Optimizing U.S Lithium Supply Chain Policies Using POMDPs

CJ Oshiro [1, 2]
cjoshiro@stanford.edu
coshiro@scu.edu

Mansur Arief [1]
ariefm@stanford.edu

Mykel Kochenderfer [1]
mykel@stanford.edu

[1] Aeronautics and Astronautics Department
Stanford University, Stanford, CA

[2] Computer Science and Mathematics Department
Santa Clara University, Santa Clara, CA

Problem

Developing a self-sufficient lithium supply chain is vital to the U.S. as the adoption of lithium-based clean energy technologies continues to increase. Currently, the volume of lithium that the U.S. imports greatly exceeds its domestic production. To become partially self-sufficient, the U.S. aims to primarily rely on international mines to satisfy lithium demand while building mining infrastructure and exploring domestic deposits. The objectives include:

- Reducing CO2 emissions to a minimum
- Satisfying lithium demand
- Making decisions when we are uncertain of the true volume of lithium contained in deposits.

Problem: Formulate the problem as a POMDP (Partially Observable Markov Decision Process and solve it to find the **optimal decisions** when mining for Lithium each year.

Background

A **Partially Observable Markov Decision Process** is a mathematical formulation that can be used to represent sequential decision-making problems. A POMDP can be formulated as the 7-tuple $(\mathcal{S}, \mathcal{A}, \mathcal{T}, \mathcal{R}, \mathcal{O}, \mathcal{Z}, \gamma)$, i.e., the *state space, action space, transition function, reward function, observation space, observation function,* and *discount factor.*

Motivation

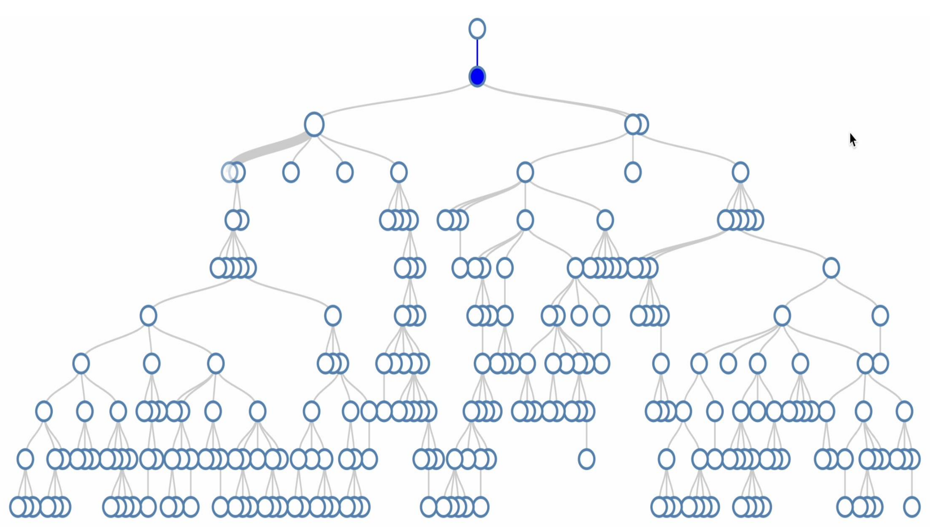
The desire for lithium self-sufficiency infrastructure is motivated by various factors:

- The adoption of **clean energy storage technologies** that require lithium increases demand.
- The volume of domestic lithium deposits is **uncertain** and substantial time is required to develop infrastructure.
- Acquiring **social licenses** to mine domestically is difficult, further complicating our decision whether to mine/not mine.

Methodology

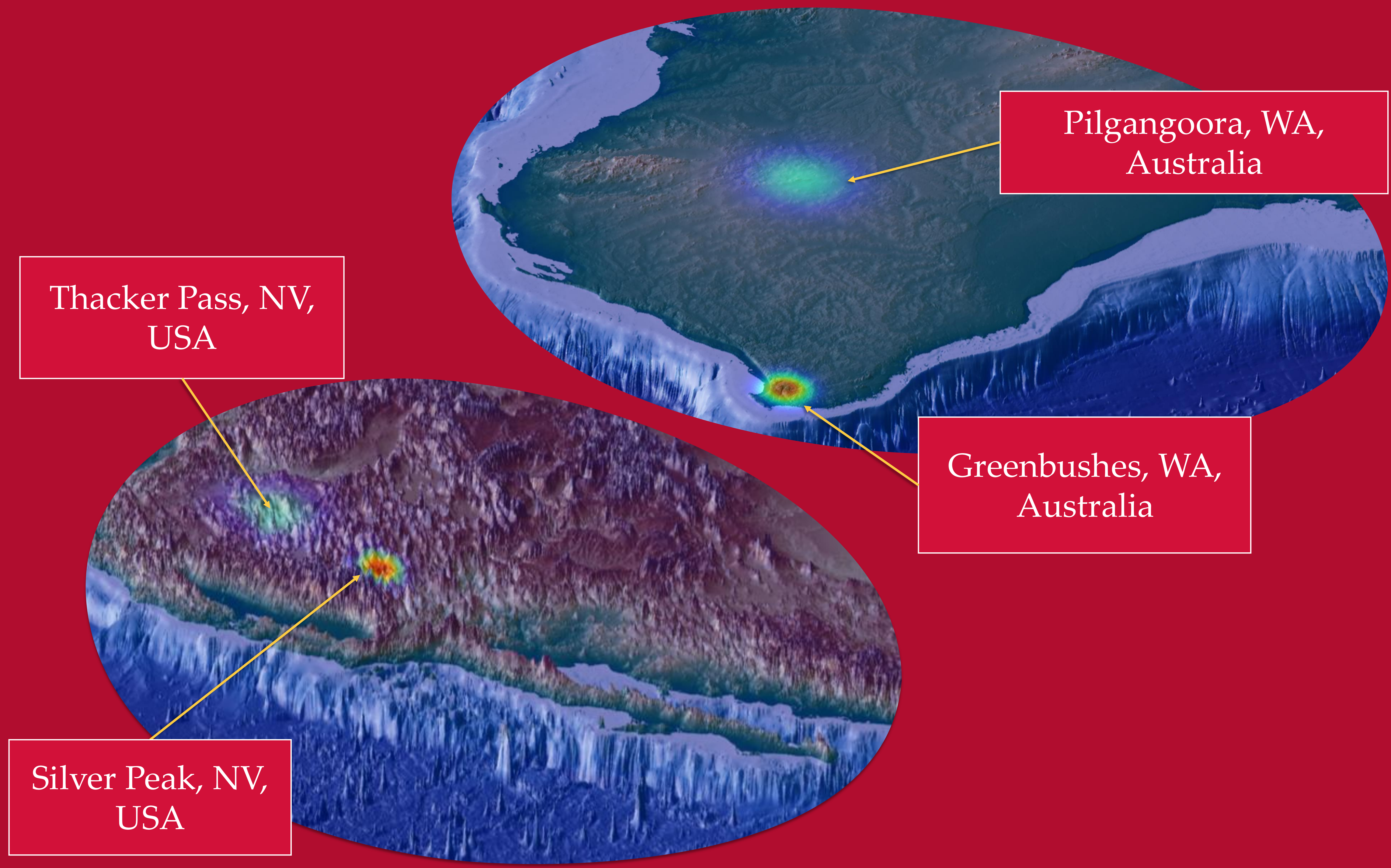
We built a model and formulated the problem as a POMDP. We solved the POMDP by using AI solvers, in particular, POMCPOW and DESPOT. POMCPOW utilizes **Monte Carlo tree search** which selects and explores the most promising nodes based on a balance between the **exploration** of new paths and **exploitation** of already explored paths with favorable outcomes. DESPOT captures the execution of all policies under a set of **sampled scenarios**. It finds the optimal policy by choosing the best action at every internal node based on the scenario it encounters.

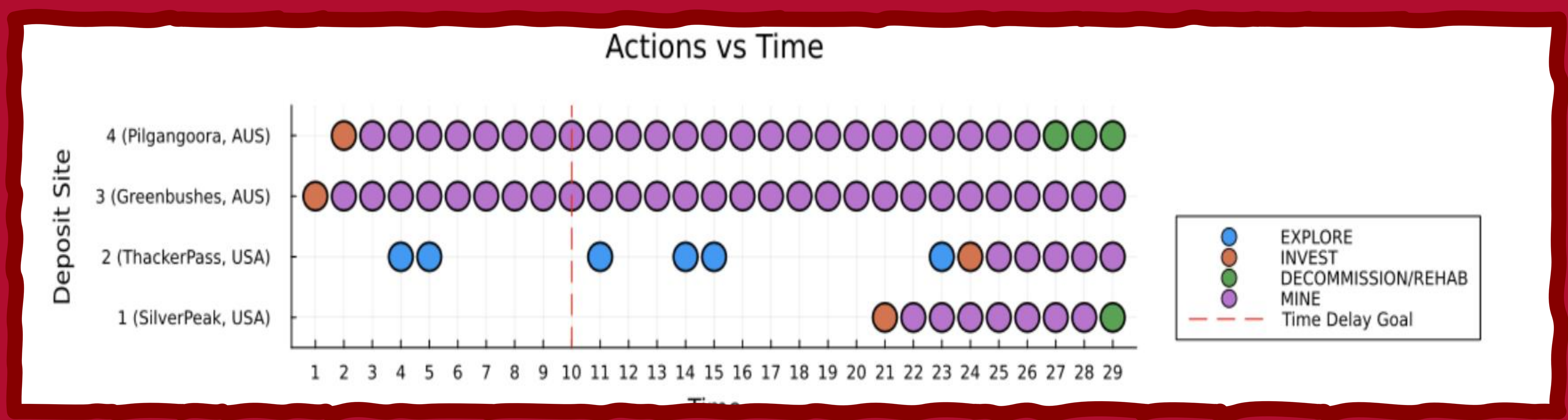
Figure 1: Illustration of Monte Carlo tree search



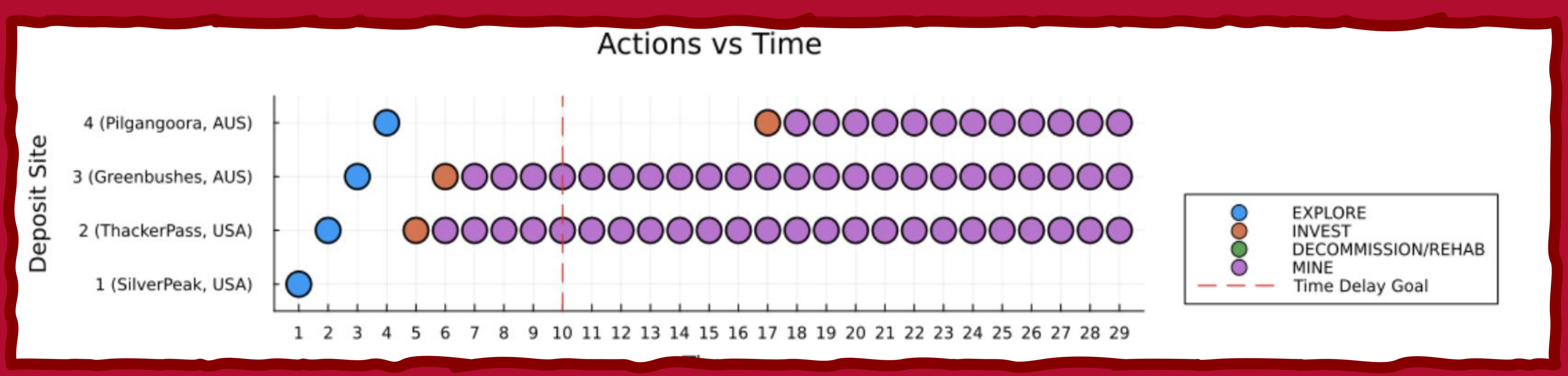
To develop a domestic and **self-sufficient** lithium **supply chain** for the U.S., we use **POMDPs** and their **solvers** to build an optimal policy.

For our analysis, we focus on four mines; two in the U.S. and two in Australia. We have varying beliefs over the volumes of each mine. Therefore, we represent our belief about the volume of each mine as a probability distribution, visualized through heat maps





POMCPOW Policy



Efficiency (W/ Uncertainty) Policy

POMDP Formulation

State Space $\mathcal{S}$:

- Collection of States, each comprised of a deposits vector - deposits, the current time t, the amount of Li mined domestically thus far V_t, the amount of Li imported thus far I_t, and vector have_mined of Boolean values indicating whether we have begun operation of a mine.

Action Space $\mathcal{A}$:

- MINE units of Li or EXPLORE to reduce uncertainty.

Transition Function $\mathcal{T}$:

- Describes the dynamics of how the current state changes when taking action *a* from state *s*
- When you MINE: subtract the appropriate volume from deposits, increase t by 1, increase V_t by 1, and update the have_mined vector if necessary.
- When you EXPLORE: increment t by 1.

Reward Function $\mathcal{R}$:

- Dependent on current state *s* and action *a*.
- 5 objectives: encourages time delay (r1), encourages Li volume maximization (r2), discourages CO2 emissions (r3), encourages annual demand satisfaction (r4), and encourages profit maximization (r5).
- Contains objective weights to determine the importance of each objective.

Observation function $\mathcal{O}$:

- When you MINE: no observation received
- When you EXPLORE: stochastically generate an observation.

Belief Updater:

- Utilizes a Kalman filter to more accurately generate beliefs of the true state despite the noisiness of our observations.

Results

Our simulations suggest that we can improve our policies regarding carbon dioxide and lithium volume production through AI solvers such as POMCPOW and DESPOT. We can specify the importance of each of our objectives to these algorithms through objective weights, signifying the respective priority of each objective.

Policy	Total CO2 Emitted
EmissionAwarePolicy	85.0 ± 23.1
RandomPolicy	72.3 ± 27.8
EfficiencyPolicy	84.3 ± 3.0
EfficiencyPolicy(w/ Uncertainty)	125.0 ± 11.1
MCTS-DPW	66.3 ± 24.2
POMCPOW	
Weights [0.5, 0.5, 0.5]	44 ± 7.0
Weights [0.9, 0.1, 0.1]	47 ± 8.6
Weights [0.1, 0.9, 0.1]	47.7 ± 13.5
Weights [0.1, 0.1, 0.9]	44.3 ± 16.0
DESPOT	
Weights [0.5, 0.5, 0.5]	20.7 ± 7.0
Weights [15.0, 0.1, 0.1]	19.3 ± 3.0
Weights [0.1, 15.0, 0.1]	25.6 ± 6.4
Weights [0.1, 0.1, 15.0]	22.0 ± 1.5

Policy	Volume Li Imported	Domestic Li Mined
EmissionAwarePolicy	25.0 ± 1.0	2.7 ± 2.5
RandomPolicy	8.0 ± 3.4	7.7 ± 4.6
EfficiencyPolicy	17.0 ± 2.0	10.7 ± 1.2
EfficiencyPolicy(w/ Uncertainty)	17.3 ± 2.0	10.3 ± 0.5
MCTS-DPW	7.0 ± 1.1	6.8 ± 3.4
POMCPOW		
Weights [0.5, 0.5, 0.5]	5.3 ± 1.5	8.7 ± 3.0
Weights [0.9, 0.1, 0.1]	7.6 ± 2.9	6.3 ± 2.9
Weights [0.1, 0.9, 0.1]	10.0 ± 2.0	4.3 ± 1.5
Weights [0.1, 0.1, 0.9]	7.3 ± 1.5	5.7 ± 3.0
DESPOT		
Weights [0.5, 0.5, 0.5]	9.3 ± 2.0	0.0 ± 0.0
Weights [15.0, 0.1, 0.1]	19.3 ± 1.5	0.0 ± 0.0
Weights [0.1, 15.0, 0.1]	10.3 ± 1.2	1.0 ± 1.7
Weights [0.1, 0.1, 15.0]	8.0 ± 1.0	5.7 ± 3.0

Figure 2: 3D Scatter Plot of Planner Performance by Emissions and Volume