Statistical Learning Applied to Climate-Water-Energy Impacts Scenarios

Mort Webster (PSU), Duc Nguyen (Ascend Analytics), Jesse Bukenberger (EPRI),
Jonathan Celaya (PSU), Renee Obringer (PSU), Karen Fisher-Vanden (PSU)

Principal Investigators Meeting
Earth and Environmental Systems Modeling (EESM) Program
U.S. Department of Energy

August 7, 2024

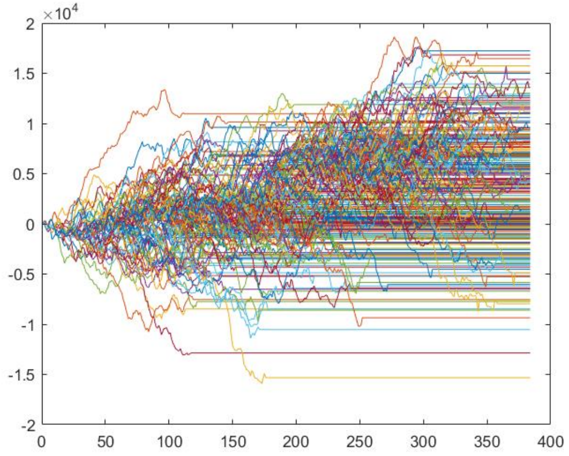
Managing Uncertainty: Scenarios

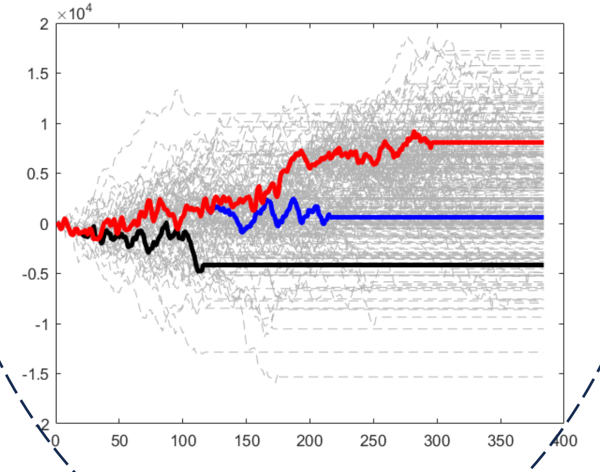
Scenario Generation

Scenario Reduction

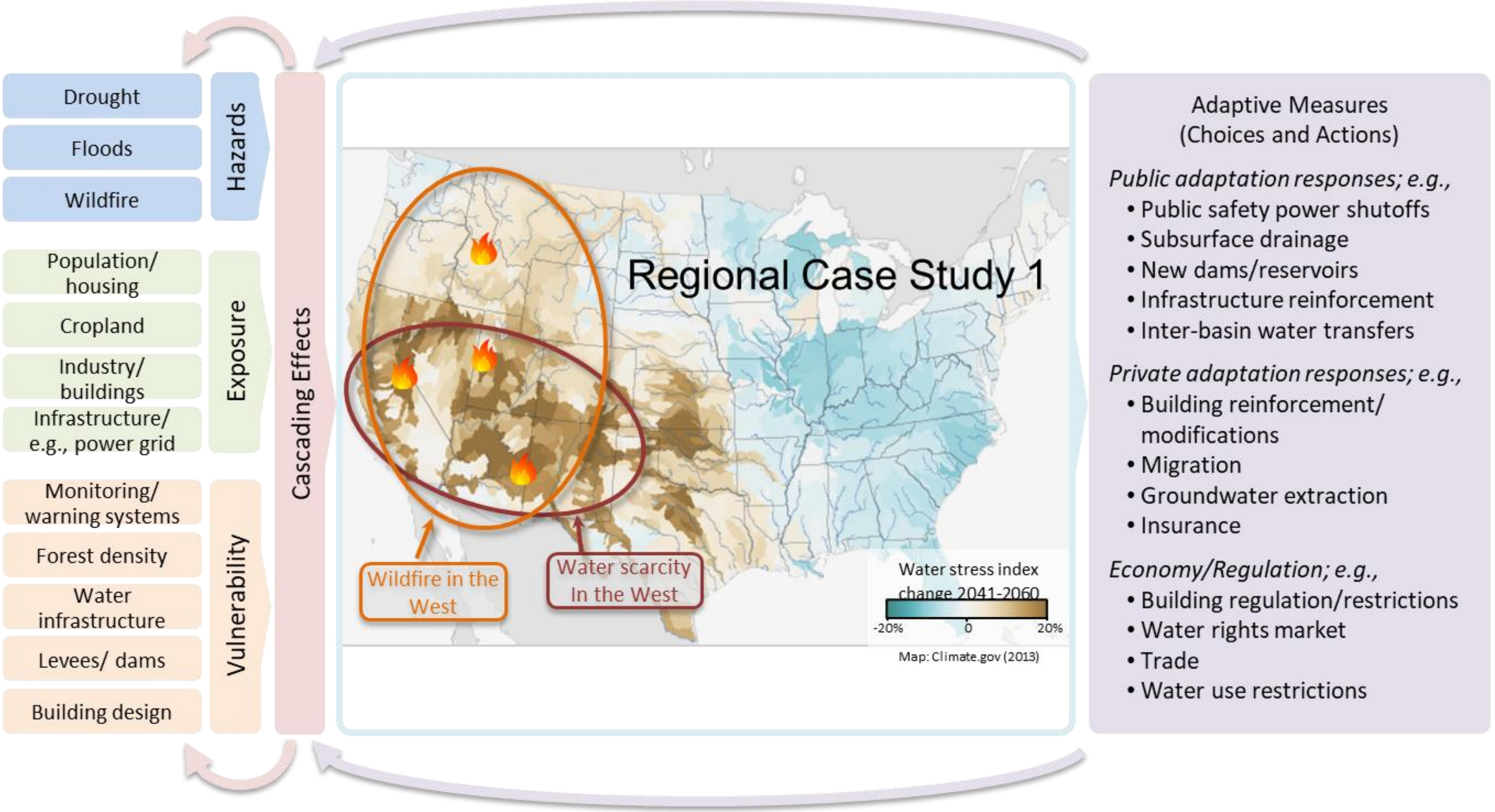
Risk Assessment: Simulation

Risk Management: Optimization





PCHES-ADAPT: visual project overview (Western US Case Study)



We seek to understand the **multisector dynamics (MSD) influences** on cascading risks associated with **hydrological stress** and **wildfire** in the U.S. West

How do **human and institutional responses** to drought and wildfire to heighten or mitigate risk to the MSD system?

To what extent can **adaptive and resilient strategies** mitigate these risks?

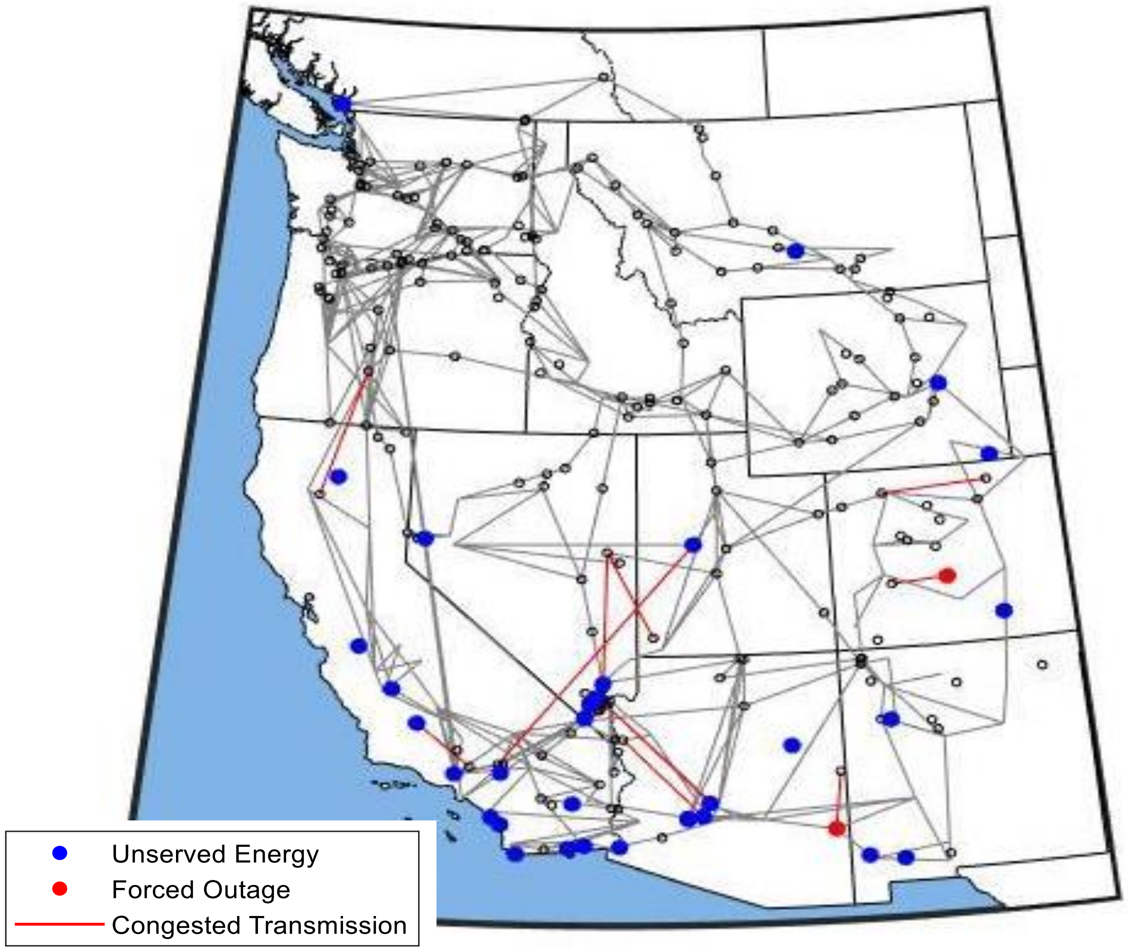
PennState · TEXAS · University of New Hampshire · Stanford University · Boston University · Cornell University · Illinois · Columbia · Purdue University · Dartmouth

Hydrology-Power-Economy Impact Example

Example:

- Future climate forcing
  - ➢ Higher surface water temperatures
- Higher water temperatures
  - ➢ Forced power plant outage
- Power rerouted
  - ➢ Transmission congestion
- Unmet electricity demand
  - ➢ Distant from the forced outages

Impacts, teleconnections for one future scenario



Hydrology-Power-Economy Impact Scenarios

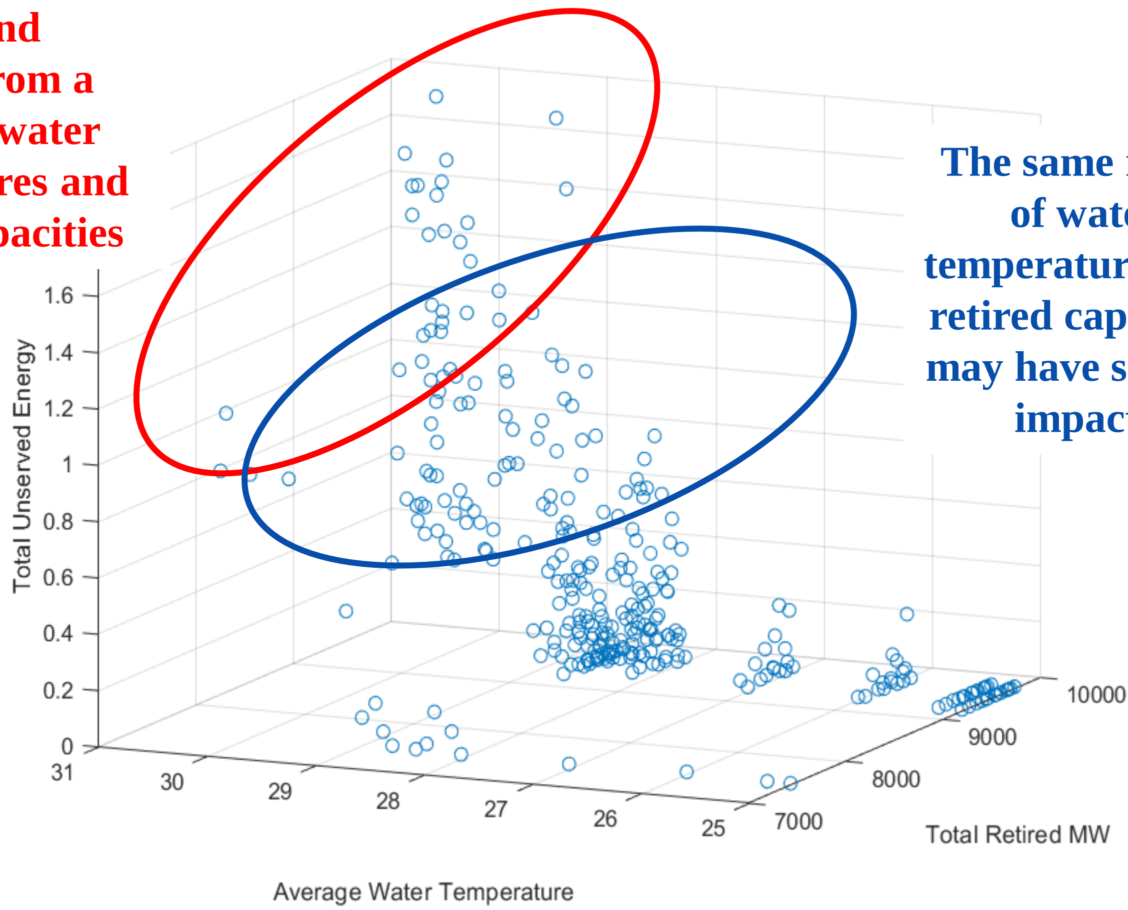
Infrastructure Resilience Planning:
- System operators must plan now:
    - Where to produce power
    - Where to add transmission lines

Uncertainty:
- Alternative futures must be planned for
- We can create many scenarios

Motivating Questions:
- How can I organize the many possible futures?
- Do I need to simulate every scenario?
- Do planners need to consider every scenario?

324 scenarios of future climate, hydrology, and changes to power grid

Greatest unserved demand results from a range of water temperatures and retired capacities



The same range of water temperatures and retired capacities may have smaller impacts

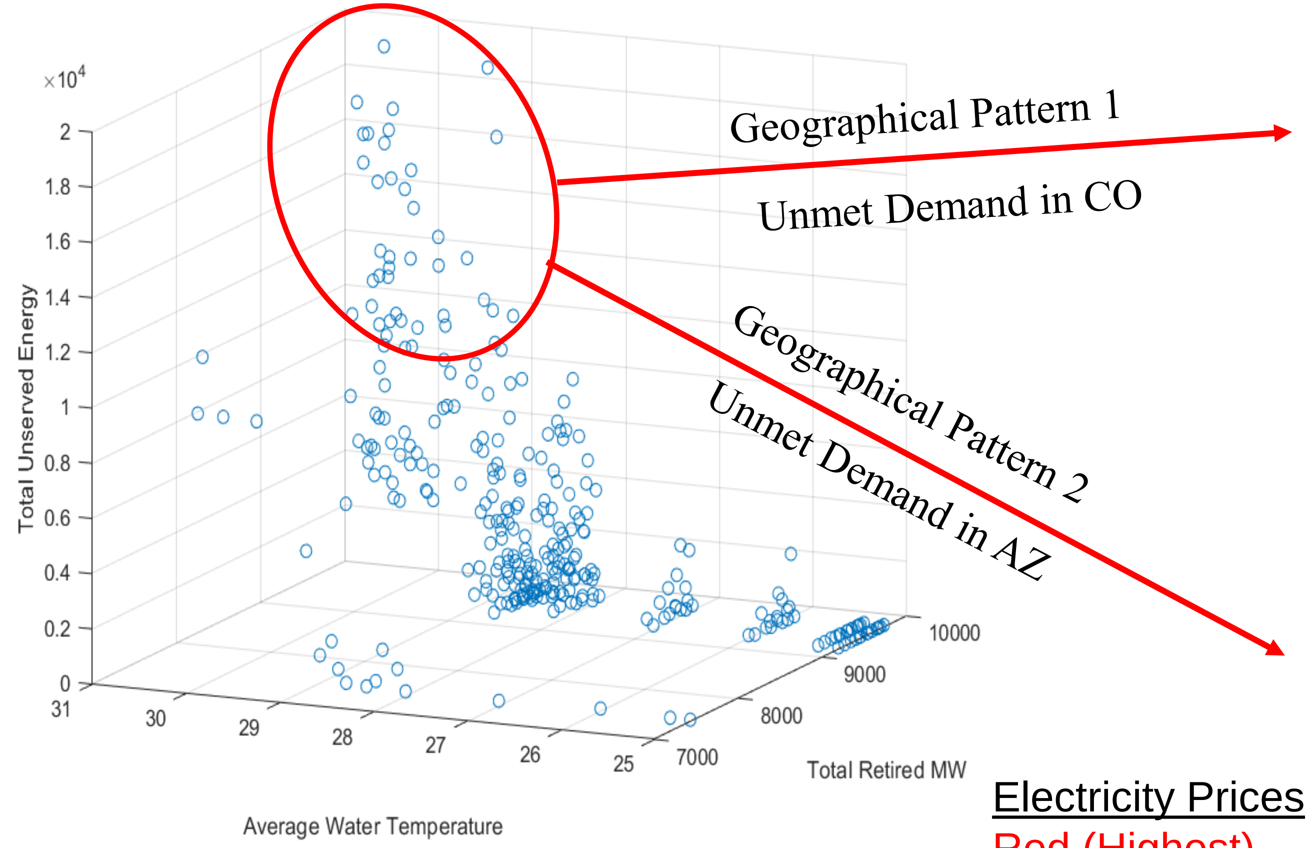
Step 1: Detect Common Patterns in Outcomes

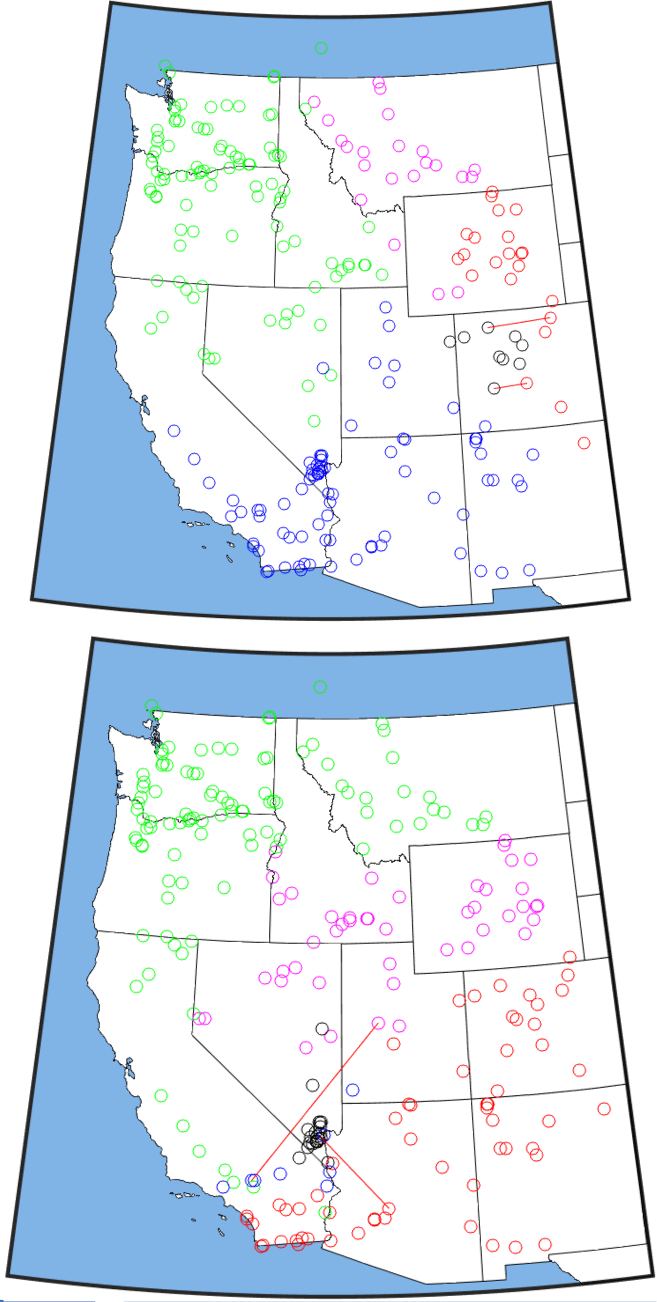
Apply *K*-means clustering to:

- Electricity Prices
- Transmission Congestion

Finding:

- Many scenarios of failure
- But only four distinct geographic patterns





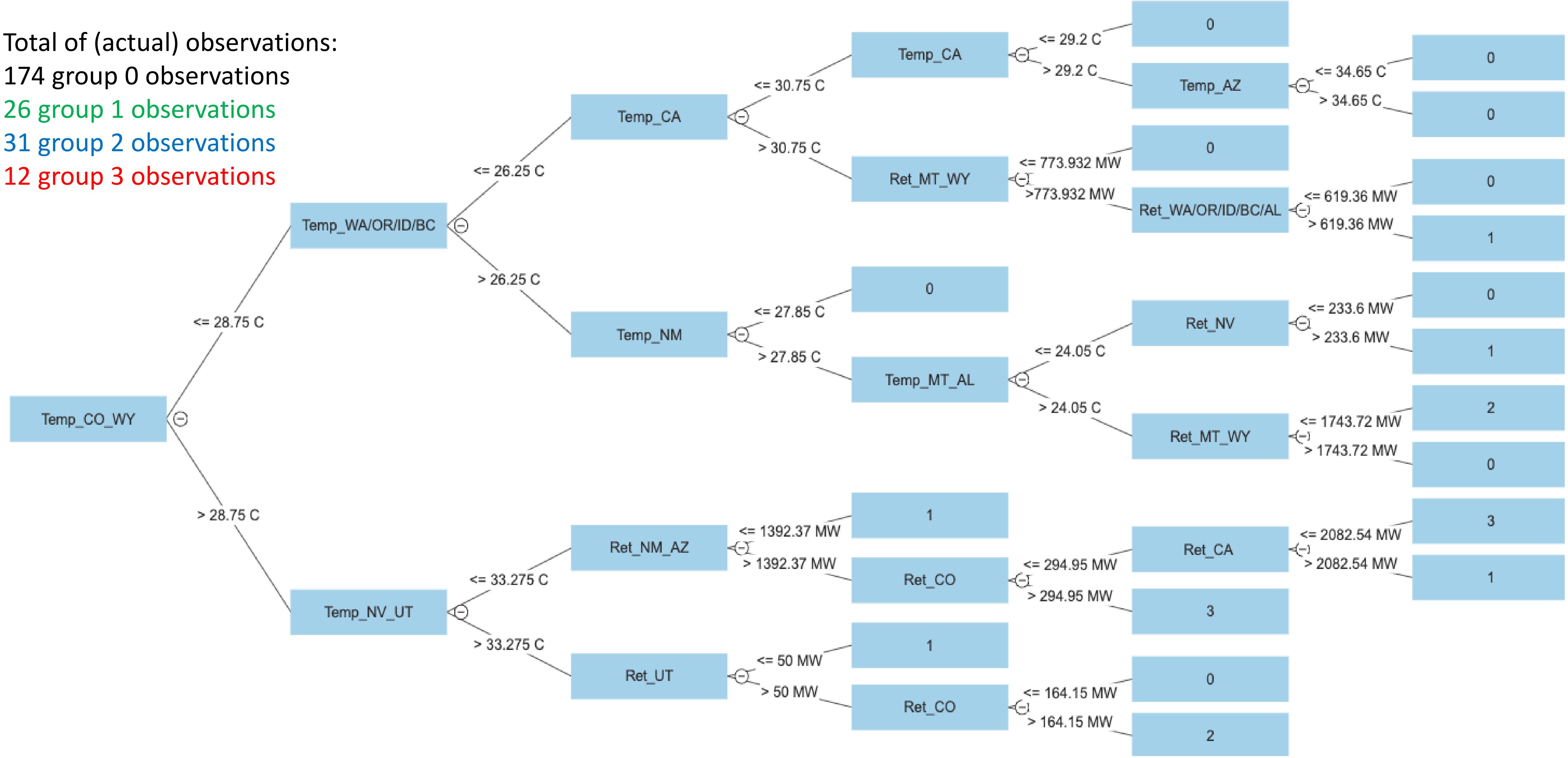
Electricity Prices
Red (Highest)
Pink
Green
Blue
Black (Lowest)

Step 2: Apply Statistical Models for Prediction

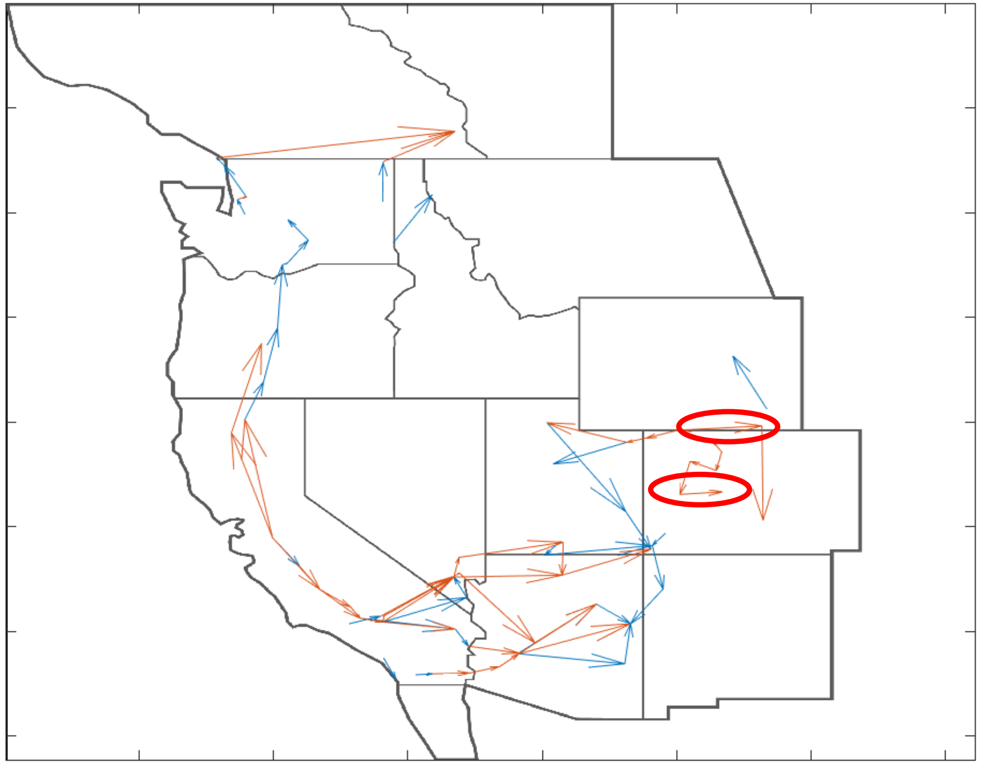
- **Next Question:**
  - How can I know which group a scenario's outcomes will belong to ***without*** simulating it?
- **Method:** Classification Tree
  - Partitions data set (all scenarios) by predictor data to differentiate groups (i.e., partition into 2 groups by one variable, then partition again based on another variable, etc.).
  - Because of randomness, ambiguity, no one tree will be a perfect predictor
- **Better Method:** Random Forest
  - Independently fit MANY classification trees
  - Use the weighted average of all trees to predict group from a scenario
- **Representative Tree**
  - Cannot visualize a random forest: good predictor but "black box"
  - Find a single tree that is minimum distance from all trees in random forest

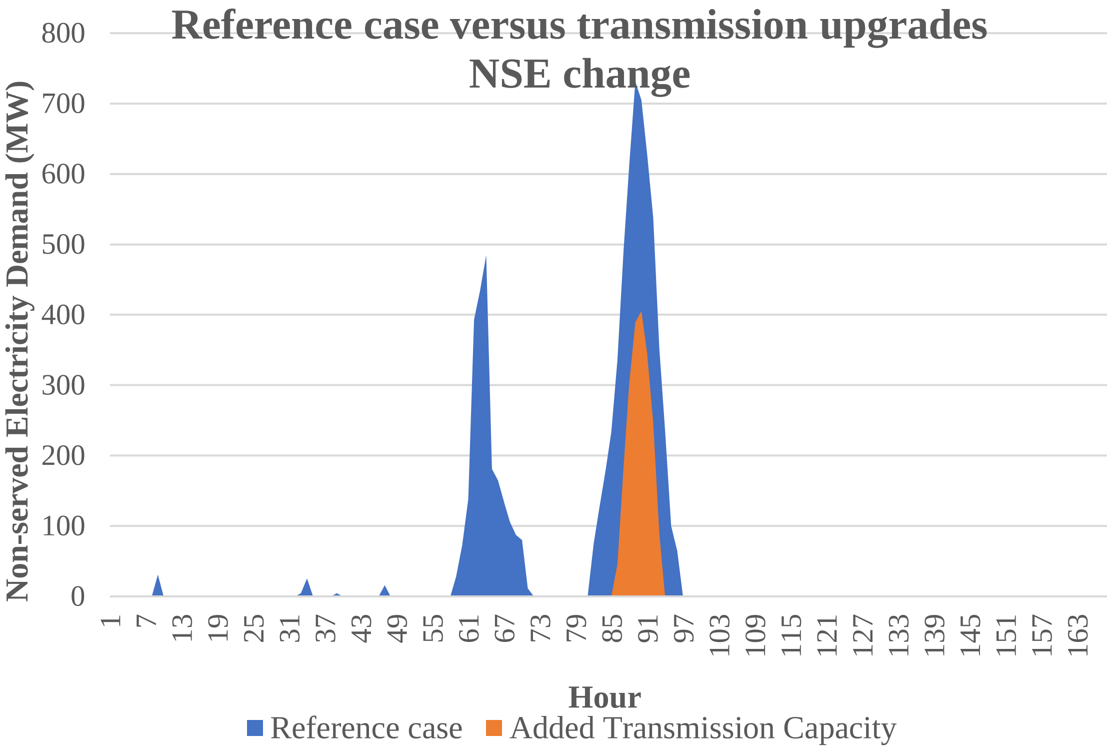
Example Classification Tree for 324 Scenarios

Total of (actual) observations:
174 group 0 observations
26 group 1 observations
31 group 2 observations
12 group 3 observations



Step 3: Test whether the Clusters are Useful for Adaptation





Example:

- Select one scenario from Group *N*
- Double transmission capacity for congested lines -> reduction in unmet electricity demand
- Test the same change for other scenarios in Group *N*
- Verify that the unmet energy is also reduced in these scenarios (yes, it is!).

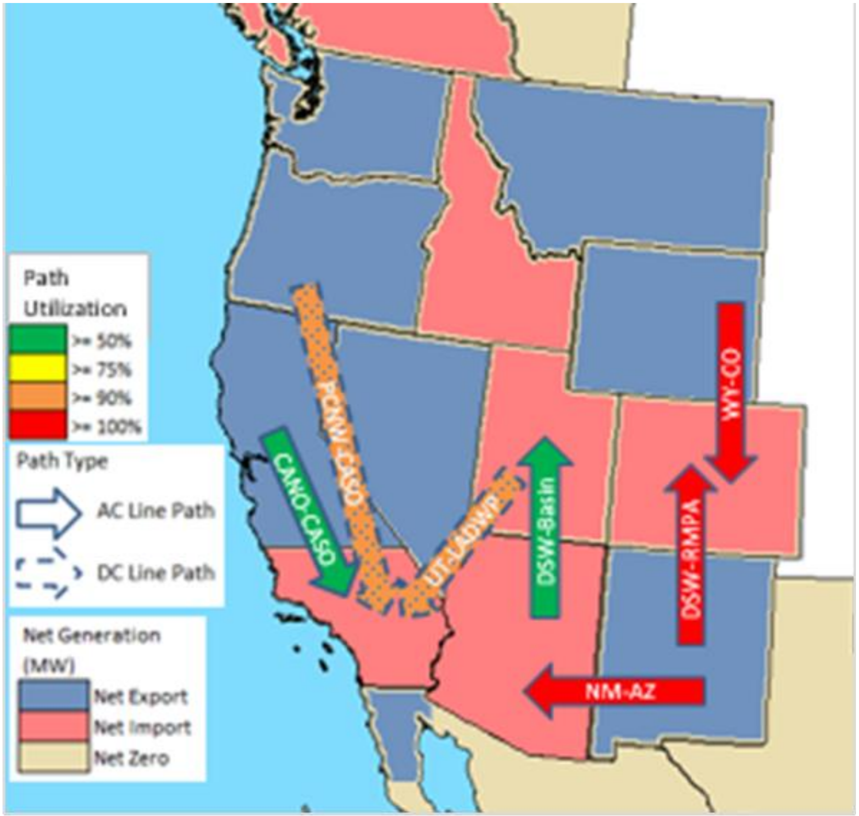
Quick Summary

Starting point:

- Unlimited number of high-resolution scenarios
- Easy to create scenarios
- Hard to simulate all of them in model
- Even harder to know what to do with the results afterwards

1. **How do I organize large set of scenarios?**
    - Cluster into groups with similar ***outcome*** patterns
2. **How can I predict which group a new scenario will belong to?**
    - Statistical learning tools (Random Forest)
3. **What to provide to infrastructure planners?**
    - Representative scenarios from each impact pattern group

Where will future transmission lines be?



Questions?

This work was supported by the U.S. Department of Energy, Office of Science, Biological and Environmental Research Program, Earth and Environmental Systems Modeling, MultiSector Dynamics under Cooperative Agreement **DE-SC0022141**.