Managing Invasive Species in a River Network

Kim Hall, Majid Alkaee Taleghan, H.J.Albers,
Mark Crowley, Tom Dietterich
Oregon State University

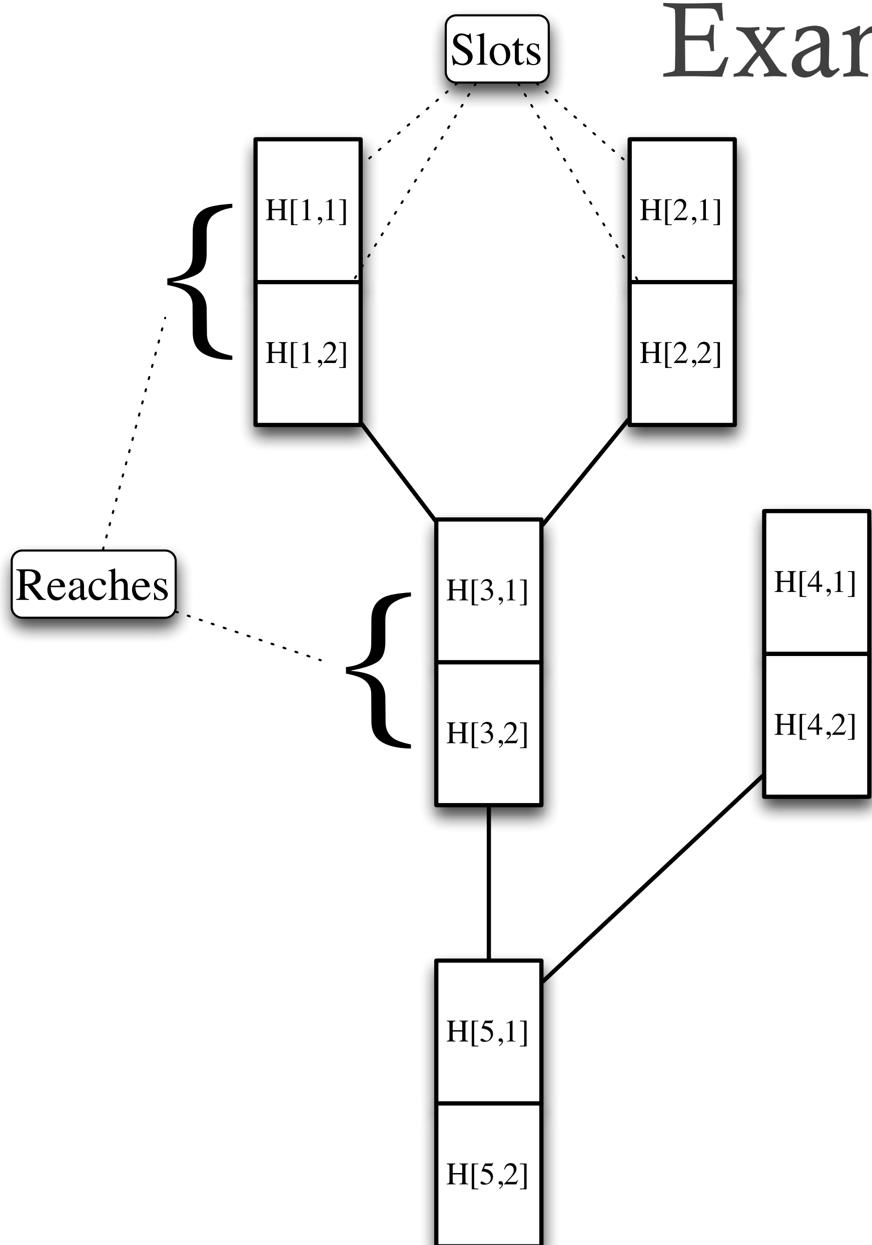
CompSust at OSU

- eBird – learning species distribution patterns from citizen science data
- BirdCast – predicting bird migration patterns
- BudID – automated categorization of bugs from image data
- wildlife corridor planning
- data cleaning for forest sensor networks
- forest fire control
- **controlling invasive species**

The Problem

- Managing plant ecosystem along a river network
- Competing native and invasive plant species
- Native and Invasive species spread dynamics
  - Local, spatial, stochastic
- Optimize for best outcome subject to budget constraints

Example River Network



- Slot States
  - Native plants, invasive plants, empty
- Actions in each Reach
  - Eradicate invasive plants
  - Eradicate and restore (replant) natives
  - Do nothing

States and Actions

- Slot States
  - Native plants, invasive plants, empty
- Actions in each Reach
  - Eradicate invasive plants
  - Eradicate and restore (replant) natives
  - Do nothing

Where We Fit In

- **Ecology :** focus on biological processes, postulate complete eradication (may be economically infeasible)

- **Economics :** focus on optimal control policy
  - Spatial spread often ignored/simplified greatly
  - Steady state analysis of spatial spread

- **Econ + CS :** Collaboration between Ecosystem Economics and Computer Science
  - Spread modelled as a conditional, spatial, stochastic process
  - Optimized as an MDP
    - Structure of the problem presents interesting computational challenges

Goal

Optimal policy describing placement of management actions over space and time

Economic Optimization

- Objective: reduce presence of invasive plants while minimizing costs
- Subject to annual management budget constraints and ecological processes

Optimization Problem

$$\min_{a_{it}} \sum_t \sum_i \gamma^t c_{it}(n_{it}, a_{it})$$

$$\text{s.t.} \sum_i c_{it}(n_{it}, a_{it}) \leq b_t \forall t$$

s. t. ecological model holds

Where:

γ Discount factor

c_{it} Cost function

n_{it} Invasive population size

a_{it} Management action

b_t Budget constraint

Markov Decision Process

States $s \in S$

Actions $a \in A$

Transition Dynamics $\mathcal{T}(s'|s,a) : \mathrm{S} \times \mathrm{A} \rightarrow \delta(\mathrm{S})$

Rewards $r(\mathbf{s}^t, \mathbf{a}^t) : \mathbf{S} \times \mathbf{A} \rightarrow \Re$

Discount Factor γ

Policy $\pi(a|s) : S \rightarrow \delta(A)$

Size of States and Actions

- Number of States and Actions are exponential in the network size
  - N is the number of states each slot can take on
  - M is the number of actions available in each reach

$$|S| = \left[\frac{(N+H-1)!}{H!(N-1)!}\right]^R$$

$$|A| = M^R$$

Number of Reaches (R)	Number of Slots (H) per Reach	Number of Actions: |A|	Number of States : |S|	Transition Model Size: |S|x|S|x|A|
3	2	27	216	1.0×10^{7}
5	2	243	7,776	1.4×10^{10}
3	3	27	1,000	2.7×10^{7}
5	3	243	10,000	2.4×10^{12}

Dynamics

- **Mortality**
  - Plants in each slot die with independently with probability *d*
  - Eradication of invasive plant, could fail stochastically
- **Propagule Generation :** Each surviving plant produces *g* propagules deterministically
- **Propagule Dispersal :** upstream or downstream with
  - P(arrive at reach *j* | started at reach *i*)
- Site **competition/colonization** at slot *h*:
  - If *h* is occupied, no effect
  - Else propagules compete to colonize slot, bias β in favour of invasive plants

Network dispersal model from Muneepeerakul et al. 2007

Dynamics Models

Dispersal Model

Probability of propagule leaving reach *i* and arriving in reach *j* is proportional to the rate of propagule survival upstream/downstream and the distance travelled

$$K_{ij} = Cu^{NU_{ij}} d^{ND_{ij}}$$

Competition Model

Probability that species k wins in slot s is equal to k's proportion of the total number of propagules arriving in slot s modified by a weighted factor β.

$$p_{\text{invasive}} = \frac{\beta g_{\text{invasive}}}{\beta g_{\text{invasive}} + g_{\text{native}}}$$

$$p_{\text{native}} = 1 - p_{\text{invasive}}$$

Dispersal Model

Probability of propagule leaving reach *i* and arriving in reach *j*

$$K_{ij} = Cu^{NU_{ij}} d^{ND_{ij}}$$

Where:

- C – is a normalization constant
- u – is the upstream propagule survival rate
- d – is the downstream propagule survival rate
- NU_{ij} – is the number of upstream reaches between reach i and j
- ND_{ij} – is the number of downstream reaches between reach i and j

Competition Model

$$p_{\text{invasive}} = \frac{\beta g_{\text{invasive}}}{\beta g_{\text{invasive}} + g_{\text{native}}}$$

$$p_{\text{native}} = 1 - p_{\text{invasive}}$$

- p_{species} - Probability that species wins
- g_{species} - Number of propagules of species
- β - "competitive advantage" of an invasive seed versus a native seed
  - 1.0, 1.5, 2.0…

Estimating the Transition Model

- It's easy to write a simulator for drawing samples of the dispersal and competition processes
- But computationally intractable to compute the exact transition probabilities $\mathcal{T}(S'|S, A)$
  - Estimate transition probabilities by drawing a large number of samples from the simulator

Error Bounds on Transition Model

$$Pr\left(\max_{s'}\left|\mathcal{T}(s'|s,a)-\mathcal{T}(s'|s,a)\right|<\epsilon\right)>1-\delta$$

- Confidence interval with width of ε
- 1-δ probability of being within interval
- This is a very loose bound : $|S|$ is large
- Future Work:
  - Tighter bounds that account for missing states from simulations
  - Approximate algorithms with PAC guarantees on bounds

Optimization

- Once estimate of $\hat{\mathcal{T}}(S'|S, A)$ is obtained perform Value Iteration on action-value function $Q^*(s, a)$

Interpreting the Policy

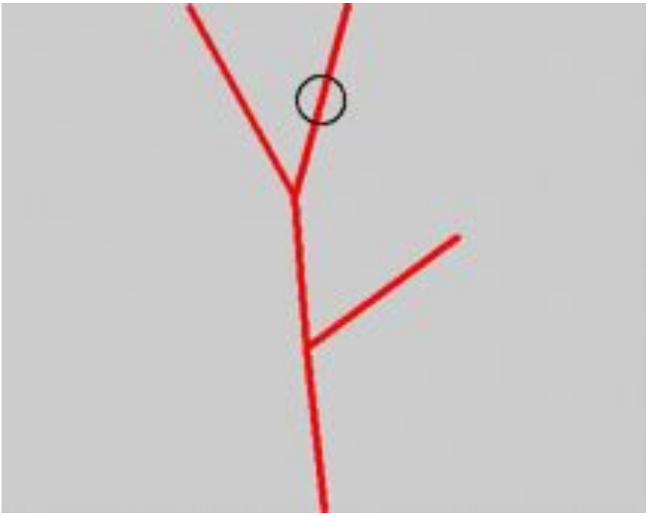
- Direct Examination of Optimal Policy
- Run optimal policy forward – collect stats from many simulated trajectories
  - Time to reach steady state
  - Frequency with which completely invaded
  - Frequency with which uninvaded states are reached
- Future Work

Comparing to Rule of Thumb Policies

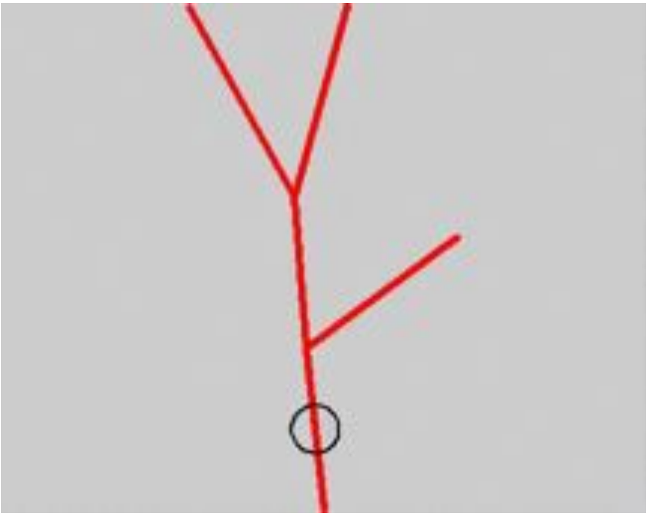
- Managers and Ecology/Economics Literature suggest:
  - **Triage**: treat most invaded reaches first
  - **Chades**, et al.: upstream first; extreme nodes first (one reach treated per period)
  - Treat **leading edge** of spread

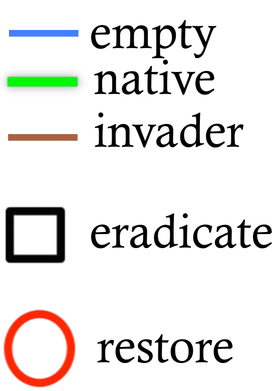
Comparing policy pathways

Chades

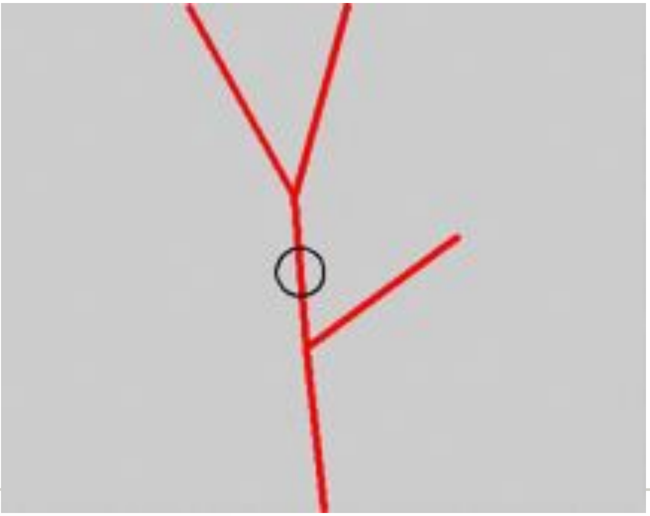


Leading edge

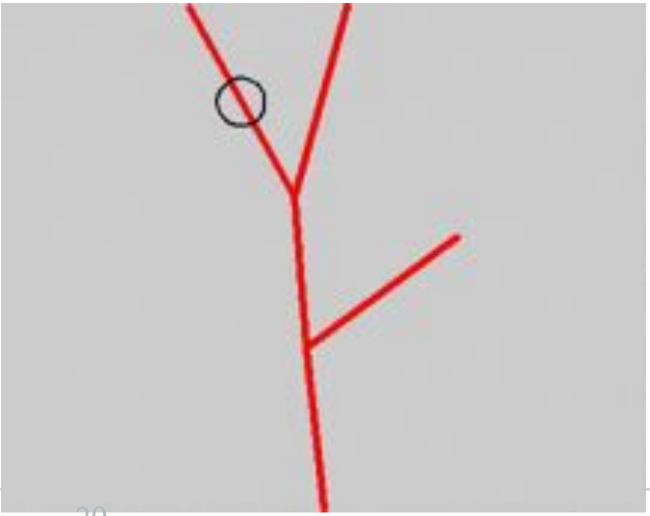




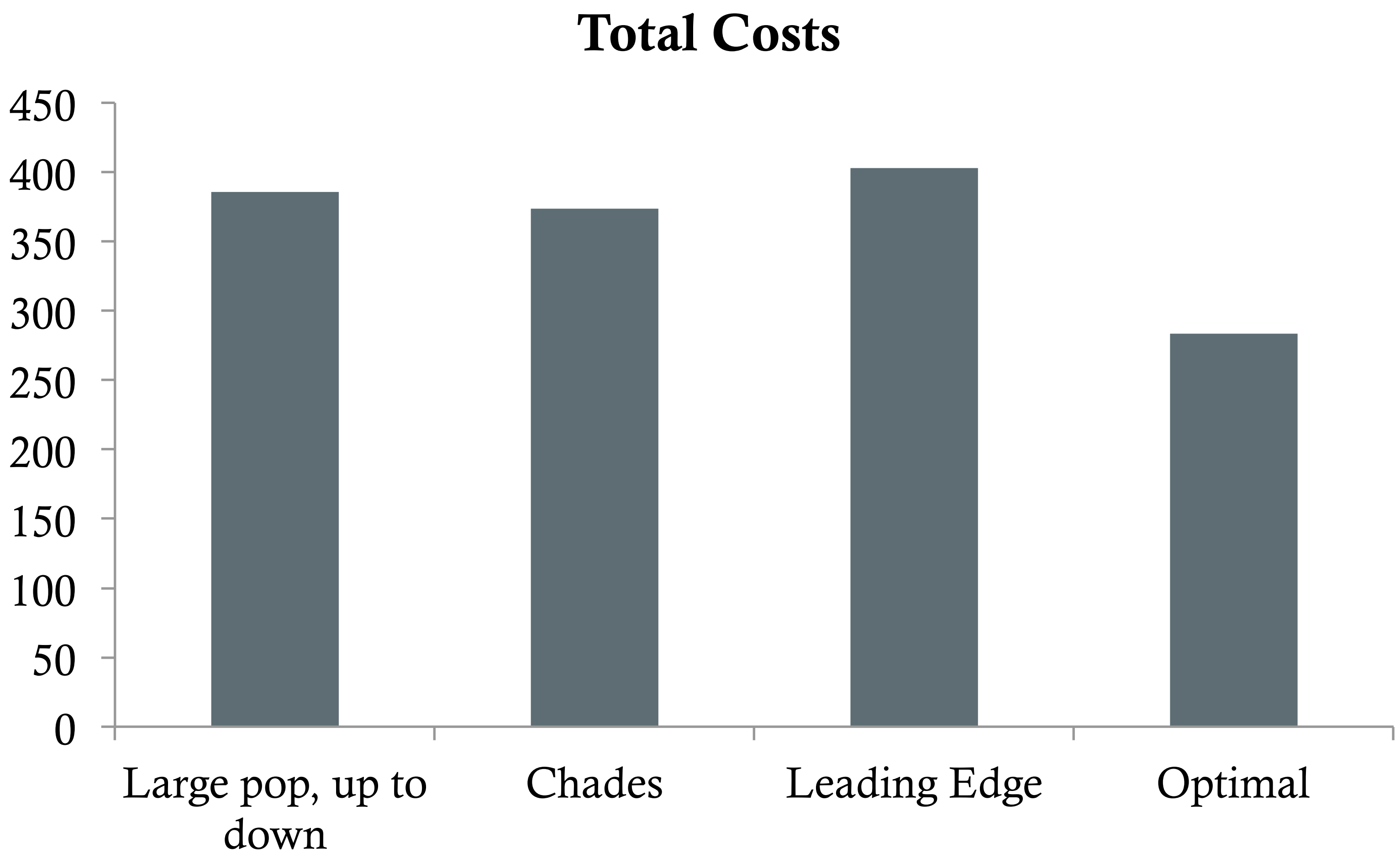
Optimal



Triage up to down



Results: Costs



Results

- Optimized spatial policy can outperform aspatial rules of thumb policies
- Spatial characteristics of the system under invasion are relevant to optimal management
    - strength of downstream vs upstream dispersal
    - presence of long distance dispersal changes policy

Future Ecology/Economics work

- Better ecological models of competition needed
- More investigation of space-time interactions
- Stochastic arrivals from outside network
- Richer objective functions
  - Separate competitiveness and colonization probabilities
- Model human dispersal of invasive plants via boating

Future Computational Work

- **Memory:** value iteration with partial transition model loaded into memory
- **Data re-use:** minimizing calls to expensive simulators when learning model
- **Compact Representations:** more compact spatial representations of states and policies
  - Larger problem sizes
  - Improved policy interpretation
  - Relational learning to distill general rules from policy
- **Bounded Approximations:** PAC-style algorithms with bounds on results
  - Estimate values of states directly through simulation

References

- Chades, I., T. G. Martin, et al. (2011). "General rules for managing and surveying networks of pests, diseases, and endangered species." Proceedings of the National Academy of Sciences of the United States of America 108(20): 8323-8328
- Muneepeerakul, R., J. S. Weitz, et al. (2007). "A neutral metapopulation model of biodiversity in river networks." Journal of Theoretical Biology 245(2): 351-363
- Strehl, A & Littman, M. (2008). An analysis of model-based Interval Estimation for Markov Decision Processes. Journal of Computer and System Sciences, 74(8), 1309-1331.
- Weissman, T., Ordentlich, E., Seroussi, G., Verdu, S., & Weinberger, M. J. (2003). Inequalities for the L1 Deviation of the Empirical Distribution. Hewlett-Packard Labs.

This work funded as part of the US National Science Foundation project in Computational Sustainability.

Thank You

Questions?