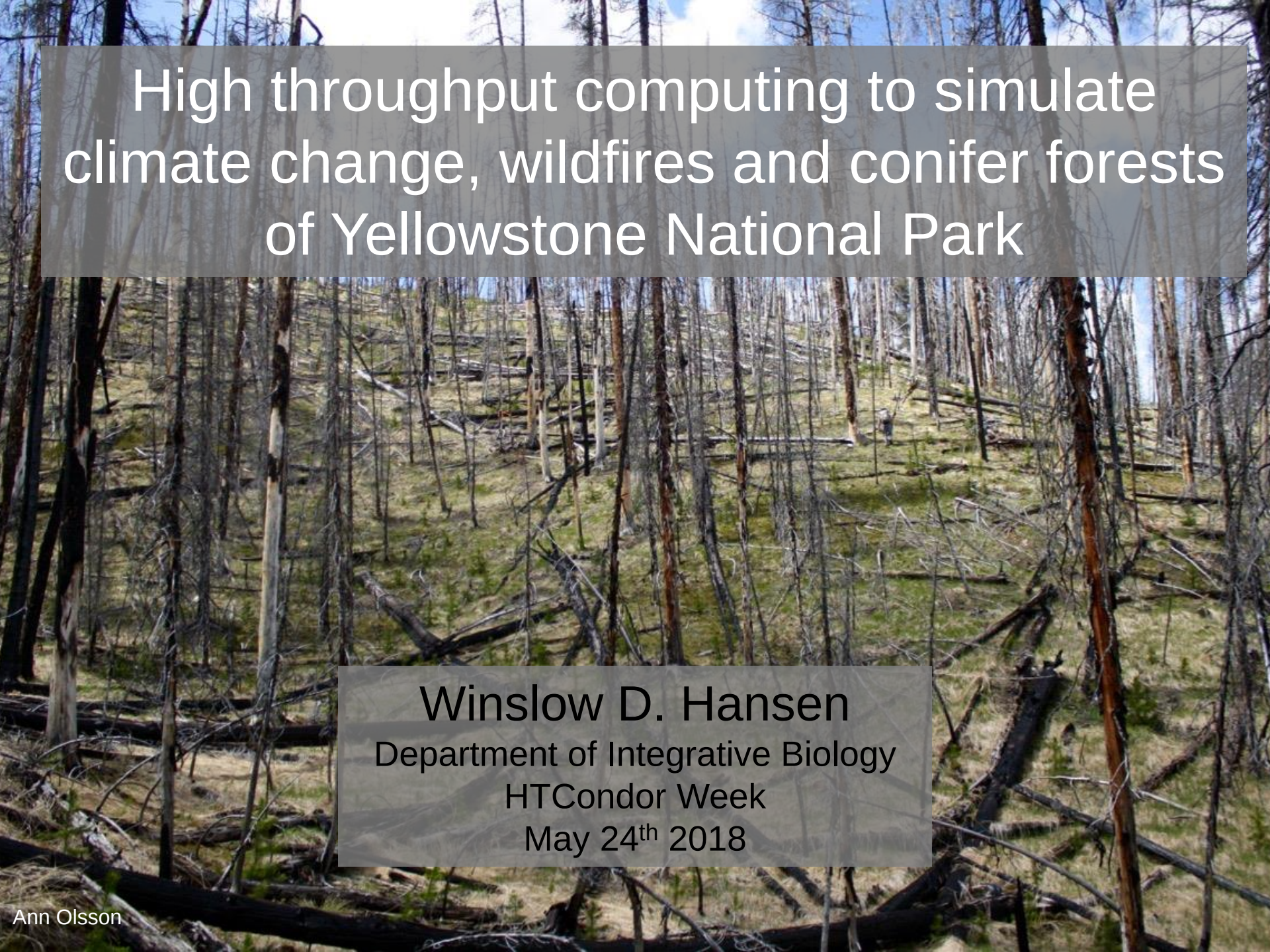
High throughput computing to simulate climate change, wildfires and conifer forests of Yellowstone National Park

Winslow D. Hansen
Department of Integrative Biology
HTCondor Week
May 24th 2018


Ann Olsson

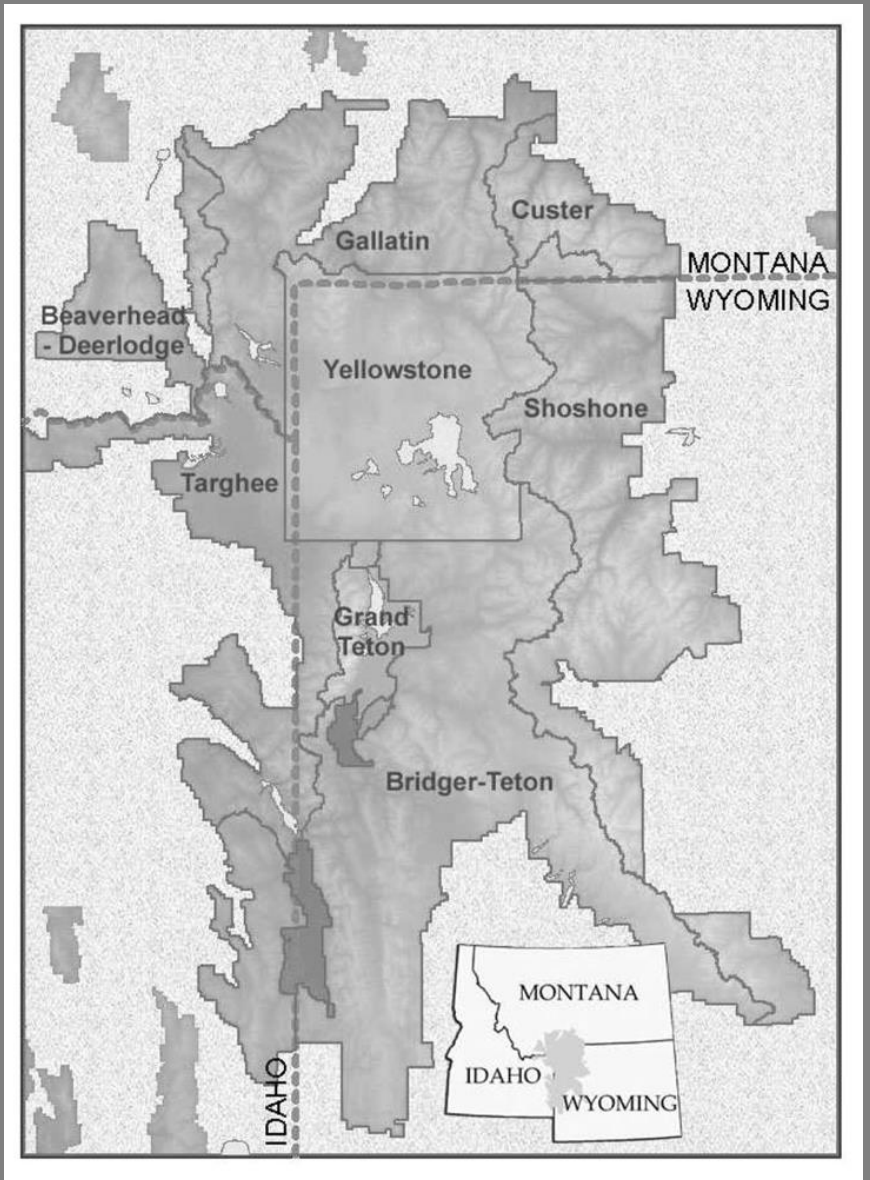
Acknowledgements

- Monica Turner
- Rupert Seidl and Werner Rammer
- Kristin Braziunas and Zak Ratajcazk
- NSF, JFSP, Alaska CSC, NPS, Vilas Research Trust

Ann Olsson

Yellowstone National Park

- Large intact wildland area
- Oldest national park
- Point farthest from a road in the contiguous US

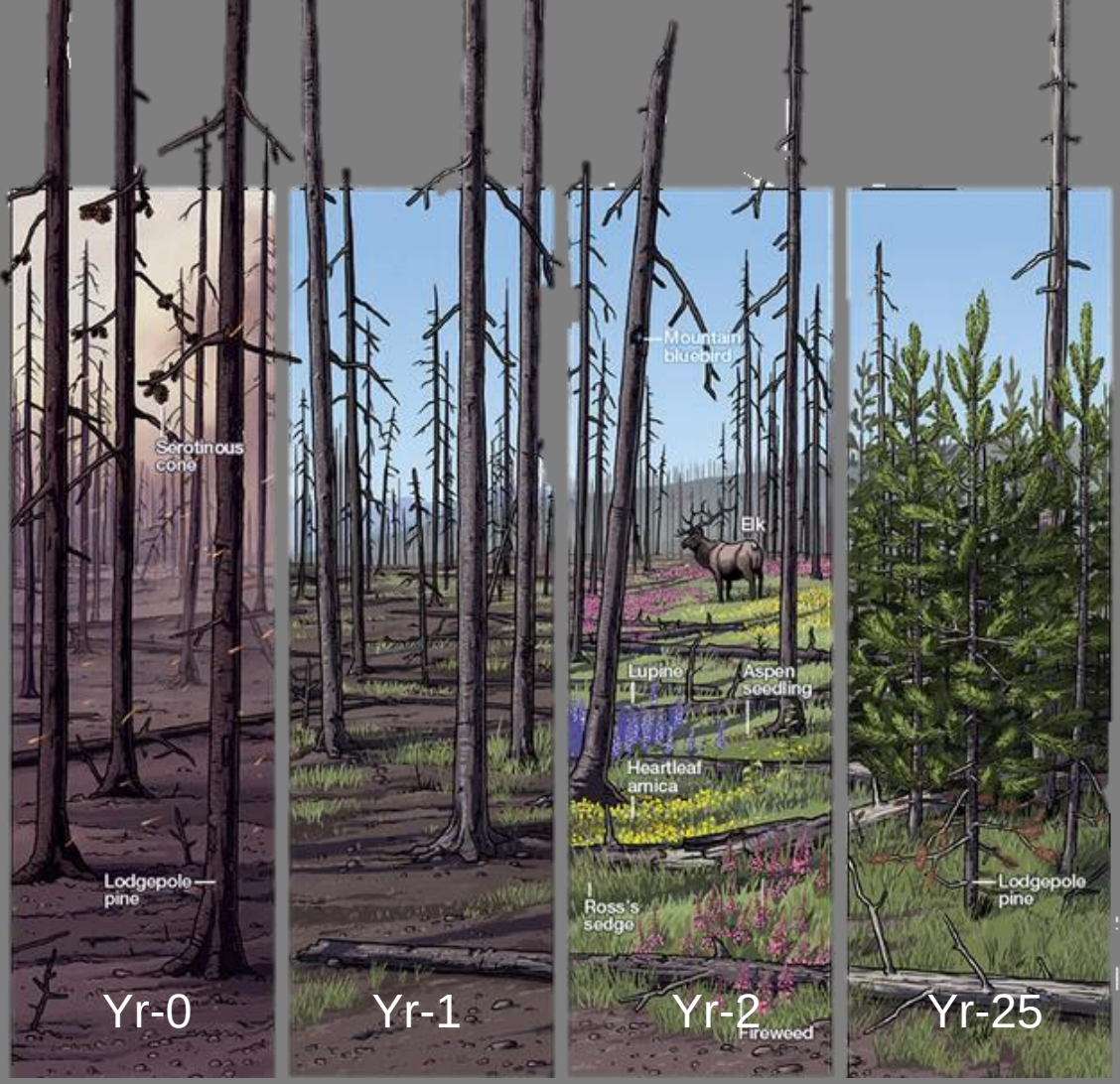


Landenburger et al. 2008 Sensing

Ann Olsson


Ann Olsson

Ann Olsson



Tree regeneration historically robust

National Geographic May 2016

Climate change causing increased fire and droughts

US News

Fires and drought may exceed forest resilience

National Geographic May 2016

Where and why might changing climate and fire may cause forests of Yellowstone to change profoundly?

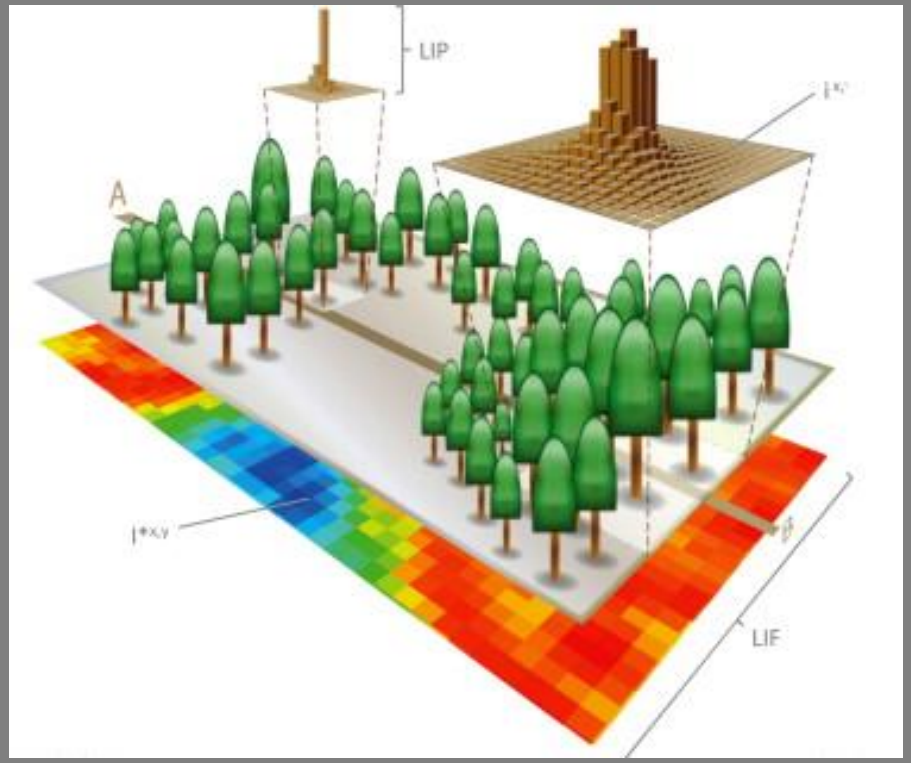
Ann Olsson

Research Challenges

- How do we study forests 100 years in the future?

Process-based models

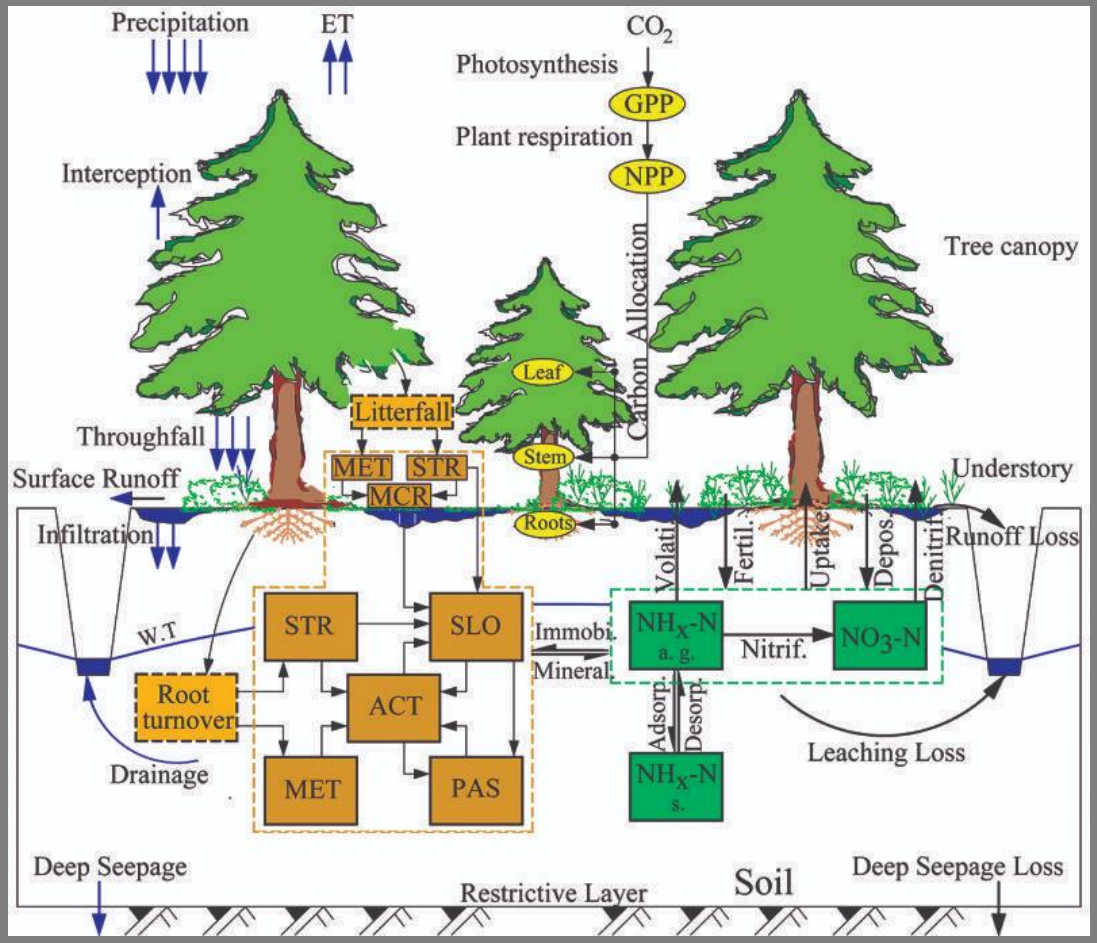
- System is represented as a set of mechanistic processes that describe cause and effect relationships between system variables



Seidl et al. 2014

Research Challenges

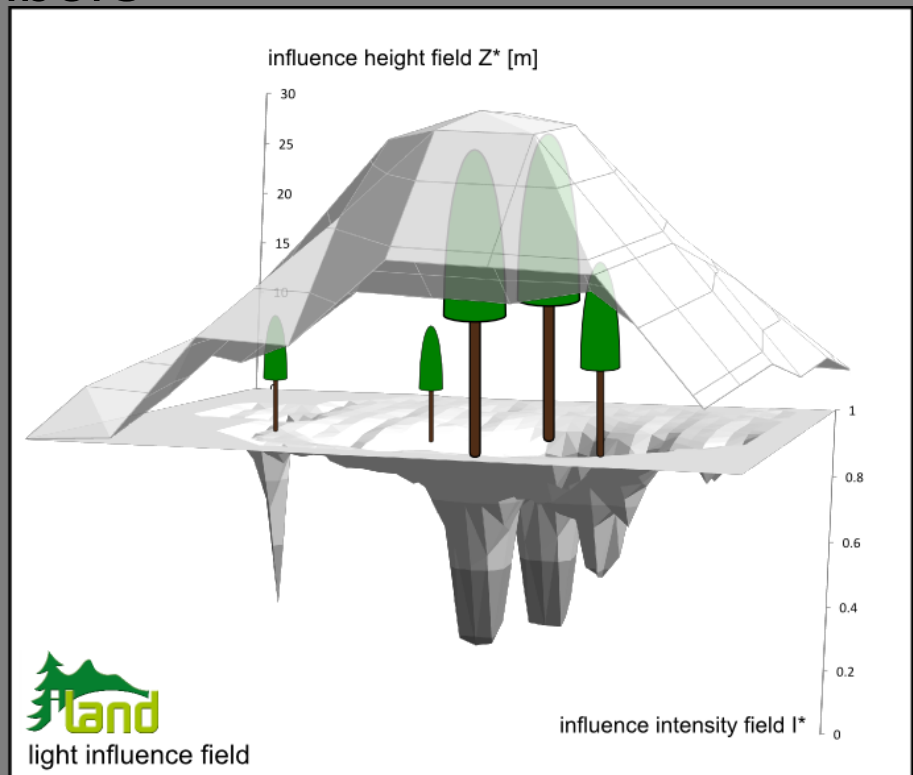
- Forests are very complex!



Tian et al. 2011

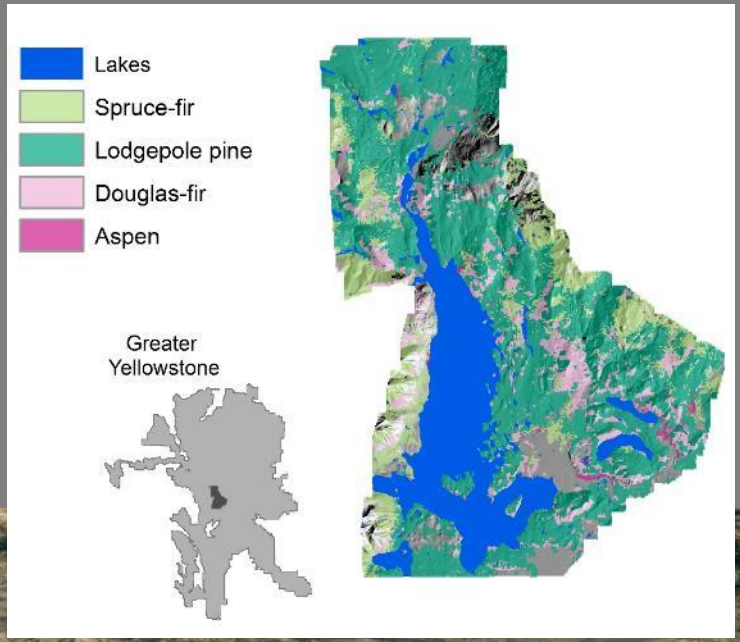
iLand

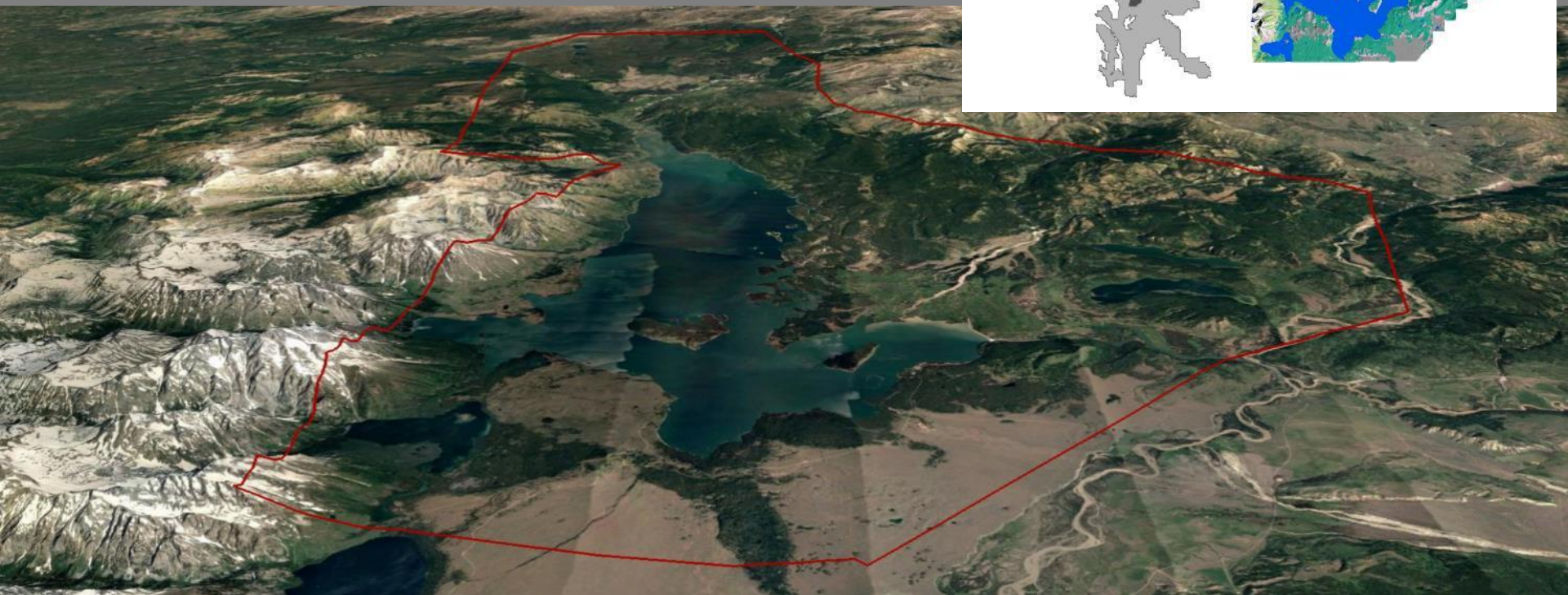
- Individual forest Landscape and Disturbance
  - Tree reproduction, growth, and mortality
    - Light availability, climate, soils, competition
  - Trees interact with their neighbors
  - Broad-scale patterns and process emerge dynamically



Spatially explicit landscapes

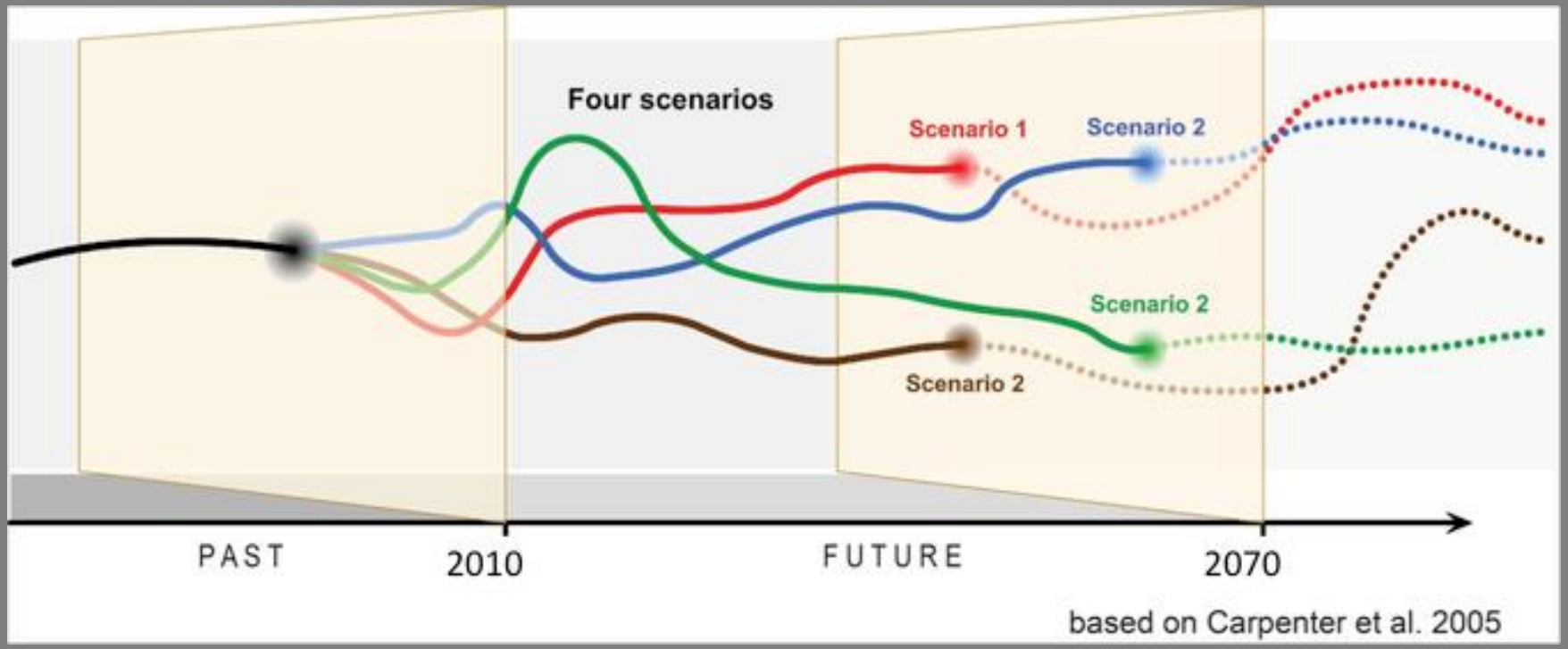
- Millions of trees
- Simulate wildfires
- ~ 3 hr runs on office computer





Scenarios and stochasticity

- Future is uncertain and so we often ask what if questions
- Run multiple replicates of each scenario



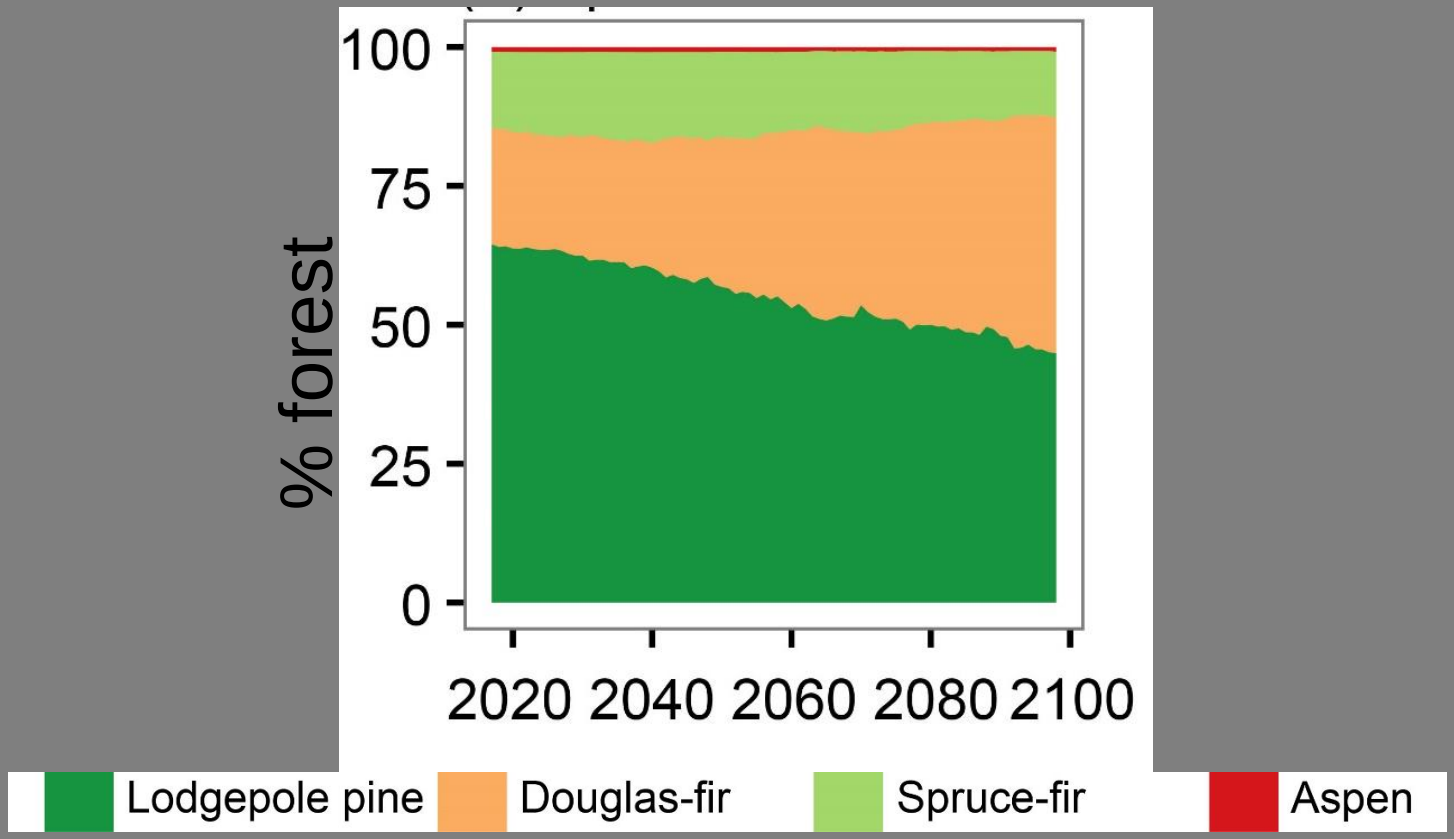
The need for high throughput

- 11 different scenarios
- 20 replicates of each scenario
- totaling 280 runs
- This would have taken me 35 days of computation working 24hrs per day
- It took 6 hrs with HTCondor

Post-processing

- Simulating individual trees every day for 100 yrs generates a lot of data
- 10 gigabytes/ run minimum
- Also used high-throughput to aggregate data to landscape scales in R

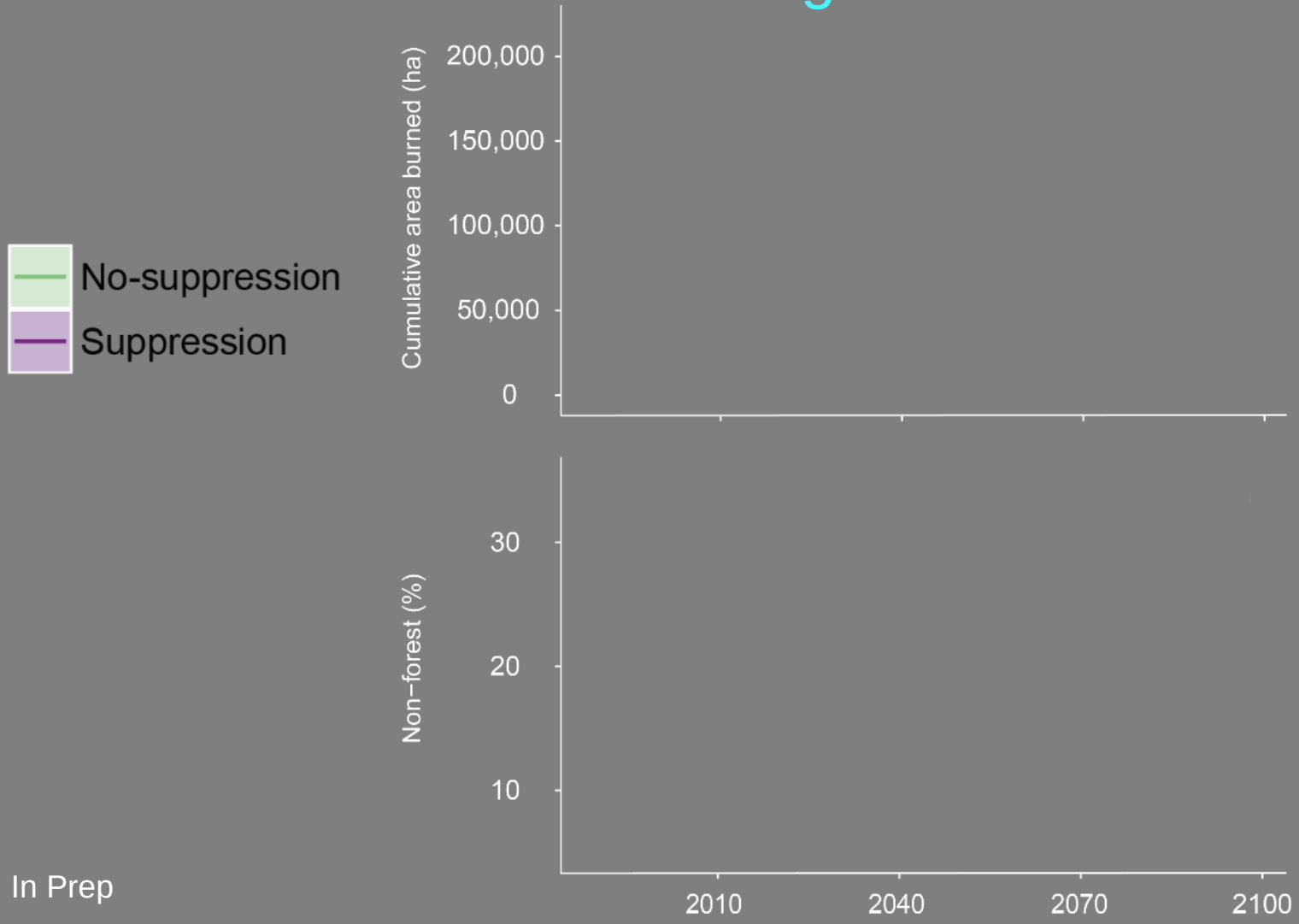
21st century fire and forests



Hansen et al. In Prep

21st century fire and forests

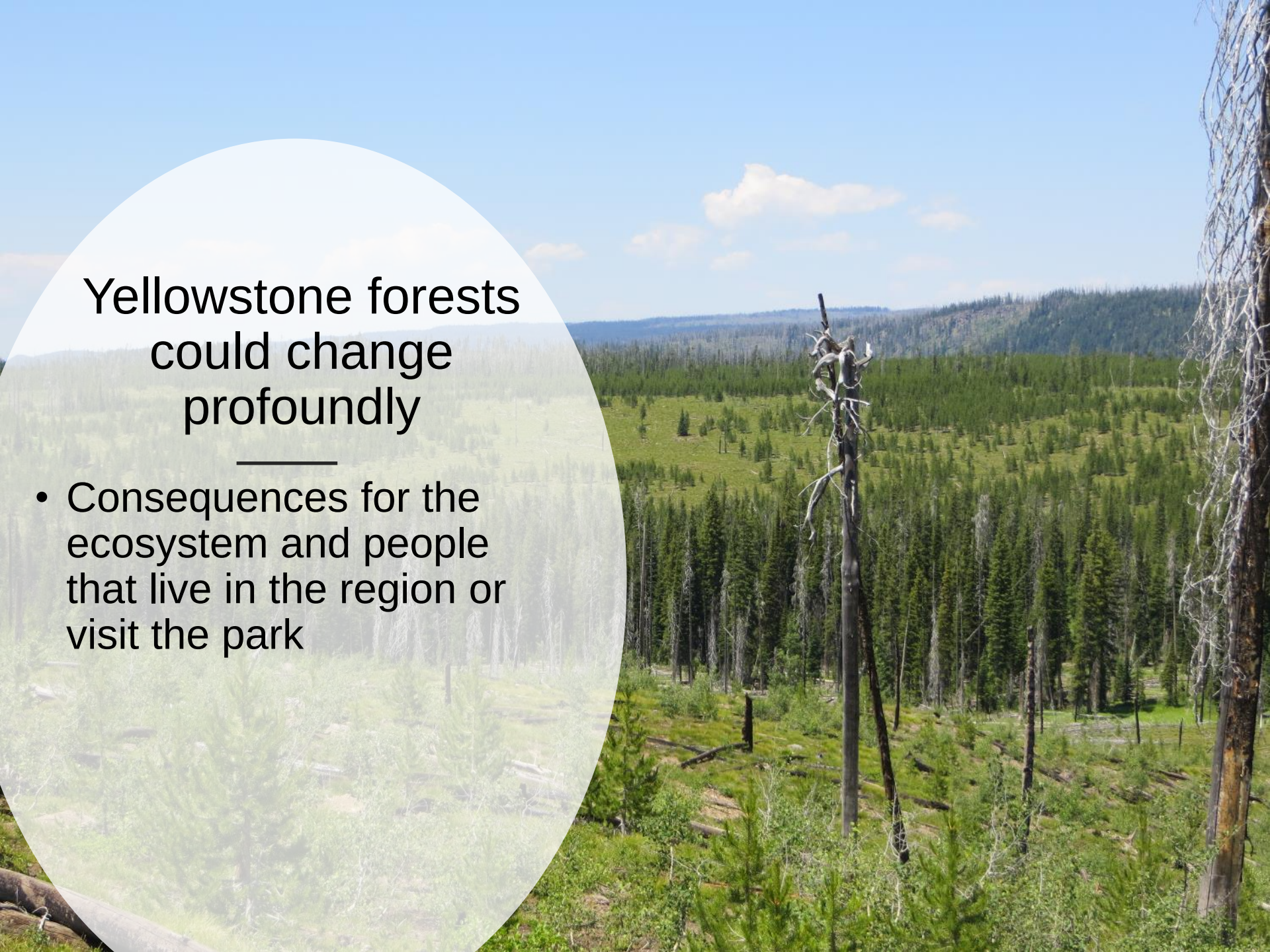
Identified thresholds of burning



Hansen et al. In Prep

Yellowstone forests could change profoundly

- Consequences for the ecosystem and people that live in the region or visit the park



- Ecology is going through a computational renaissance
- Gaining unprecedented insights

- A lot to how the computing sausage is made
- Ecologists are not well trained for this type of work

Take aways

Computing newbies do not know what they are doing and that is okay

- High throughput computing is intimidating for beginners
- Experts often lack a frame of reference for where beginners are starting from
- Patience is required to break barriers

Document everything to save time later

- When one task is conquered, often its on to the next
  - Insufficient repetition to master each skill
- Critical to document step by step for repeating computing tasks in months or years
- Helpful for institutional memory as new students begin in labs and old ones leave

Computing facilitators are essential

- True innovation of CHTC
- Facilitators maximize the value to university and scientific communities