Reflections on Environmental Decision Making

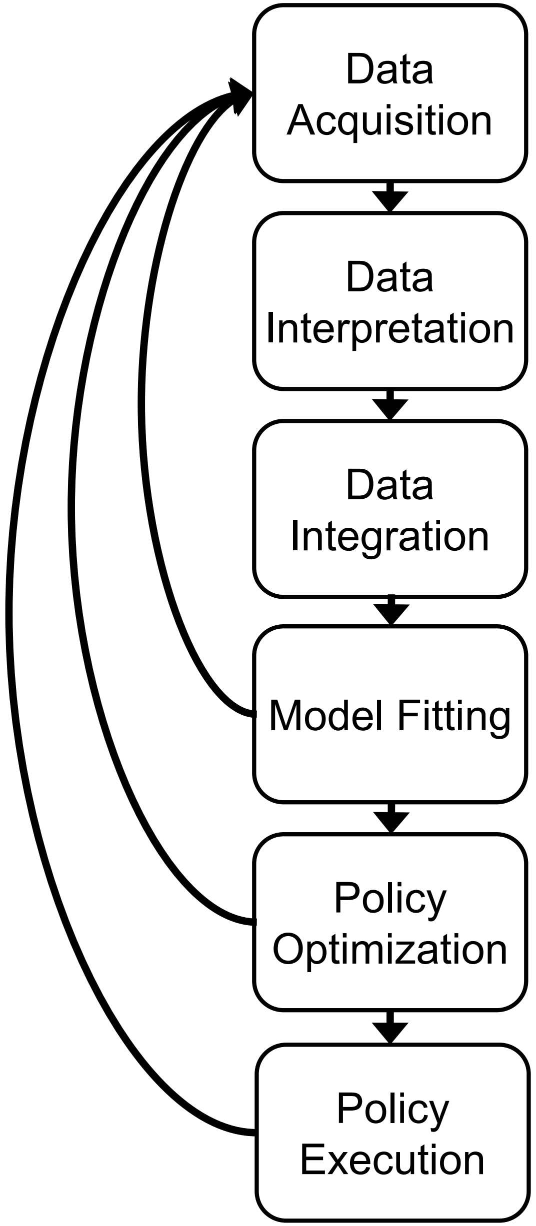
Tom Dietterich

Distinguished Professor Emeritus

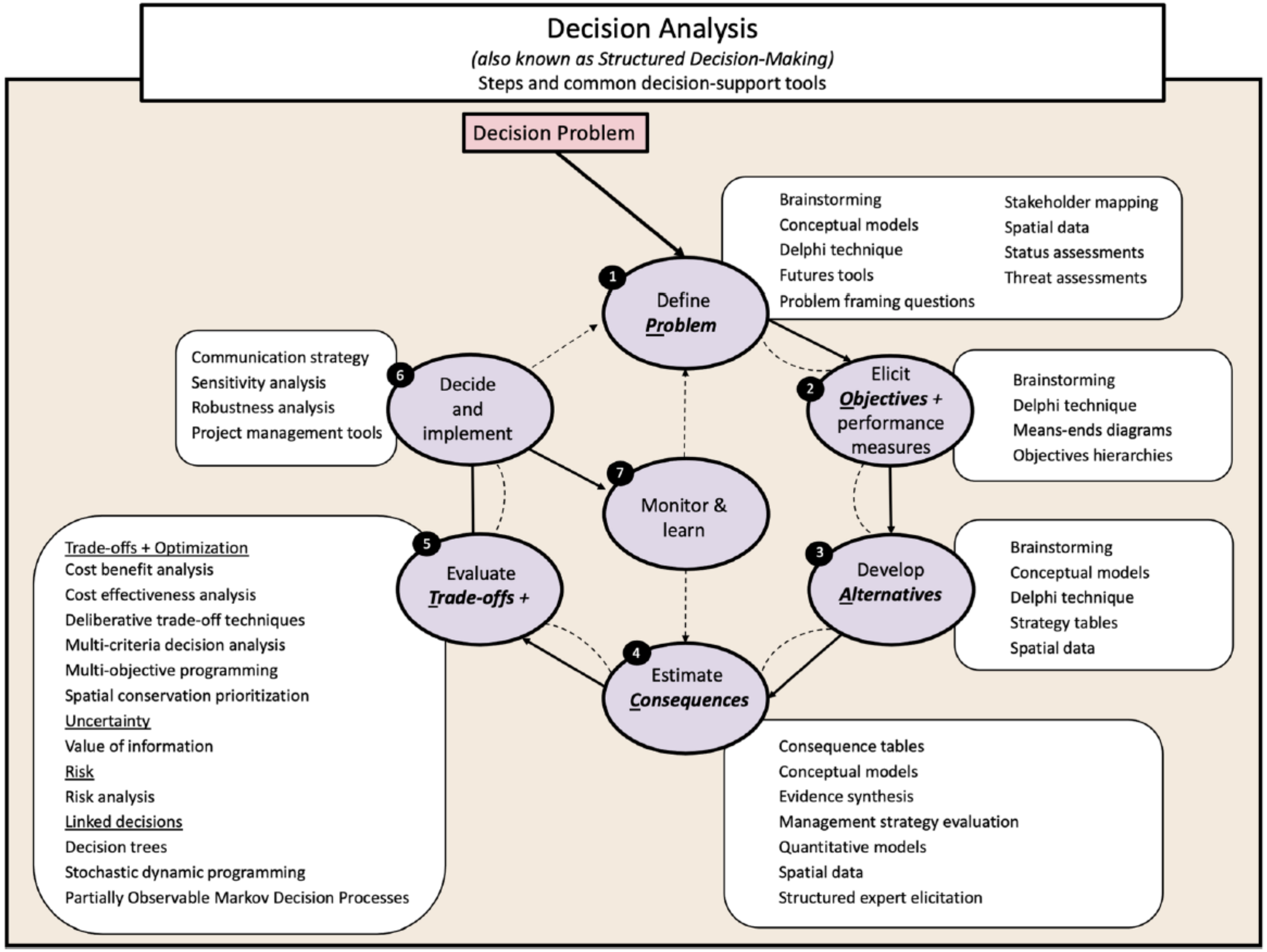
Oregon State University

My Naïve Pipeline View (2012)

- Everything starts with data collection
- Policy / decision making is tacked on at the end

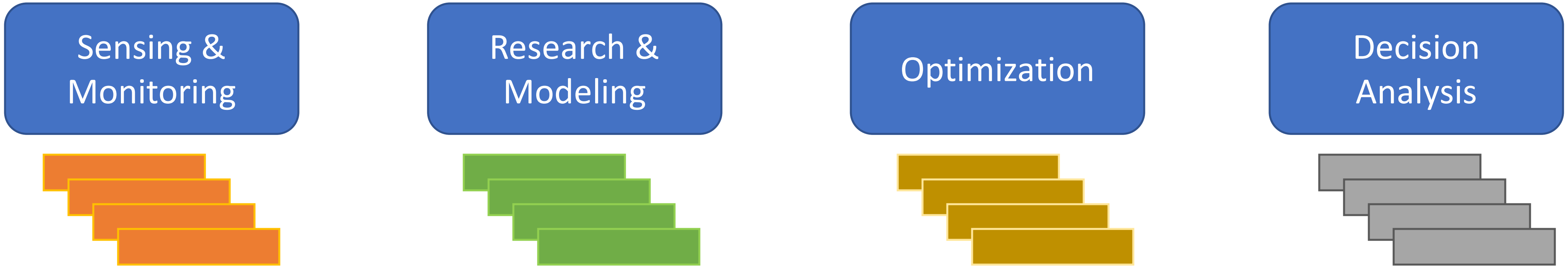


Opposite View: Focus on the decision problem



Hemming, et al., An introduction to decision science for conservation

A synthesis?



- Concurrent efforts:
  - Sensing: Modalities, locations, scale, observers, …
  - Research: Improving the science base (e.g., community dynamics, more effective/cheaper interventions)
  - Modeling: Advances in ML and statistics
  - Optimization: Robust, constrained, multi-objective, interactive speeds
  - Decision analysis: A network of decisions at multiple scales in time, space, and organization
- Information should flow within and between these activities

Three Projects and Lessons Learned

- Project 1: Invasive species management: Tamarisk
  - Lesson: Chance constraints provide an interesting alternative to species valuation
- Project 2: Forest fire management: LETBURN
  - Lesson: Visualization driven by interactive optimization can potentially help multiple stakeholders refine their decision options
  - Lesson: We can provide conformal guarantees for system trajectories
- Project 3: Forest fire liability rules
  - Lesson: Multi-agent reinforcement learning (MARL) provides a tool for policy analysis

Project 1: Controlling Tamarisk invasion in river networks

- In most environmental decision problems the model contains two components:
  - Biological model
  - Economic model
- Quantifying cost for the economic model is typically easy
- But it is difficult to assign dollar values to some biological models
  - Ecosystem services is an attempt to get around this, but it is not a complete solution
  - What is the dollar value of a species extinction? The cultural value of an ecosystem?
- Another potential solution: Constrained optimization
  - Minimize *economic cost* subject to a *chance constraint*
  - Probability of species extinction in 100 years is ≤ 0.01
- We now have a rich set of algorithms for optimizing Constrained-MDPs (CMDPs)
  - See Gattami et al. (2021)

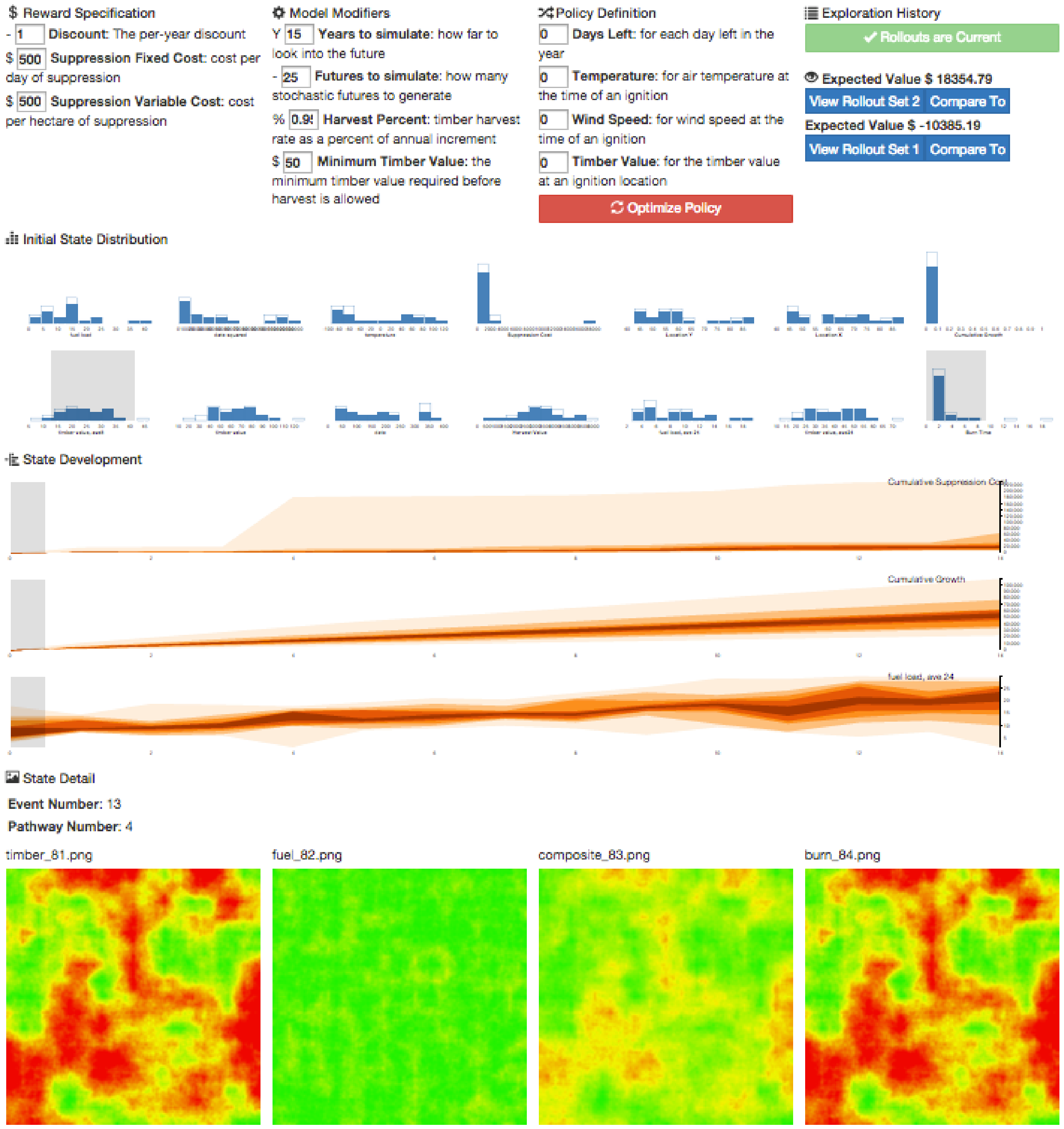
Project 2: LETBURN Decision Problem

- When lightning ignites a wildland fire, should we
  - SUPPRESS
    - Benefits realized immediately, but may increase long-term risks
  - LETBURN
    - Losses realized immediately, but may reduce long-term risks
- Stakeholders:
  - Home owners
  - Timber companies
  - Politicians
- Agreeing on policies can be easier than agreeing on reward functions
  - Interactive optimization-assisted search for acceptable policies

Constituency	Suppression Costs	Timber Revenues	Ecology Target	Air Quality	Recreation Target
Composite	✓	✓	✓	✓	✓
Politics	-	✓	✓	✓	✓
Home Owners	-	-	-	✓	✓
Timber	✓	✓	-	-	-

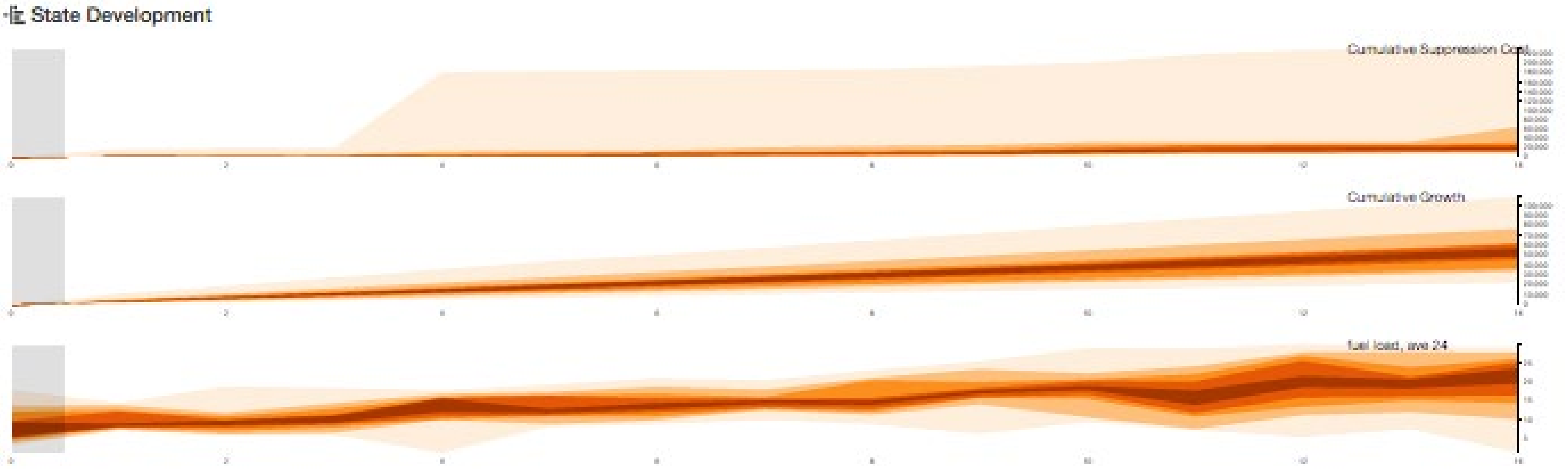
Visualization of rollout quantiles

- User selects a few state variables
- User either
  - Defines a policy
  - Defines a reward function and invokes RL to optimize it
- System generates 1000 rollouts and visualizes various quantiles for the chosen state variables



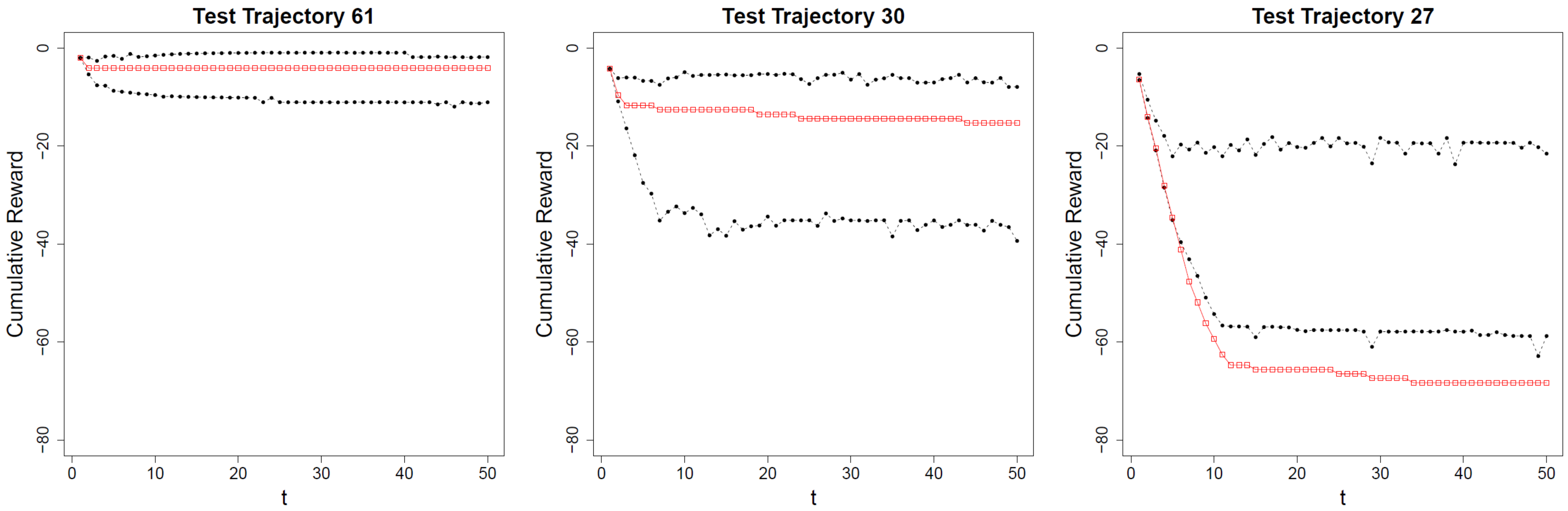
MDPvis (McGregor, et al. 2017)

Quantile Visualization



Time steps (years)

Recent Work: Conformal Guarantees on Trajectories
(D & Hostetler, arXiv 2206.04860; D & Guyer, forthcoming)



Project 3: Forest Fire Liability Rules

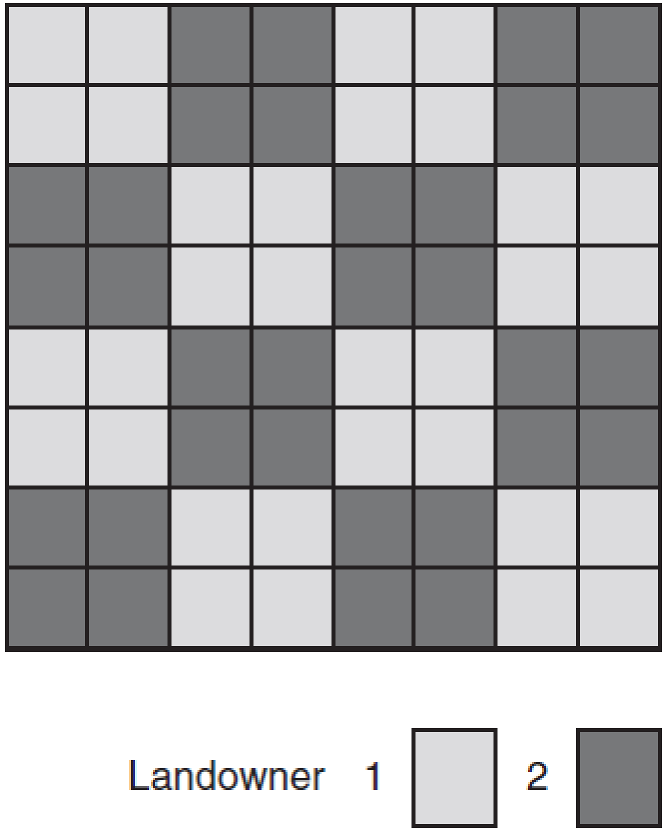
(Lauer, Montgomery, D. (2020))

- Setting
  - Multiple land owners
  - Fire starts on the property of owner A and burns property of owner B
  - Who should pay for the damage to owner B?
- Candidate regulations:
  - **NoReg**: Each owner is responsible for any damage that occurs on their property
  - **Strict**: A must pay B according to estimated timber value lost
  - **Negligence Standard (Neg)**: A pays nothing if A has maintained their property to a basic level of fire safety (fuel levels)
- Apply multi-agent reinforcement learning to compare expected value and variance for each agent and for the total welfare
  - Compare against a "social welfare" single agent land manager
- Note:
  - Landowner behavior changes under the different rules
  - Higher fire risk ➔ harvest trees earlier ➔reduced carbon storage and wood products

Land ownership patterns in western Oregon

- Checkerboard pattern from 19^{th} C railroad grants
  - Government is landowner 1 (BLM)
  - Timber companies (or investor group) is landowner 2

- Bio-economic model
  - Timber value and optimal harvest age policies are well-understood
  - Fire fuel models and fuel reduction treatments are fairly well-understood
  - Lightning ignitions (historical data is available)



Results

- Negligence Rule is almost as good as the Social Planner
  - Harvest ages increase vs. NoReg and Strict
- Variance increases for the landowners because they don't pay for some fires
  - Variance: 1.89→3.84
- Strict liability is worse than no regulation
  - Optimal owner behavior is "do nothing" because probability of fire starting on their land AND burning onto the other landowner's property is small

	Landscape		Landowner 1		Landowner 2	
	$mean\ V_r$	$\bar{V}$	$mean\ V_r^{n=1}$	$\bar{V}^{n=1}$	$mean\ V_r^{n=2}$	$\bar{V}^{n=2}$
Soc	15.08	15.14	7.19	7.25	7.88	7.88
	(7.23)		(1.89)		(1.93)	
NoReg	14.52	14.61	6.93	6.94	7.59	7.66
	(7.35)		(1.82)		(1.99)	
Strict	14.30	14.60	6.76	6.99	7.55	7.60
	(8.46)		(4.85)		(3.73)	
Neg	15.03	14.91	7.16	7.21	7.87	7.70
	(7.00)		(3.84)		(2.85)	

Summary

- Project 1: Invasive species management: Tamarisk
  - Lesson: Chance constraints provide an interesting alternative to species valuation
- Project 2: Forest fire management: LETBURN
  - Lesson: Visualization driven by interactive optimization can potentially help multiple stakeholders refine their decision options
  - Lesson: We can provide conformal guarantees for system trajectories
- Project 3: Forest fire liability rules
  - Lesson: Multi-agent reinforcement learning (MARL) provides a tool for policy analysis

References

- Dietterich, T. G., & Hostetler, J. (2022). Conformal Prediction Intervals for Markov Decision Process Trajectories. *ArXiv*, *2206.04860*(v2). http://arxiv.org/abs/2206.04860
- Gattami, A., Bai, Q., & Aggarwal, V. (2021). Reinforcement Learning for Constrained Markov Decision Processes. *AI STATS 2021*.
- Hemming, V., Burgman, M., Chalifour, L., Garrard, G. E., Camaclang, A. E., Adams, M. S., Carbeck, K., Carwardine, J., Chadès, I., Converse, S. J., Finn, R., Davidson, L. N. K., Huard, J., Fleri, J. R., Mayfield, H. J., Possingham, H. P., Tulloch, V. J. D., Madden, E. M., Rumpff, L., … Martin, T. G. (2021). An introduction to decision science for conservation. *Conservation Biology*, *36:e13868*, 1–16. https://doi.org/10.1111/cobi.13868
- Lauer, C. J., Montgomery, C. A., & Dietterich, T. G. (2020). Evaluating wildland fire liability standards-does regulation incentivise good management? *International Journal of Wildland Fire*, *29*(7), 572–580. https://doi.org/10.1071/WF19090
- McGregor, S., Buckingham, H., Dietterich, T. G., Houtman, R., Montgomery, C., Metoyer, R. (2017). Interactive Visualization for Testing Markov Decision Processes: MDPvis. *Journal of Visual Languages and Computing, 39,* 93–106 .

Project 2a: Conformal Trajectory Guarantees

- Given:
  - A set of rollouts generated by sampling a starting state $s_0 \sim P_0(s_0)$ and then following a fixed policy π for T time steps
  - A behavior function $b(s)$ that returns a real-valued quantity
- Output:
  - A "tube" $[\mathbf{b}_{\boldsymbol{lo}}, \mathbf{b}_{\boldsymbol{hi}}]$ from $t = 0, \dots, T$ such that with probability $1 - \delta$ over $P_0(s_0) \prod_t P(s_t | s_{t-1}, \pi)$ the true trajectory of b will satisfy

$$\mathbf{b}_{lo}(t) \leq b(s_t) \leq \mathbf{b}_{hi}(t) \ \ \forall t \in [0, T]$$