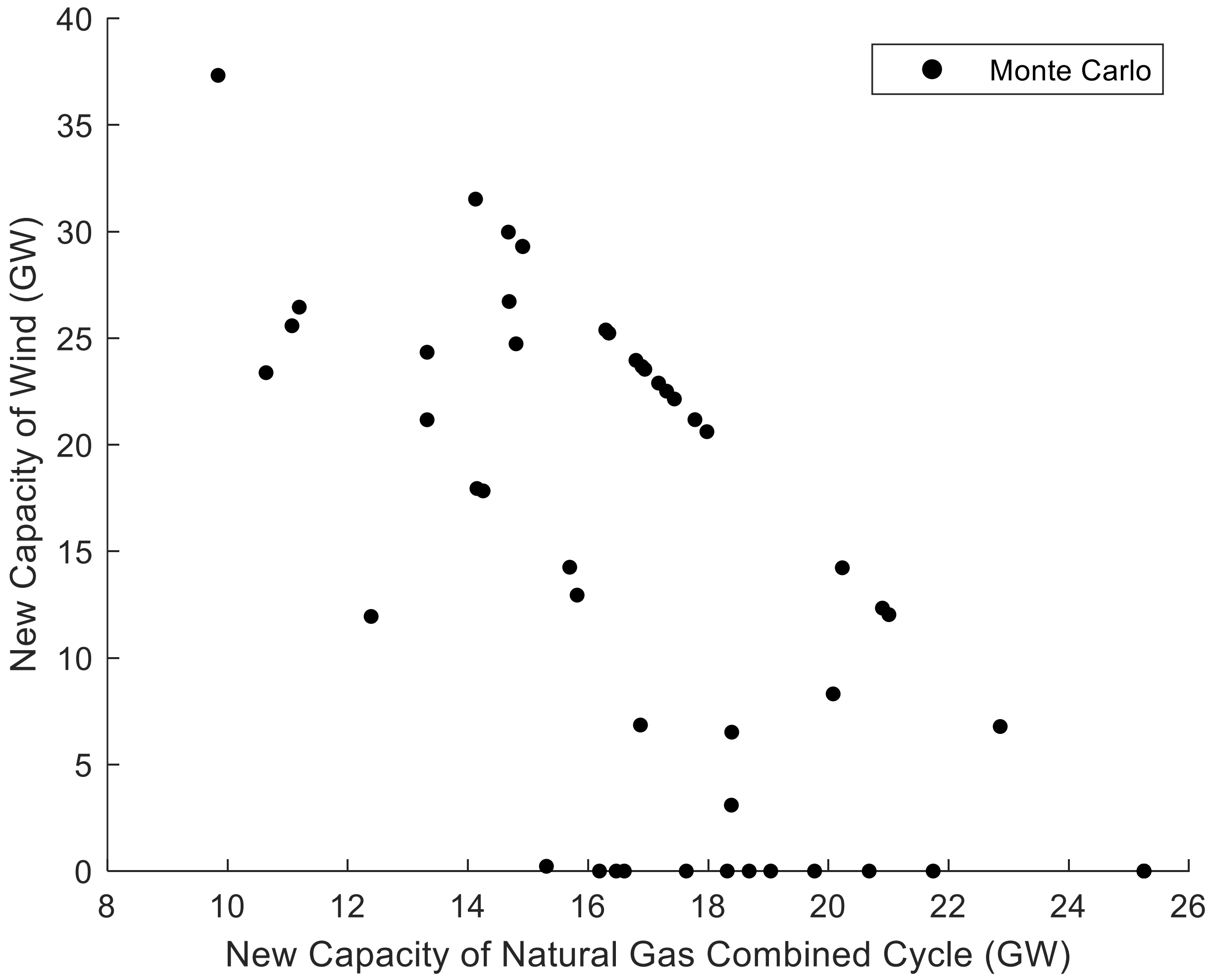


Analysis Frameworks

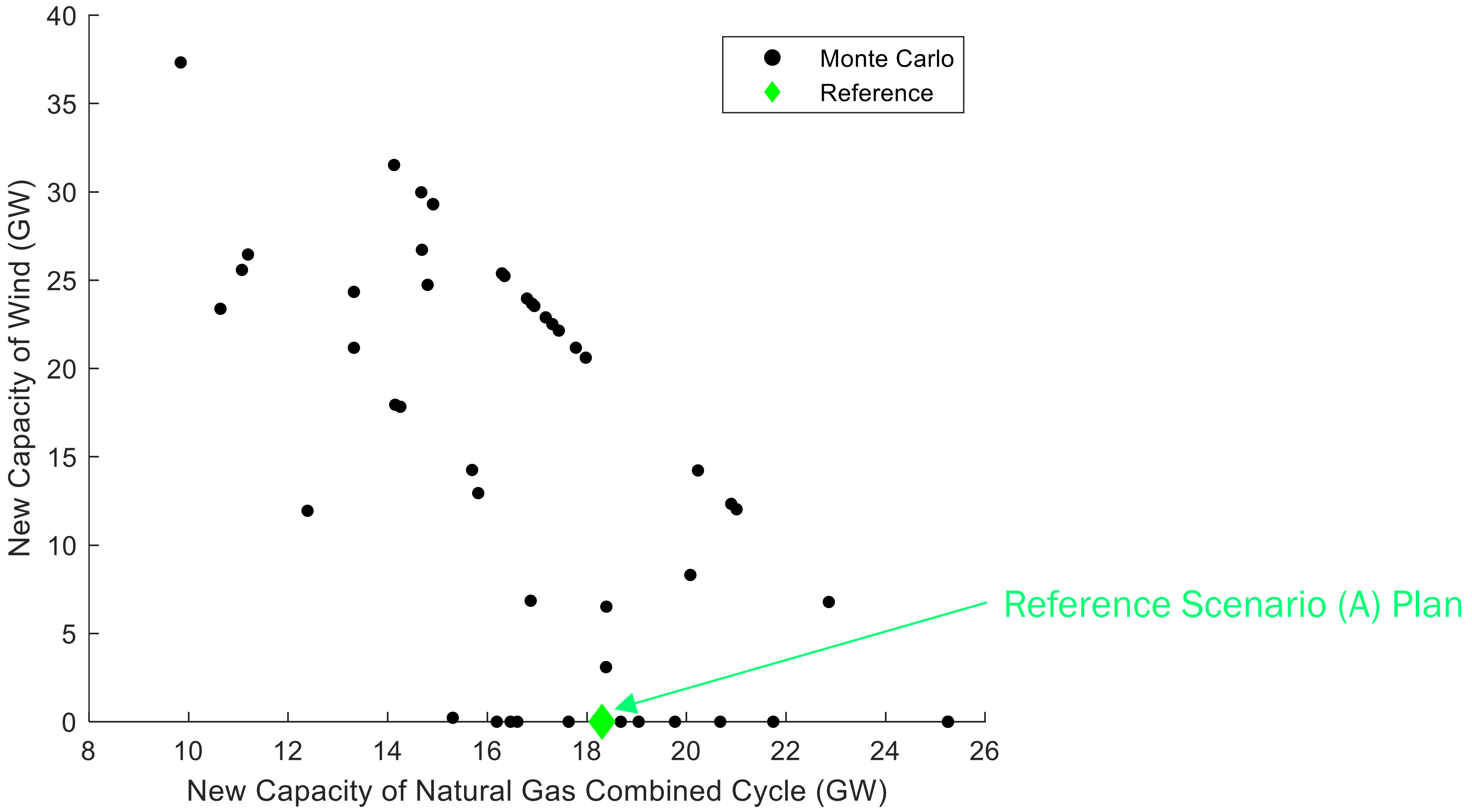
1. Scenario Analysis (Three scenarios)
   - Three investment plans
2. Monte Carlo Simulation (50 scenarios)
   - 50 investment plans
3. **Two-Stage Stochastic**
   - **2030 investments common across scenarios**



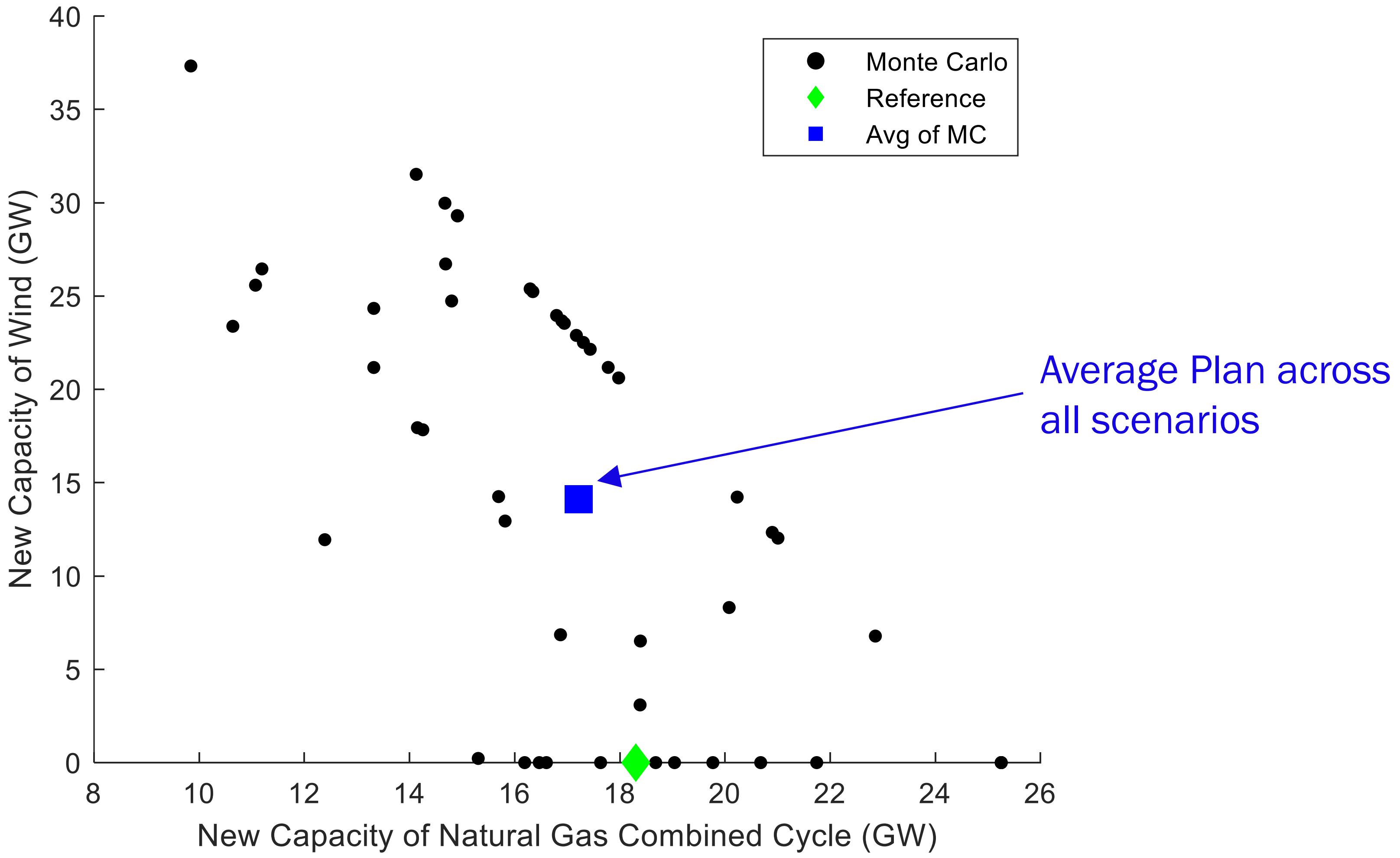
Optimal Investment Plan: Stage 1 (2030)



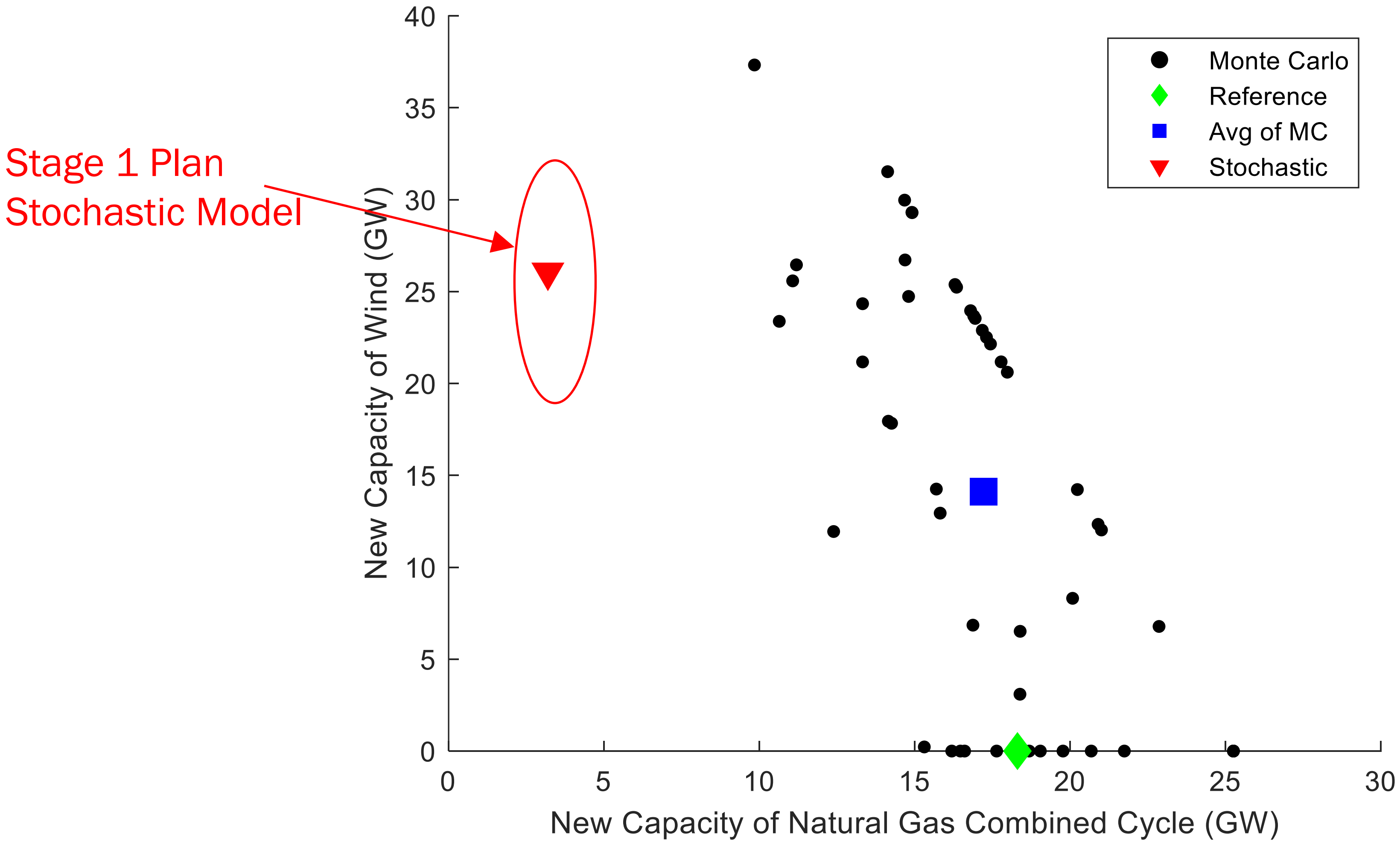
Optimal Investment Plan: Stage 1 (2030)



Optimal Investment Plan: Stage 1 (2030)



Optimal Investment Plan: Stage 1 (2030)



Option-value to investing in less CCGT in first stage; can build more later if needed

Before Selecting Scenarios: Setting up the Problem

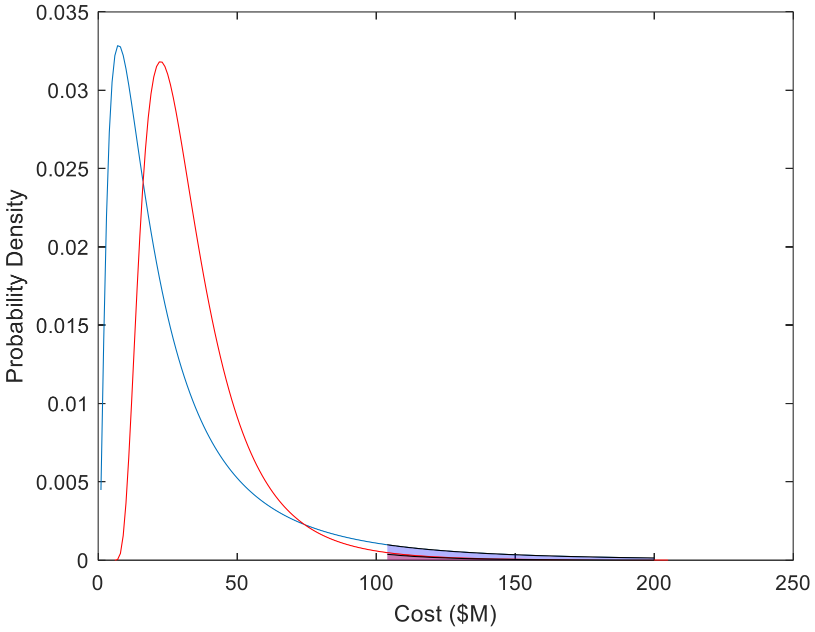
- Problem Formulation
  - What is the question the analysis should address?
- Temporal structure
  - When will information be updated? When can decisions be made?
- Treatment of Constraints
  - Hard constraint? Penalty for violation? Uncertainty?
- Treatment of risk
  - Risk measures in objective function
  - Risk measures in constraints
  - Alternative formulations for cost / constraint tradeoffs

Managing Uncertainty: Question Types

- How do I compare performance of alternative investment plans?
  - Simulate candidate plans under many future scenarios
  - Construct risk profile (e.g., distributions of cost, reliability) for each plan
  - ➢Monte Carlo Simulation
- How do I find a plan that does well on average?
  - ➢Stochastic Optimization – minimize expected costs
- How do I prepare for the worst-case?
  - ➢Robust Optimization
- What if I am risk-averse, but RO is too extreme?
  - ➢Stochastic Optimization with Contingent Value at Risk (CVaR)
  - ➢Stochastic Optimization with Chance Constraints
  - ➢Other hybrid approaches

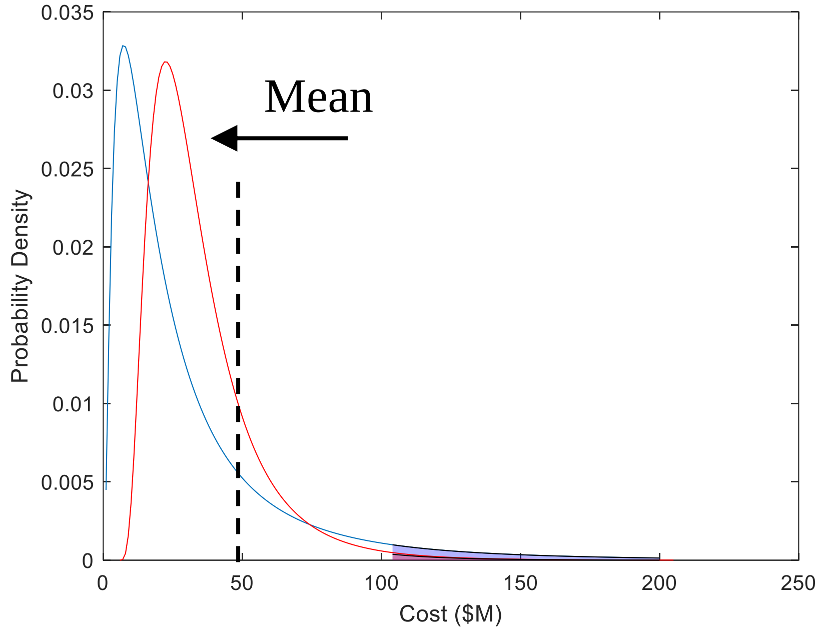
Treatment of Risk

- Risk measures in objective function
  - Distribution of total cost across scenarios
- Risk measures in constraints
  - Distribution of violations across scenarios
- Depends on which constraints
  - Meet demand
  - Capacity reserve margin
  - Emissions targets
- Alternative formulations for cost / constraint tradeoffs



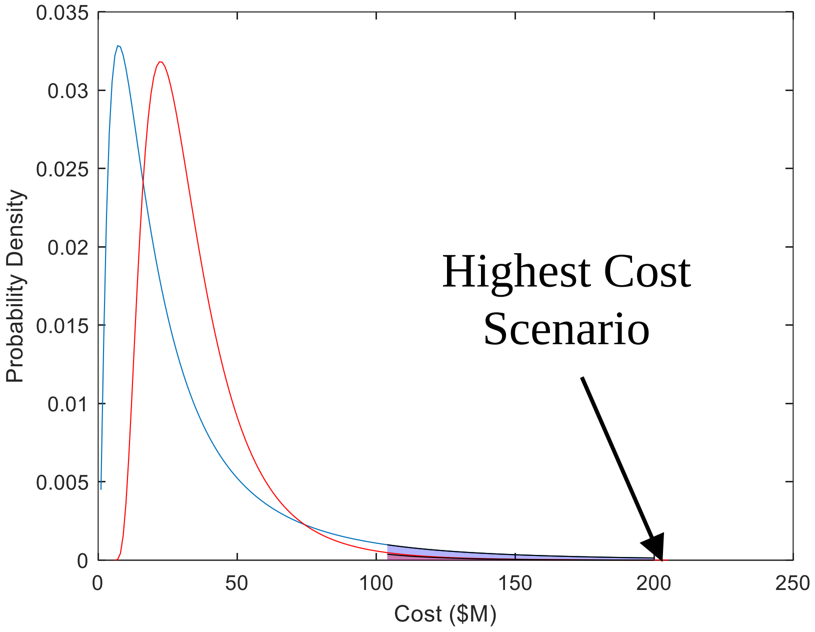
Treatment of Risk in Objective Function

- Given the cost from every scenario, what do you want to minimize?
- Expected Costs



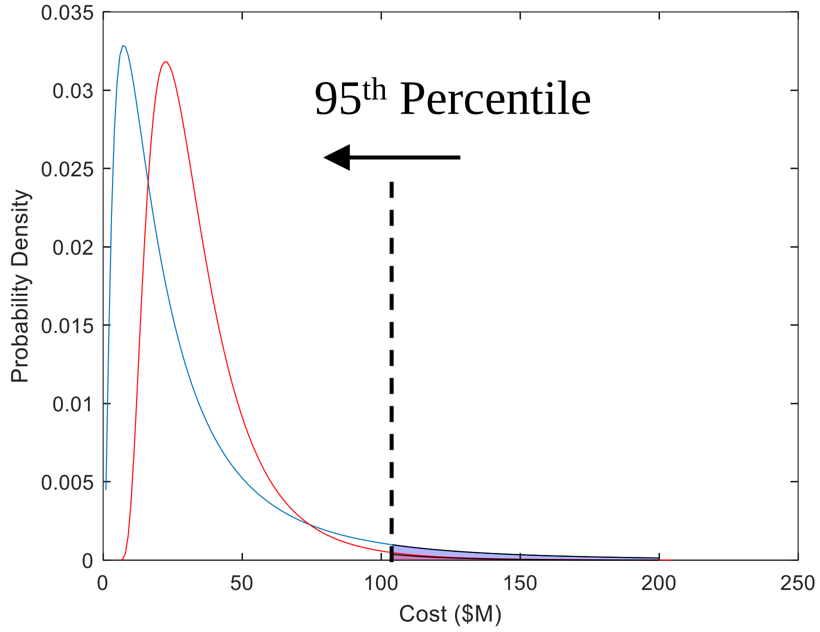
Treatment of Risk in Objective Function

- Given the cost from every scenario, what do you want to minimize?
- Expected Costs
- Minimize cost of worst scenario



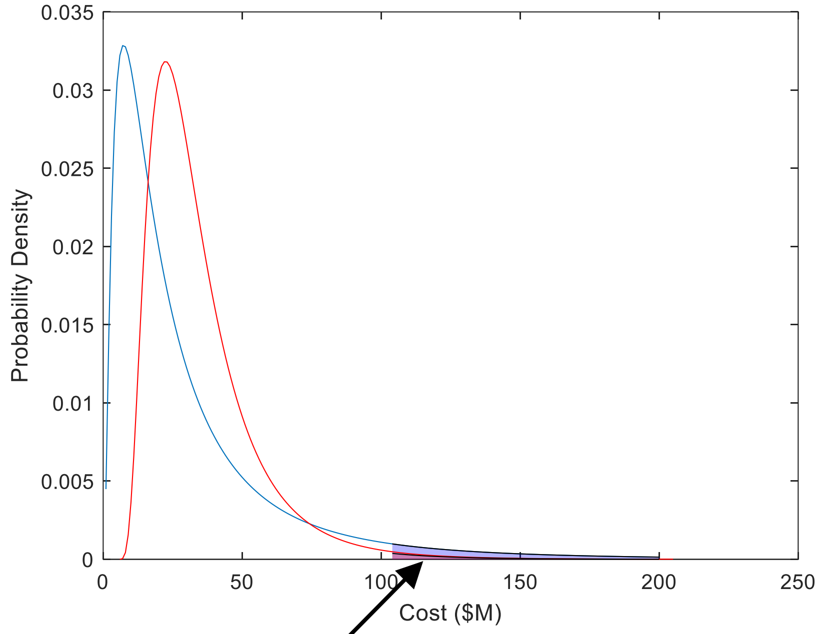
Treatment of Risk in Objective Function

- Given the cost from every scenario, what do you want to minimize?
- Expected Costs
- Minimize cost of worst scenario
- Minimize a percentile of the cost distribution (VaR)



Treatment of Risk in Objective Function

- Given the cost from every scenario, what do you want to minimize?
- Expected Costs
- Minimize cost of worst scenario
- Target a percentile of the cost distribution:
- Contingent Value at Risk (CVaR)



Average Loss
Above $100M

Dimensionality Reduction: The Full Problem

Find minimum average cost investment plan considering:

- All possible long-term future scenarios (infinite)
- All planning periods (annual for 20 years)
  - Recourse decisions every period
- All hours of each year for operations (8760)
- Many samples of forced outages for each hour/year/scenario
- All candidate units for addition or retirement
- Fully detailed operations model with all constraints (UC/OPF)

We cannot solve the full problem -> too large!

Dimensionality Reduction / Model Tractability

How can I solve GEP under uncertainty in a reasonable amount of time?

- Reduce the number of elements in one or more of :
  - Number of future scenarios of long-term uncertain parameters
  - Number of operational hours per planning period
  - Number of planning periods
  - Number/resolution of candidate resources
- Simplify operations model
  - Fewer constraints
  - Aggregate resolution (time, spatial)
- Use decomposition scheme to solve large problem efficiently
  - Can include more scenarios / hours

Dimensionality Reduction / Model Tractability

How can I solve GEP under uncertainty in a reasonable amount of time?

- Reduce the number of elements in one or more of :
  - **Number of future scenarios of long-term uncertain parameters**
  - **Number of operational hours per planning period**
  - Number of planning periods
  - Number/resolution of candidate resources
- Simplify operations model
  - Fewer constraints
  - Aggregate resolution (time, spatial)
- Use decomposition scheme to solve large problem efficiently
  - Can include more scenarios / hours

Dimensionality Reduction: The Goal

- Goal:
  - Solution to the Approximate Problem
  - Should be “close to” the solution of the “Full Problem”

➢ ***Which result do you want to approximate?***
  - The optimal total cost?
  - The Stage 1 investment plan?
  - The risk of not meeting a constraint?

Long-Term Scenario Selection

- Create very large scenario set, and down-select
- Repeated sampling subsets from full set
  - Forward selection: Add scenarios until some objective is met
  - Backward selection: Remove scenarios until some objective is met
- Clustering-based reduction methods
  - Cluster based on similar inputs (e.g., similar load/wind/solar patterns)
  - Probability distance methods
    - Select a subset that approximates the same outcome (e.g., expected cost)
  - Decision covariance methods
    - Cluster scenarios to maximize variance across candidate decisions
- Importance Sampling
  - Ensure sufficient samples to represent the “tail”

Short-Term Uncertainty (Operational Hours)

- Within any planning year, need operations cost for candidate plans
  - Within-year variability in load, wind, solar, forced outages, etc.
  - Using 8760 hours may be prohibitive
  - How to select a subset of hours, and how to weight them?
- Traditional approach: select representative hours (LDC)
  - Because of expected increase in renewable generation, energy storage
  - Requires chronological sequences of hours
- Select some number of segments (days, weeks) of chronological hours of operating conditions, with an associated weight

Short-Term Uncertainty: Methods

Assume the goal is to select representative days (24 hours) or weeks (168)

1. Random Sampling
   - Select a subset of days or weeks
2. Clustering
   - Solve all days for one or more plans (get operation cost)
   - Cluster/select subset of days to approximate the operation cost
   - Various clustering methods: similar to those used for long-term scenarios
3. Chronological Time Period Clustering
   - Solve all days for one or more plans (to obtain operation cost)
   - Merge consecutive time periods that are "similar"
   - Same idea as Network Reduction methods for OPF.

➢ ***Should long-term and short-term clustering be independent?***

Summary

- Planning under uncertainty encompasses many questions
  - Each analysis requires a different scenario set
- Critical assumptions to think about:
  - The timing of information and decisions
  - The relevant constraints and their representation
  - The appropriate degree of risk aversion and its representation
- Scenario reduction methods (long-term)
  - Select the subset that approximates ***your objective*** in the analysis
- Representative hours selection (short-term)
  - Best selection varies across long-term scenarios

Thank you

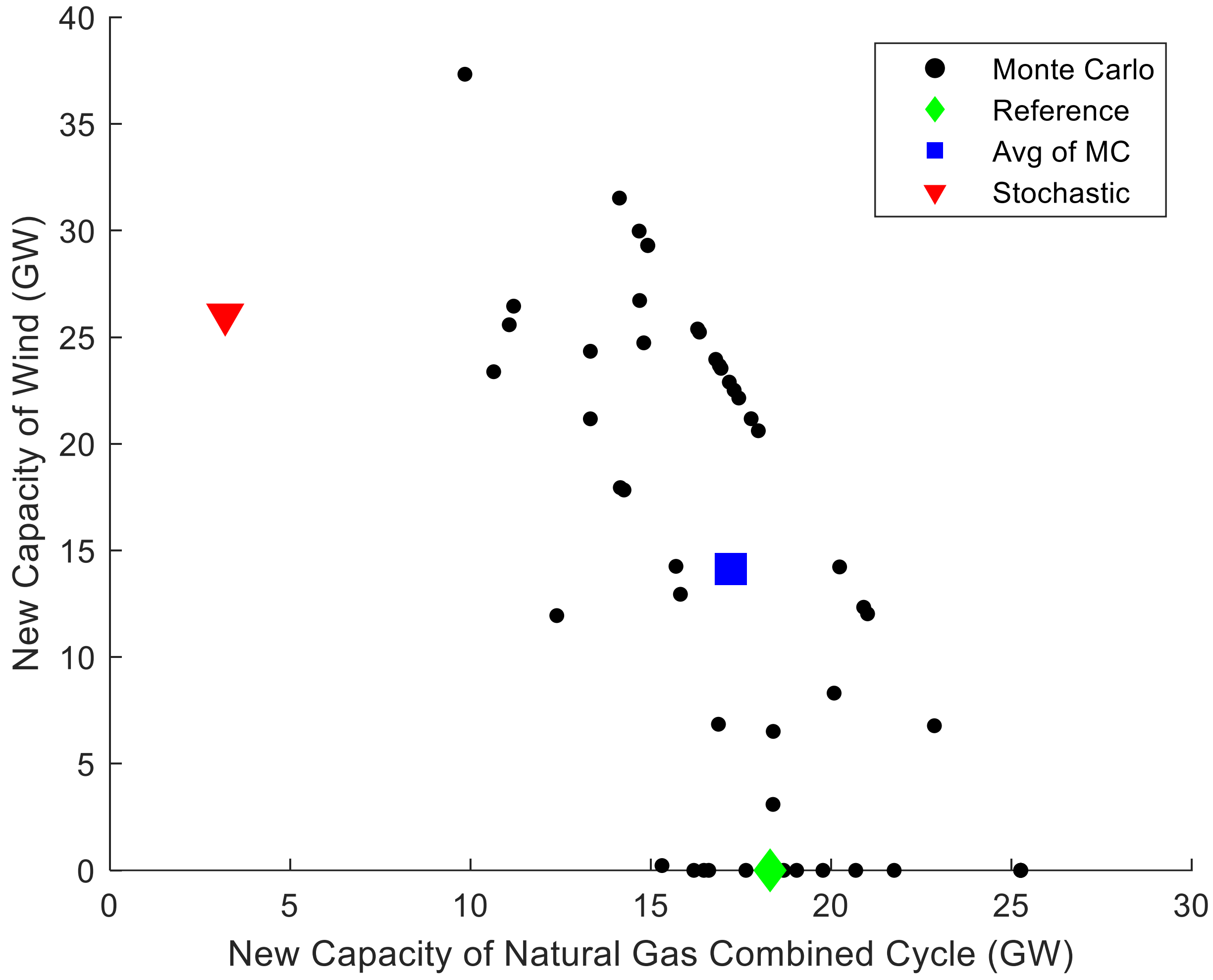
Contact: Mort Webster
mdw18@psu.edu
Professor of Energy Engineering
Dept. of Energy and Mineral Engineering
Pennsylvania State University

PennState

Extra: Example of Cost Risk

ERCOT Example: Stage 1 Build

- Scenarios / Monte Carlo:
  - Plan is optimal only in that scenario
  - Does not consider risk
- Stochastic Solution
  - Considered all scenarios
  - Lowest average cost across scenarios
- There is still a distribution of costs over the scenarios
- How do the plans differ in terms of the entire risk distribution?



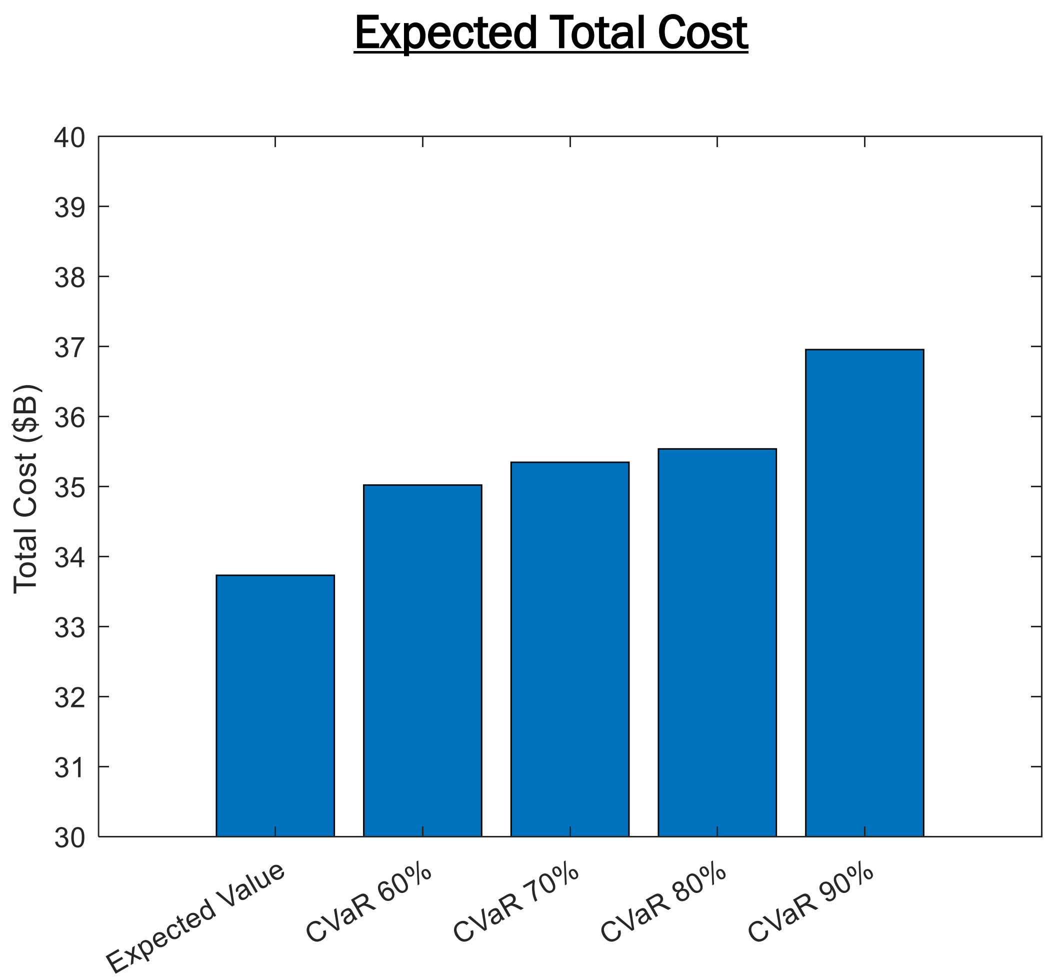
Risk-Averse Objective Function

- What if you care more about the higher cost scenarios?
- Traditionally, focus on the Value-at-Risk (VaR)
  - This is just the $1 - \alpha$ percentile out of the cumulative distribution
  - E.g., minimize the 90$^{\text{th}}$ percentile cost
- Contingent Value-at-Risk
  - Expected value for all scenarios above the target percentile
  - E.g., minimize average losses greater than X

ERCOT Example: CVaR in Objective Function

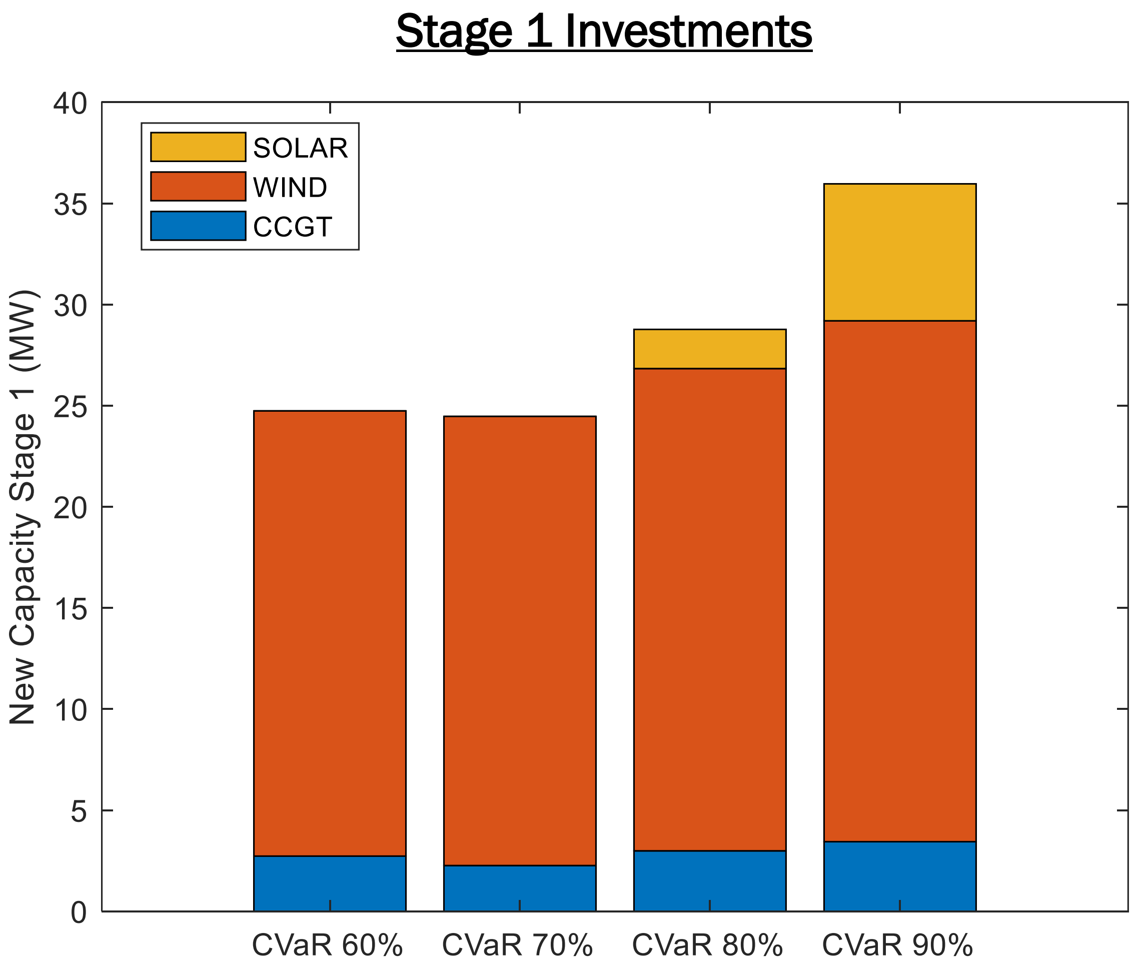
- Tested several different target levels:
  - 60%
  - 70%
  - 80%
  - 90%

Increasing Degrees of Risk-Aversion



Risk-Averse Solutions Have Higher *Average* Costs

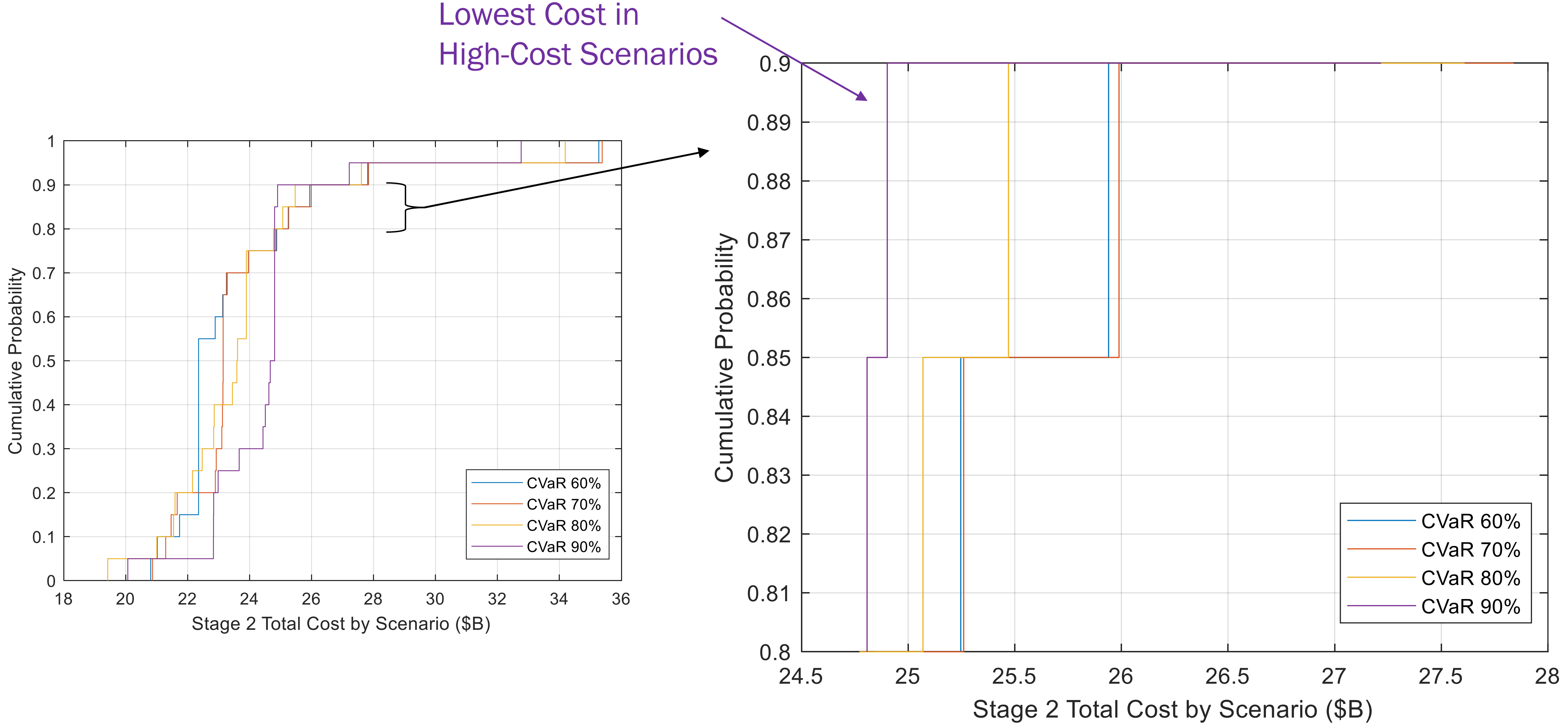
ERCOT Example: Impact of Risk Aversion on Investments



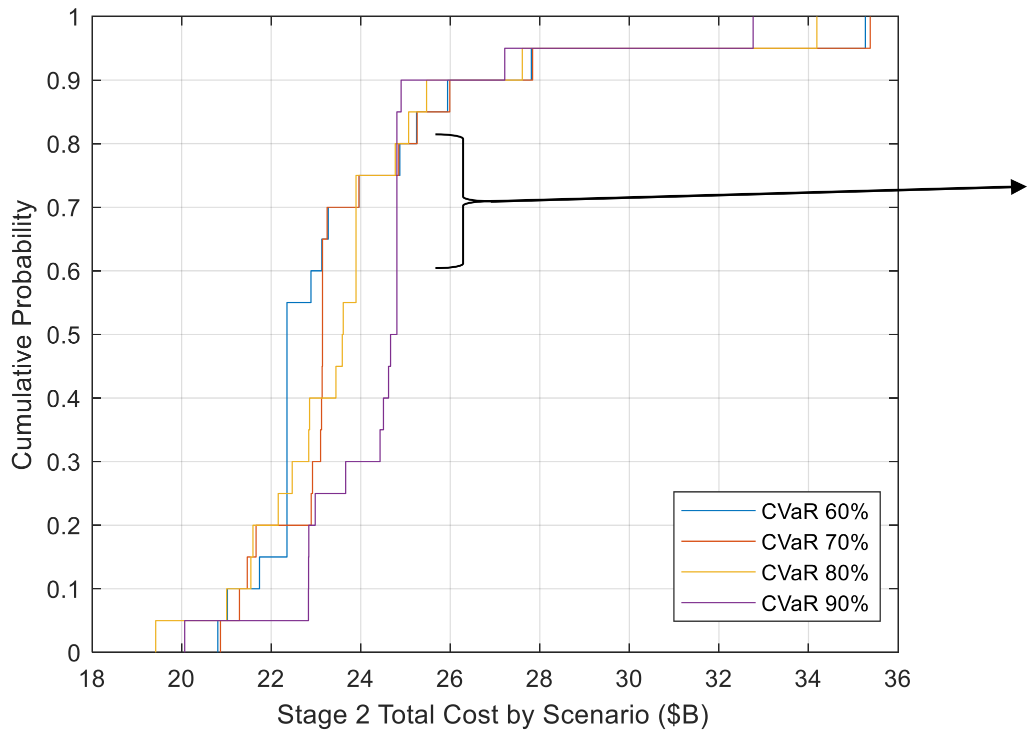
Greater Risk-Aversion: Additional Investments for “worst” scenarios

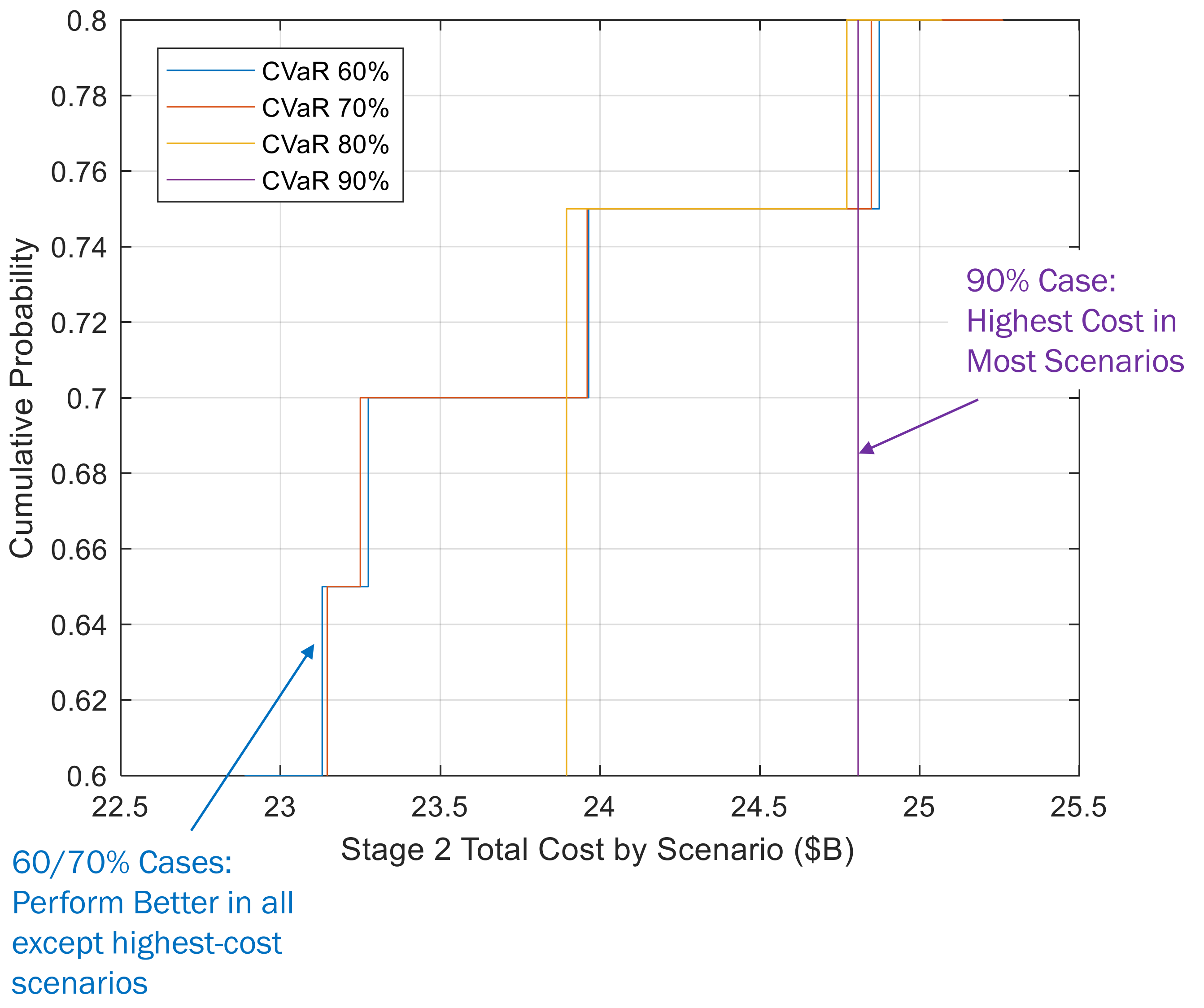
ERCOT Example: Impact of CVaR on Distribution of Costs

90% Solution:
Lowest Cost in
High-Cost Scenarios



ERCOT Example: Impact of CVaR on Distribution of Costs





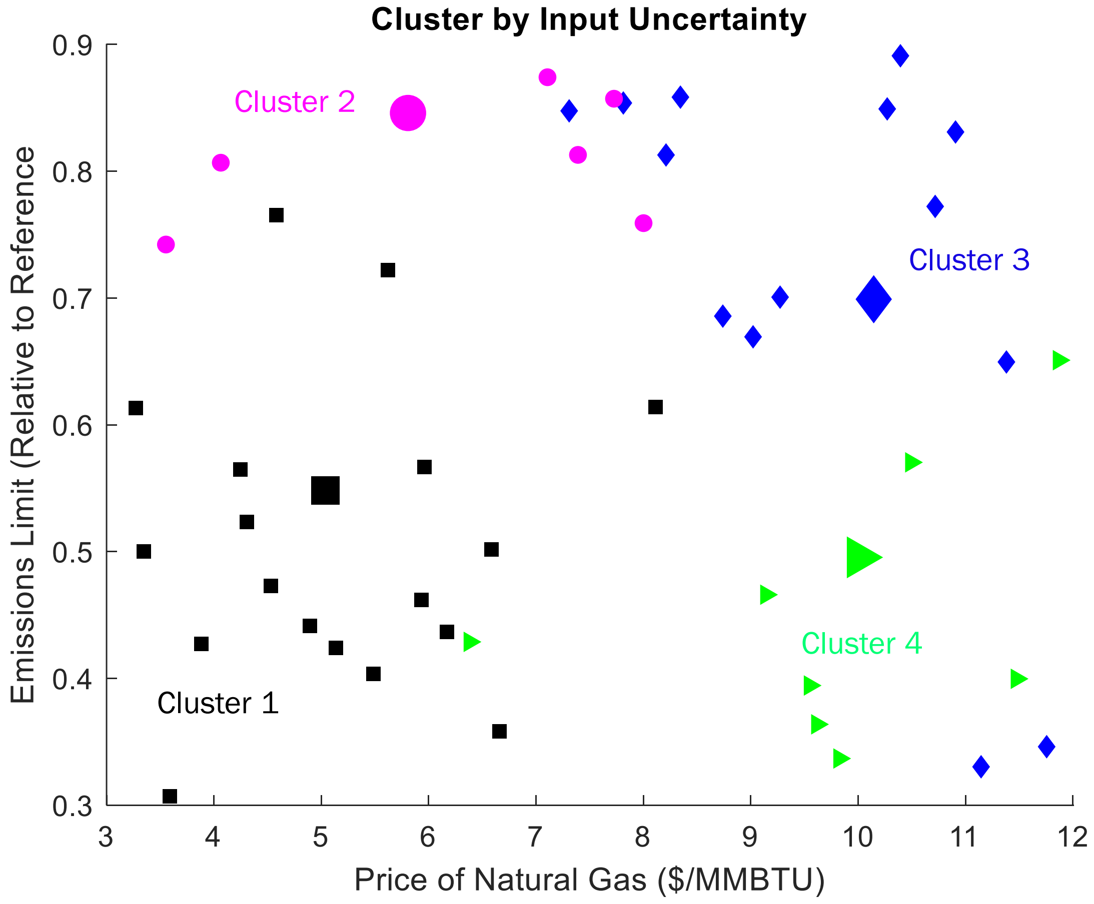
Extra: Long-Term Scenario Clustering Methods

Clustering for Long-Term Scenario Reduction

- Too many scenarios to include all
- Find the subset of scenarios that approximate the “true” solution
- Want to include
  - At least one scenario that needs a different solution
  - Do not include multiple scenarios that need the same solution
- Example:
  - Assume the 50 scenarios is the “full set” of uncertainty
  - Assume you can only include 4 scenarios in the stochastic model

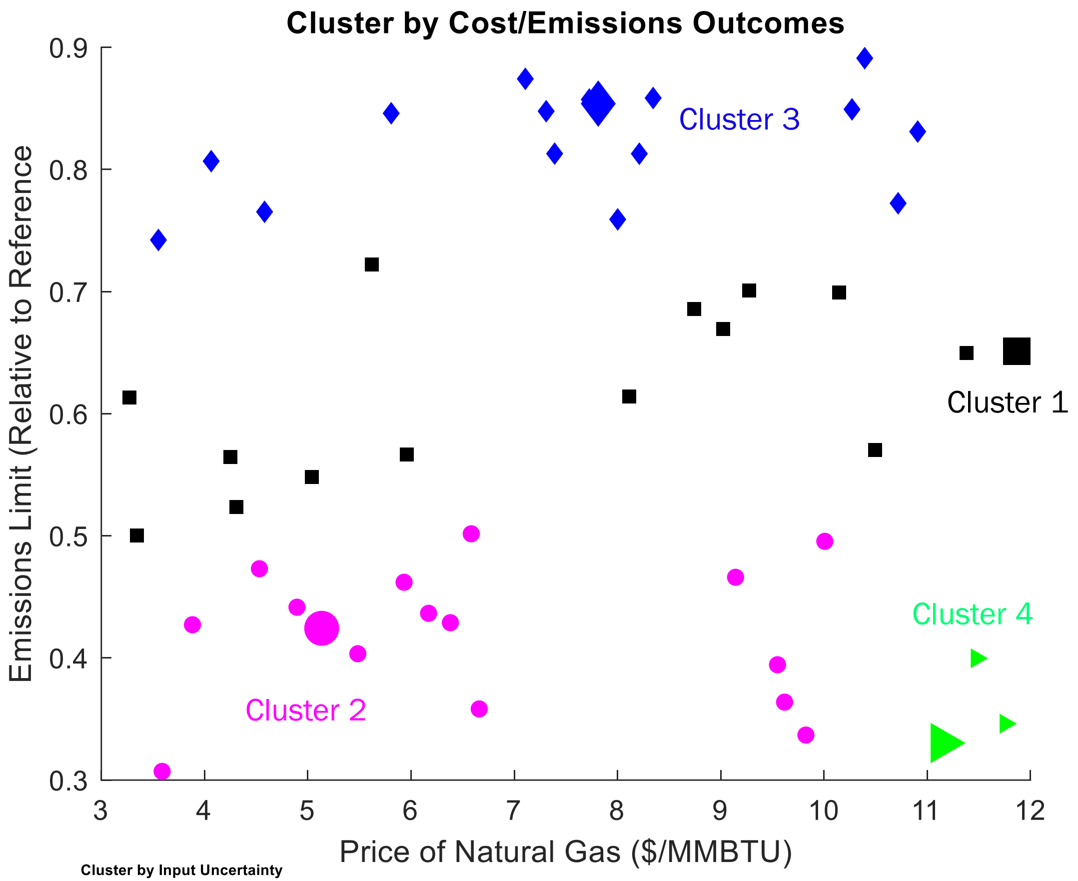
1) Clustering on Input Uncertainties

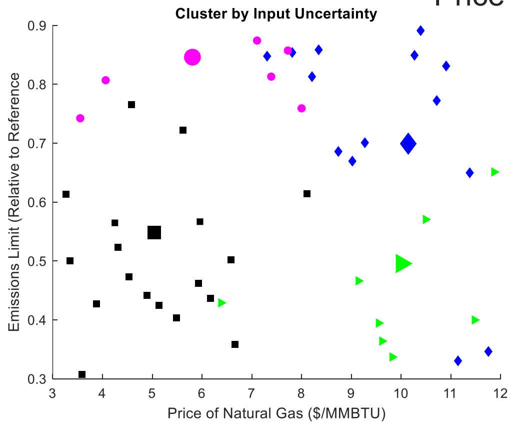
- Apply K-means clustering to the full sample set
  - Group into 4 clusters
  - Identify the "medoid" scenario
- Minimizes the distance within each group from medoid
- Maximizes the distance between medoids across groups
- "Weight" of each medoid: how many scenarios in its cluster



2) Clustering on Outcomes

- Problem: Clustering by inputs may not map directly to cost impacts
- Distance-weighted or probability-weighted methods
- Cluster by Total Cost across scenarios
- Can consider multiple outcomes
  - e.g., Cost and emissions
- Clusters by relative impact on cost
- The reduced set will provide a better approximation of total cost from the full uncertainty set



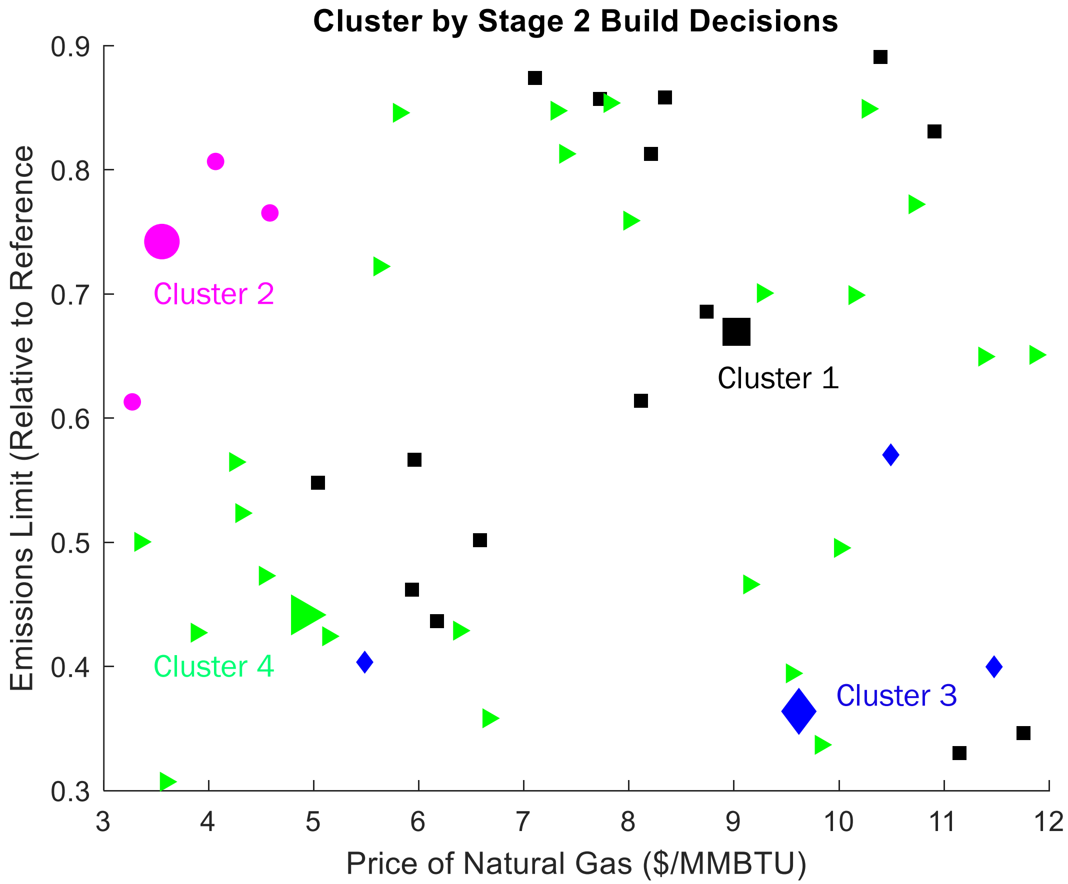


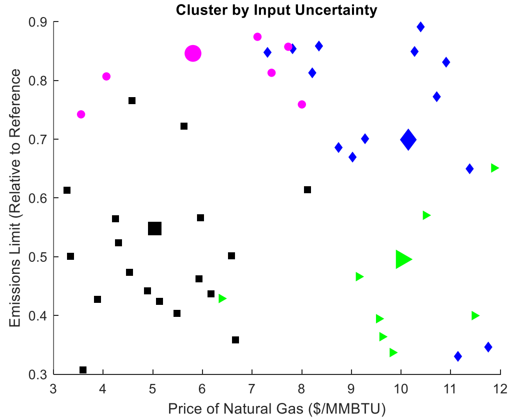
NOTE:
Different partitions the methods

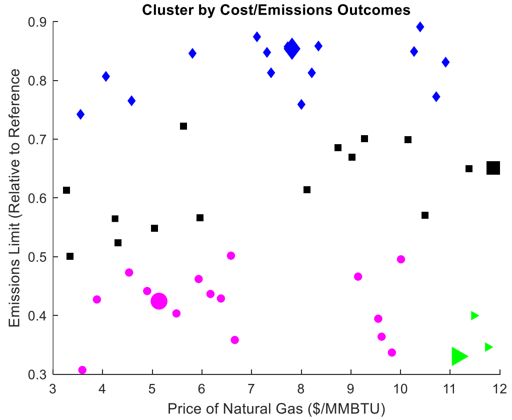
3) Clustering on Decisions

- Problem: Some high-cost scenarios might not be addressed by the decisions
  - Distance-weighted approximates the cost only in the reference case
  - Selected scenarios might not distinguish between investments
- Decision-based clustering
  - Identify groups that favor a different investment plan
  - Include one scenario from each group

NOTE: Does not partition the input space into distinct regions

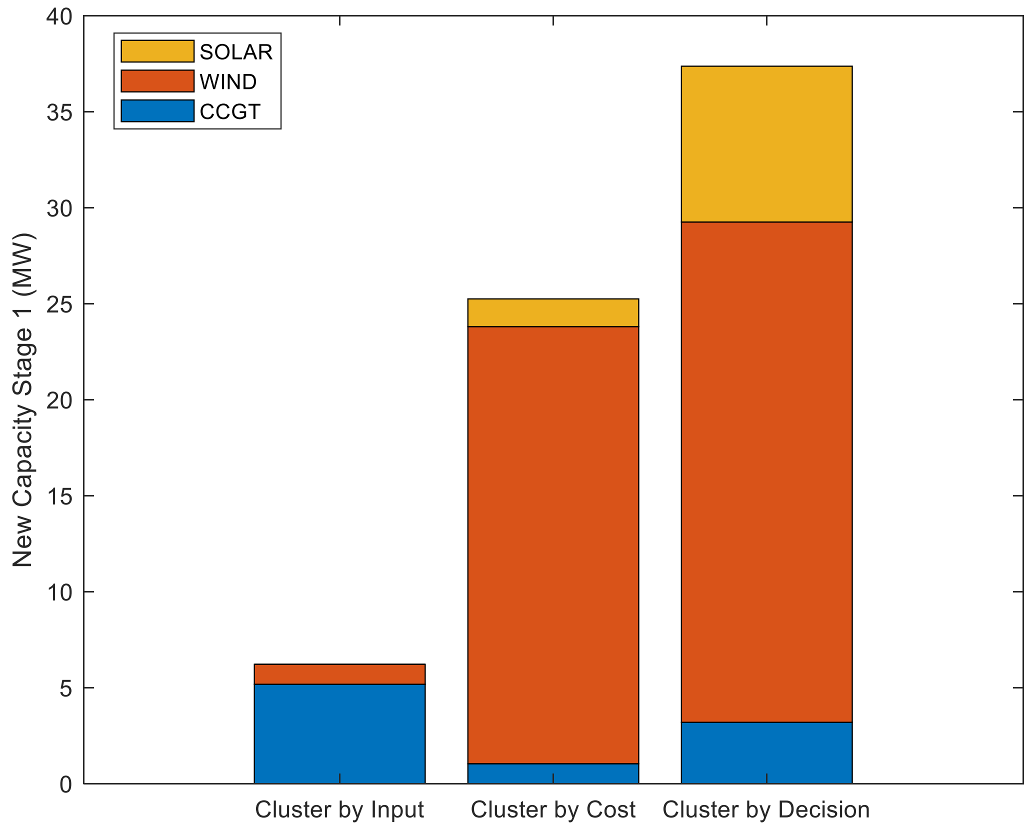




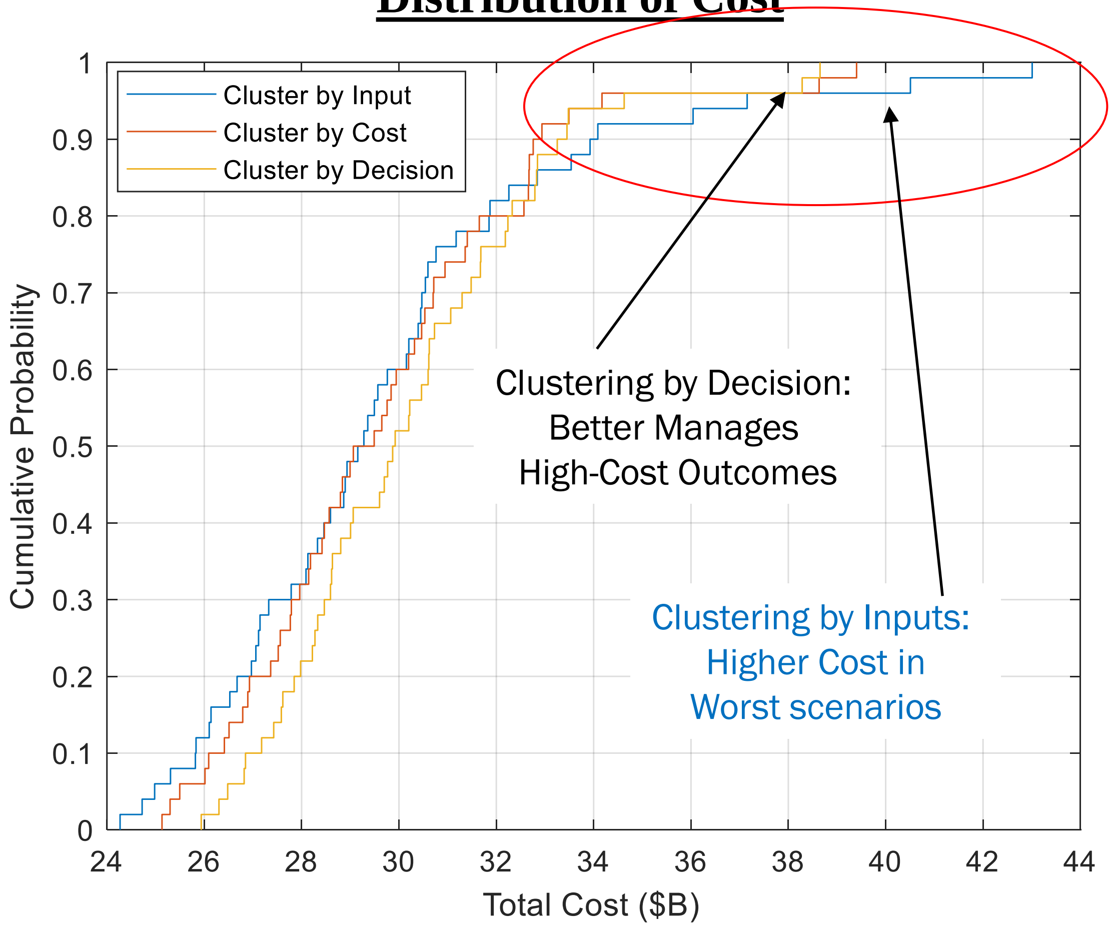


Clustering Methods: Impact on Cost, Risk, and Decisions

Stage 1 Investments



Distribution of Cost



Tradeoffs in Choosing a Clustering Method

- Computation time tradeoffs:
  - Cluster by Inputs:
    - No additional model runs needed for setup
  - Cluster by Cost:
    - Need solution from deterministic model for all scenarios, base system
  - Cluster by Decision:
    - Need solution from deterministic model for all scenarios, sample plans
- Given enough scenarios (clusters), any method works well
  - More clusters = more computation time for stochastic model