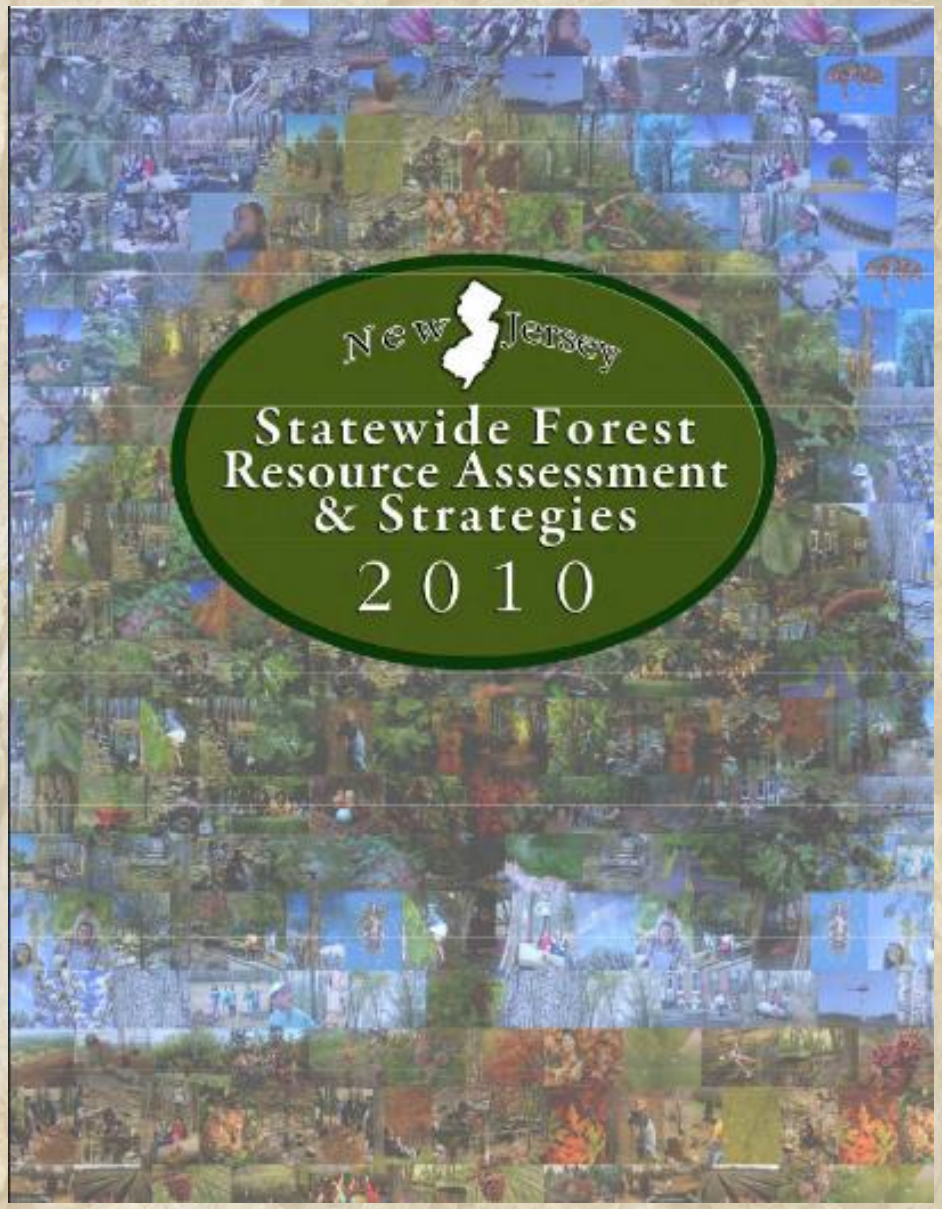
Modeling and Data Mining Concepts Applied to Forest Resource Management

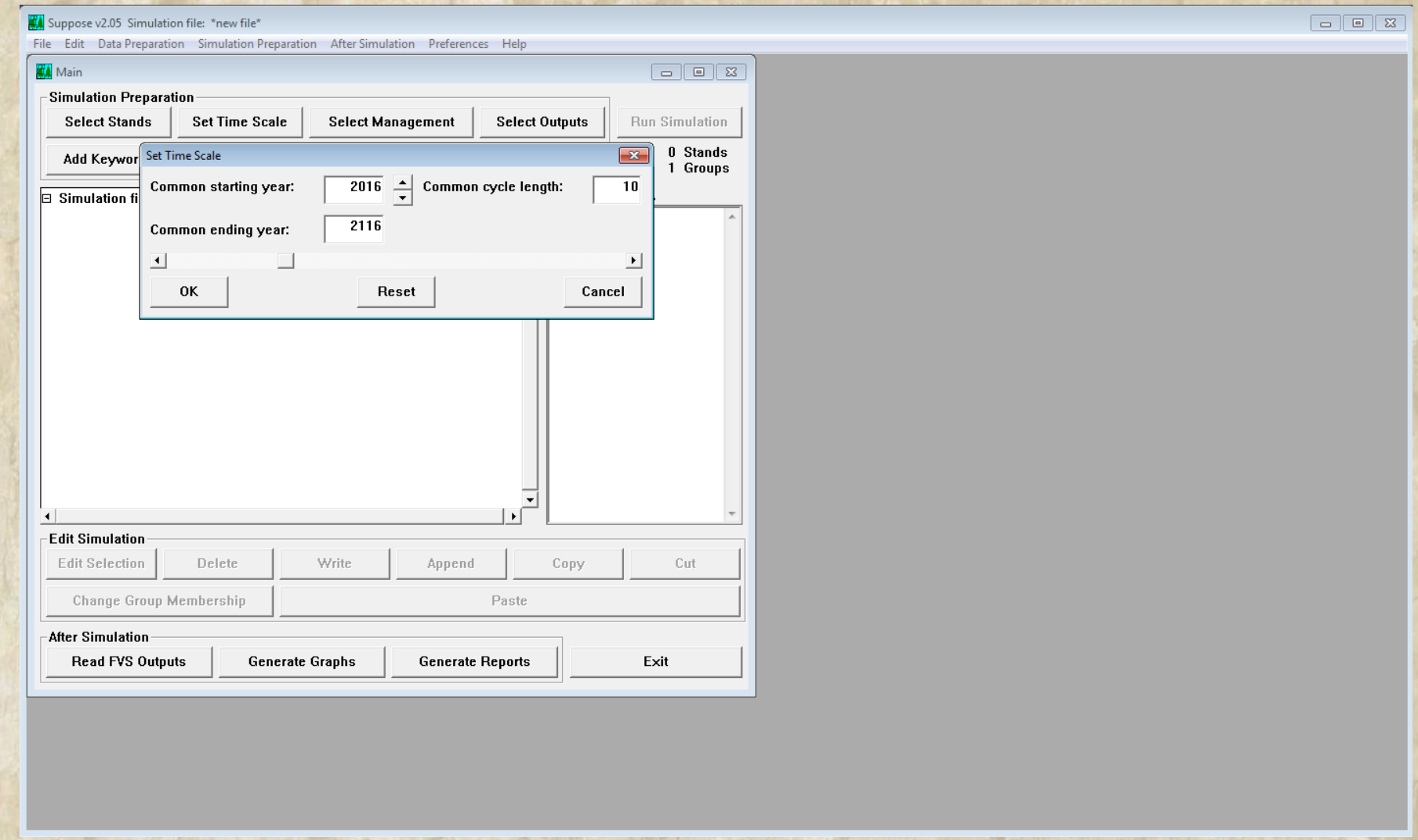
Image from: http://alvenka.deviantart.com/art/fractal-tree-34-366989431

William Zipse
NJ State Forest Service

NJ Forest Action Plan/Montreal Process



Part I – Present Capabilities



In Situ Data Collection Efforts

- State Lands Inventory
  - Snapshot in time
  - State Forest Level (Forest Type)
  - Variable Radius
- USFS FIA
  - http://www.fia.fs.fed.us/tools-data/
  - Continuous (entire state measured every 5 years)
  - 2X national grid intensity in NJ
  - Measure Change (growth, mortality, removals, etc.)

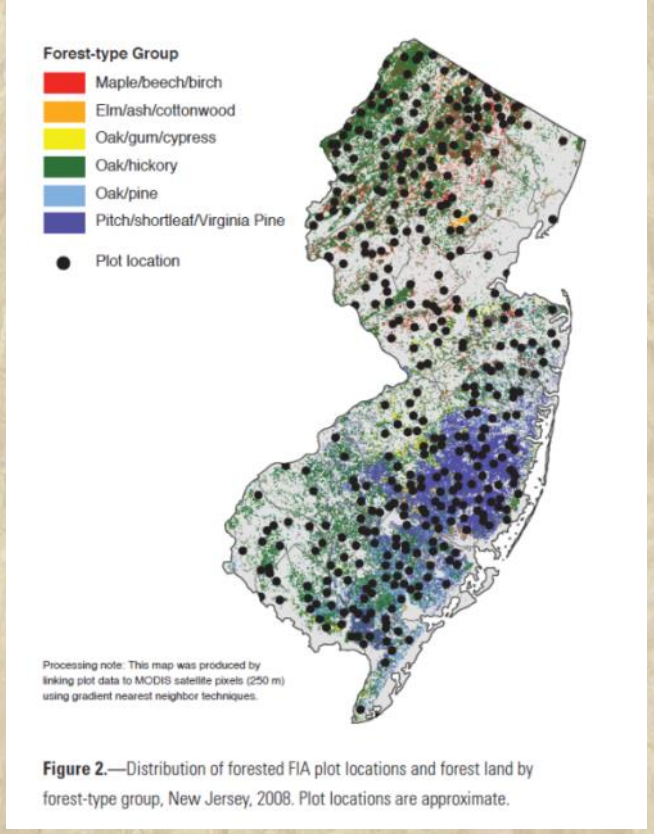


Figure 2.—Distribution of forested FIA plot locations and forest land by forest-type group, New Jersey, 2008. Plot locations are approximate.

Forest Modeling

- Ecosystem processes and management responses are complex!

- Models are mathematical representations of objects, processes, and/or interactions
  - They are abstractions of reality!
  - Often deliberately emphasize one aspect of the system at the expense of others

- Models take very complex problems and make them more manageable relative to our scope of understanding.
  - Large areas
  - Long time frames
  - Complex interactions

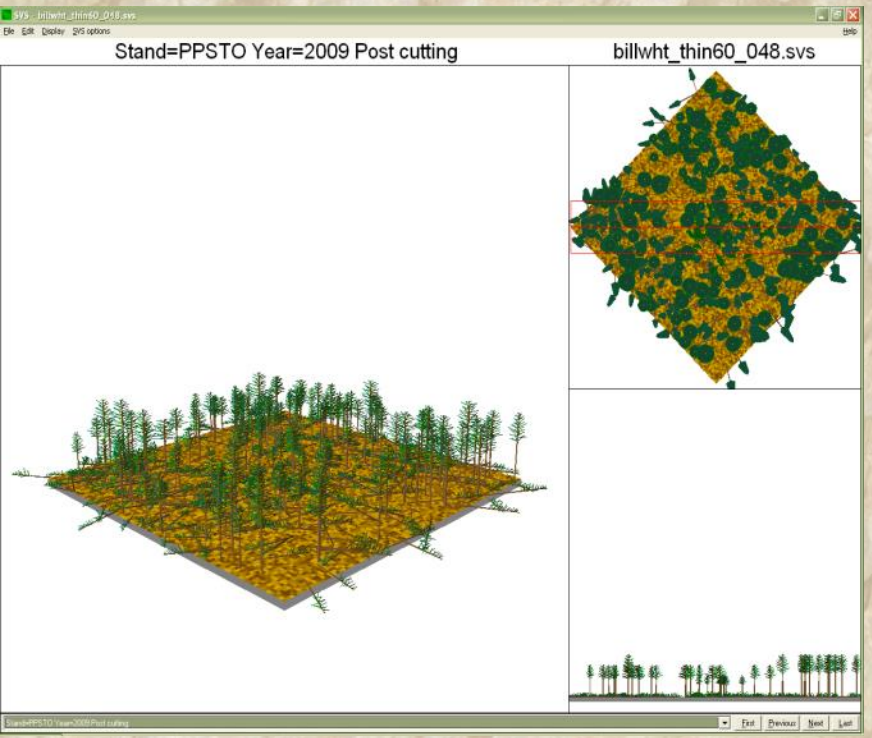
- Requires iteration, not a crystal ball!

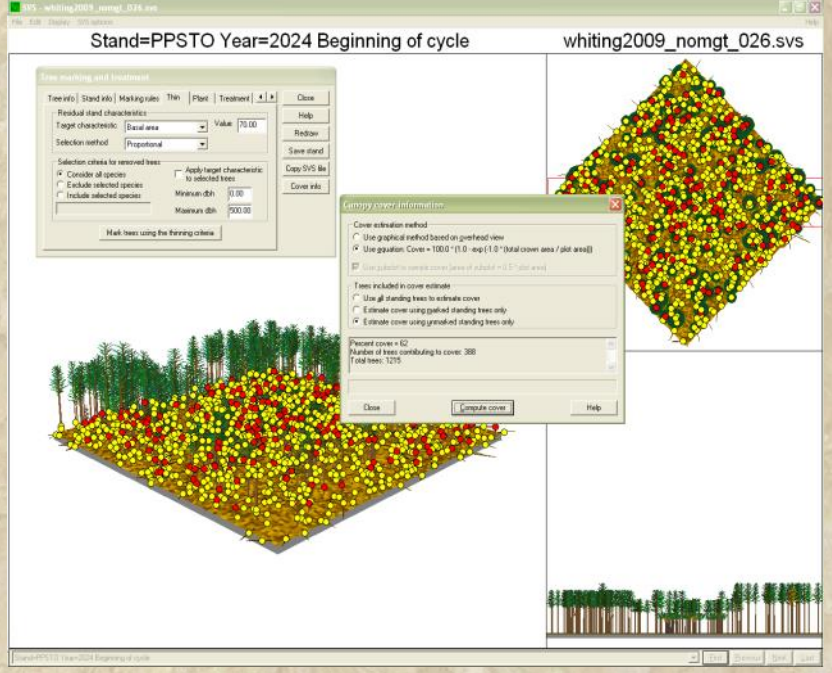
Simulation Modeling

Simulation Models

- Specific application of models to arrive at an OUTCOME
- Can evaluate outcomes of actions or lack thereof at varying time intervals and scales
- "Simming" does not generally tell you directly whether the outcome of a particular set of actions is "better" or "worse"
- Advantages
  - Test alternatives prior to doing
  - Economical
  - Time compression
  - Team development
  - Safety

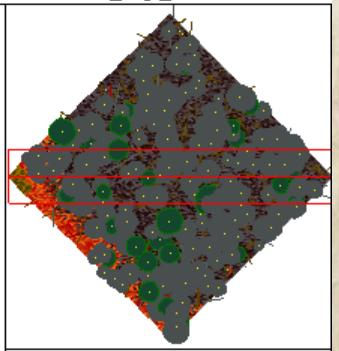
Treatment Design





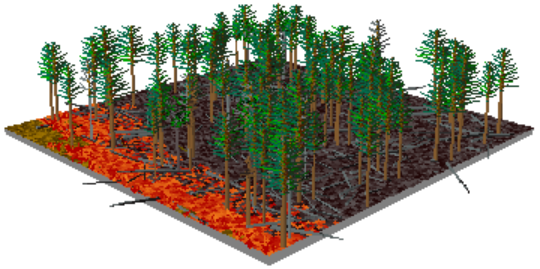
Stand=PPSTOb Year=2053 During the fire (02/03)

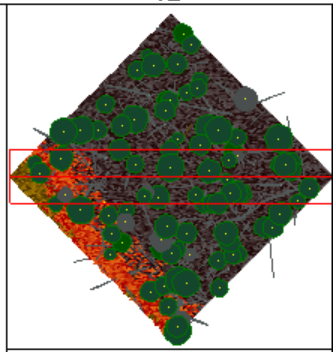
no_mgt_009.svs

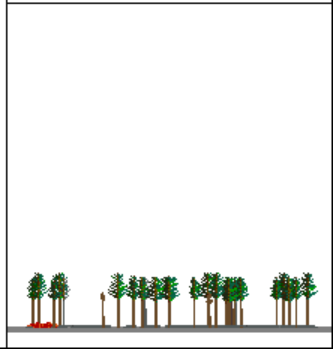


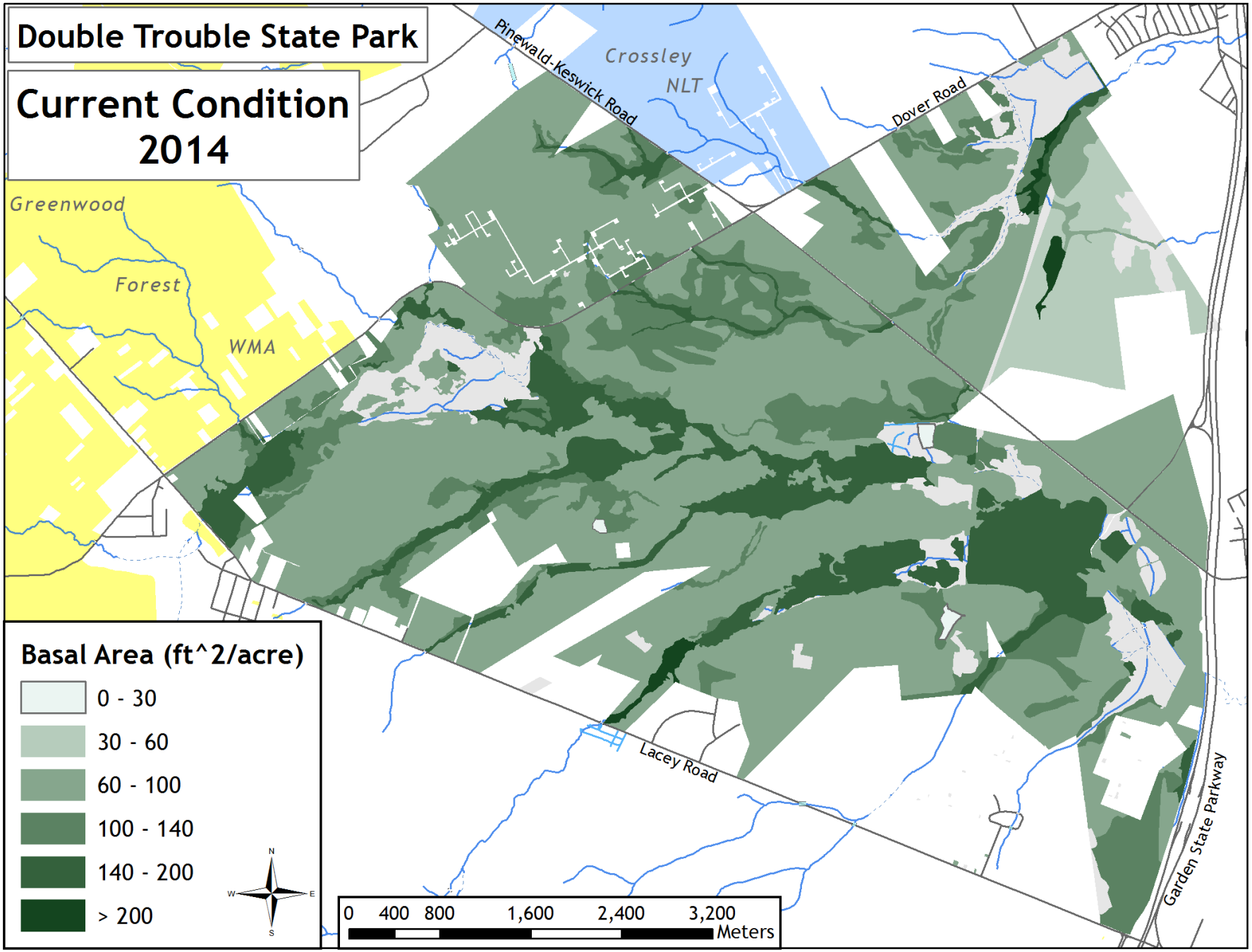
Stand=PPSTOb Year=2053 During the fire (02/03)

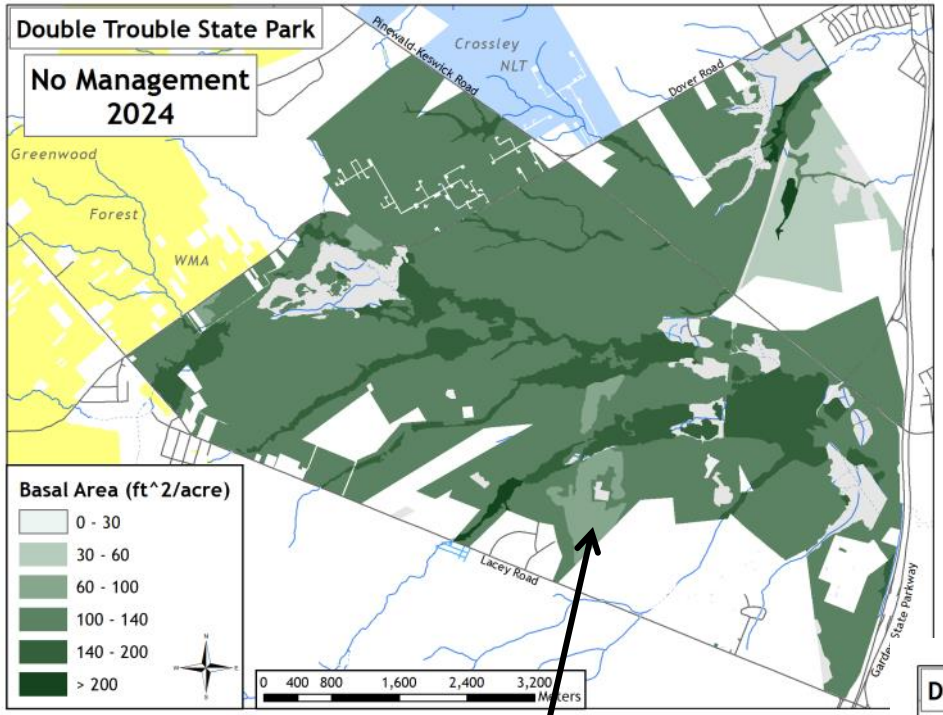
RxBOnly_027.svs





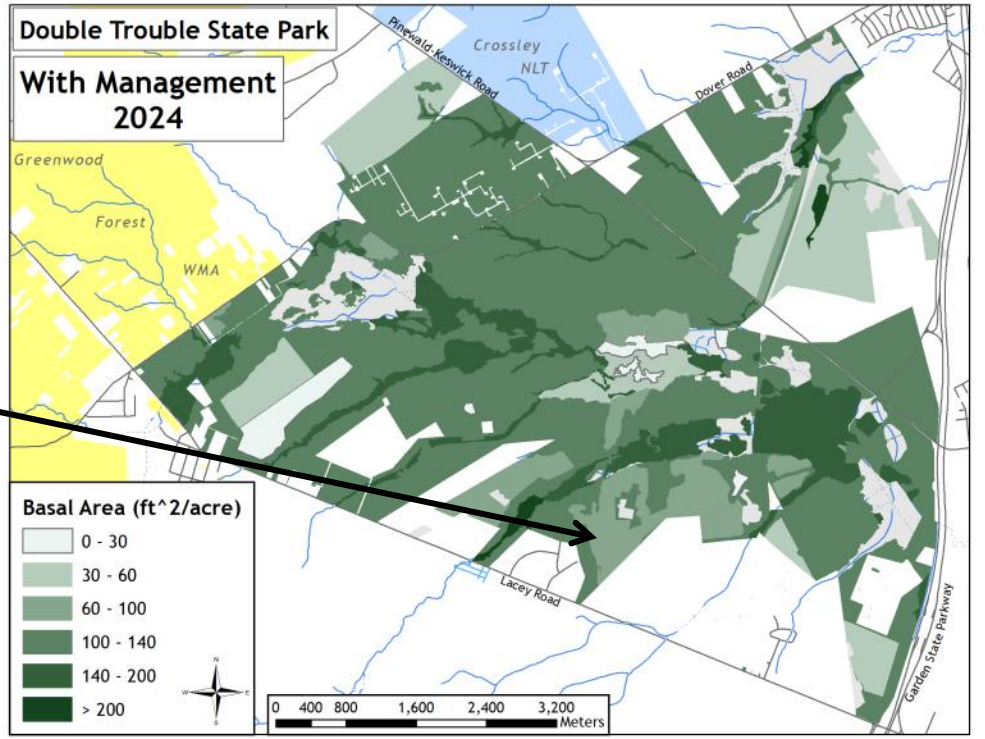


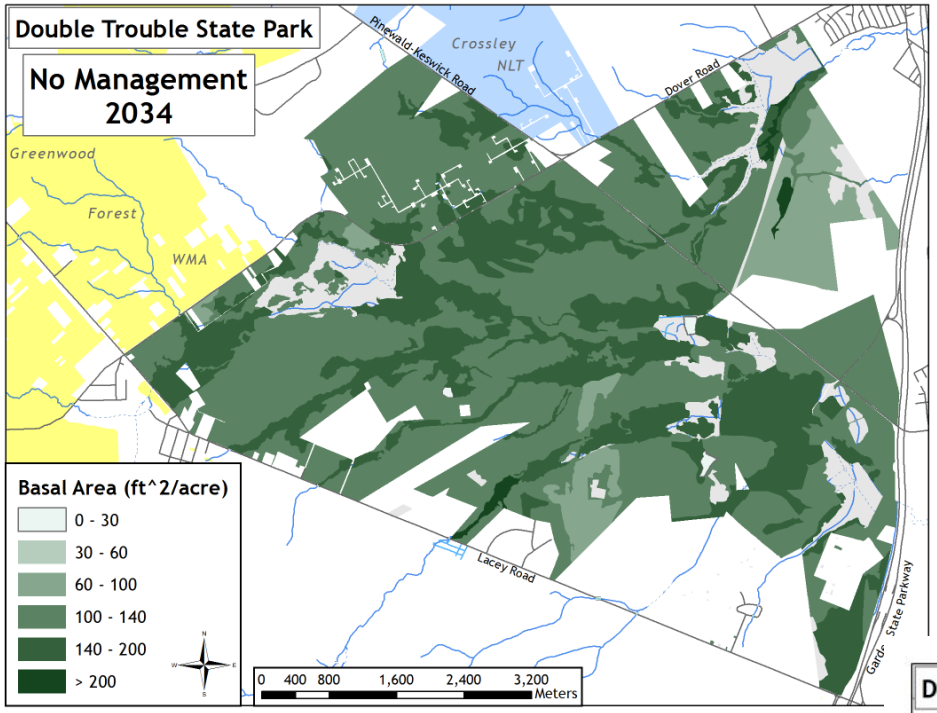

Double Trouble State Park
Current Condition
2014
Pinewald-Keswick Road
Crossley
NLT
Dover Road
Greenwood
Forest
WMA
Lacey Road
Garden State Parkway
Basal Area (ft^2/acre)
0 - 30
30 - 60
60 - 100
100 - 140
140 - 200
> 200
N
W
E
S
0 400 800 1,600 2,400 3,200
Meters

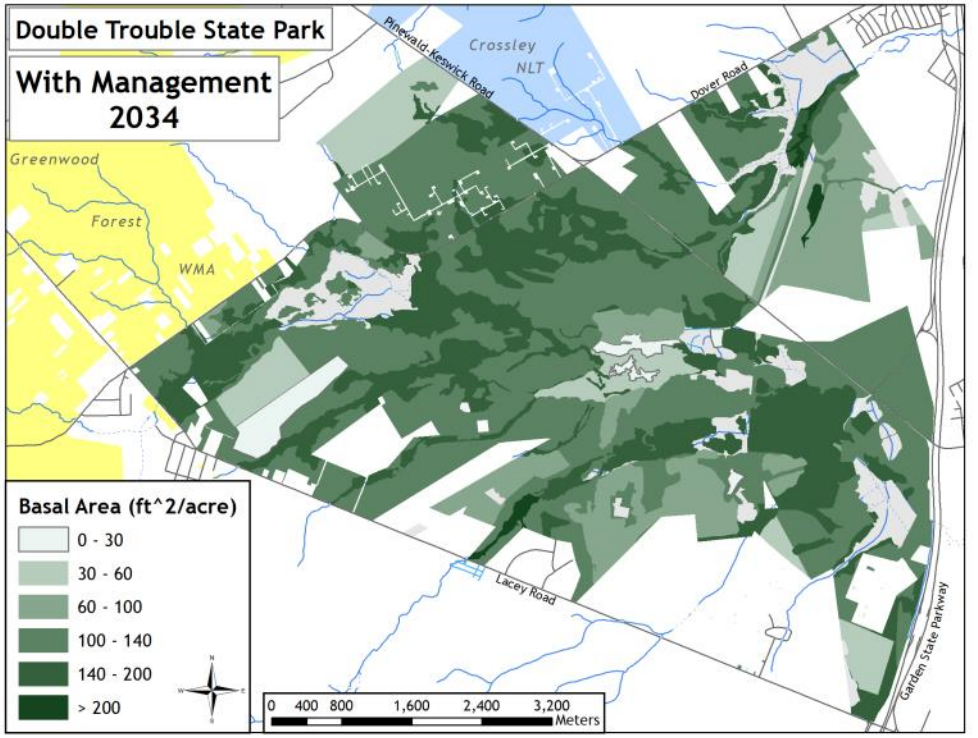


At Plan Completion

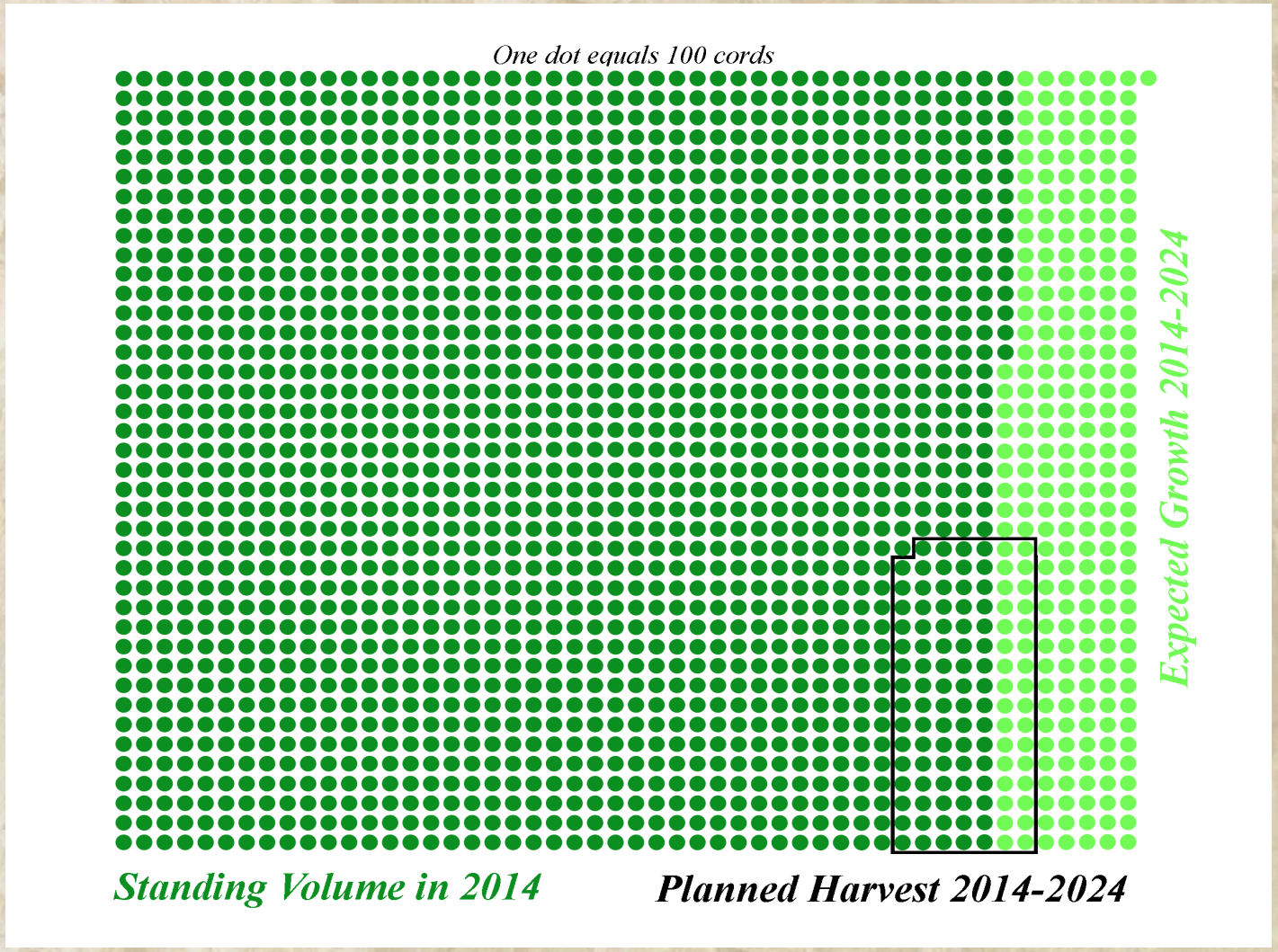
Check out wildfire impacts
From 1990's wildfire!




Double Trouble State Park
No Management
2034
Crossley
NLT
Pinewald Keswick Road
Dover Road
Greenwood
Forest
WMA
Lacey Road
Garden State Parkway
Basal Area (ft^2/acre)
0 - 30
30 - 60
60 - 100
100 - 140
140 - 200
> 200
0 400 800 1,600 2,400 3,200
Meters


Double Trouble State Park
With Management
2034
Crossley
NLT
Pinewald Keswick Road
Dover Road
Greenwood
Forest
WMA
Lacey Road
Garden State Parkway
Basal Area (ft^2/acre)
0 - 30
30 - 60
60 - 100
100 - 140
140 - 200
> 200
0 400 800 1,600 2,400 3,200
Meters

Sustainability



Part II – Future Possibilities

http://www.nanoday.com/single/224/top-10-future-technologies-that-will-change-the-world

Data Mining

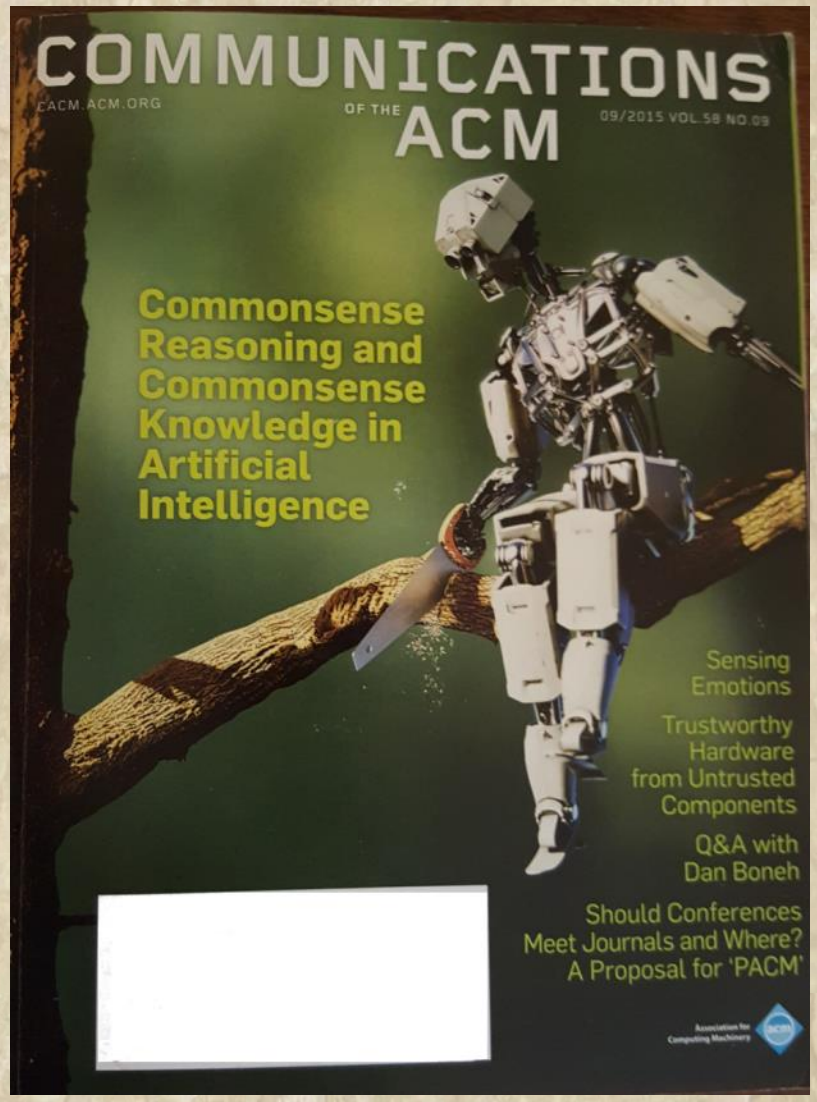
- Knowledge discovery from large data repositories
  - Extraction of interesting information or patterns from large collections of data
- Knowledge Discovery in Databases (KDD)
- Involves multiple disciplines
  - Pattern Recognition
  - Machine Learning
  - Databases
  - Statistics
  - Information Visualization

Association Mining

- Information extraction looking at association rules
  - Involves looking at support and confidence
  - i.e. Aprioi Algorithm
- Legend of Beer and Diapers

Machine Learning

- Models learned from data
- Supervised classification
- Unsupervised classification



J48 Decision Tree Example

J48 Decision Tree 10-Fold Cross Validation Results:

=== Stratified cross-validation ===

=== Summary ===

Correctly Classified Instances	2455	92.3627 %
Incorrectly Classified Instances	203	7.6373 %
Kappa statistic	0.6309	
Mean absolute error	0.116	
Root mean squared error	0.2466	
Relative absolute error	49.9371 %	
Root relative squared error	72.397 %	
Total Number of Instances	2658	

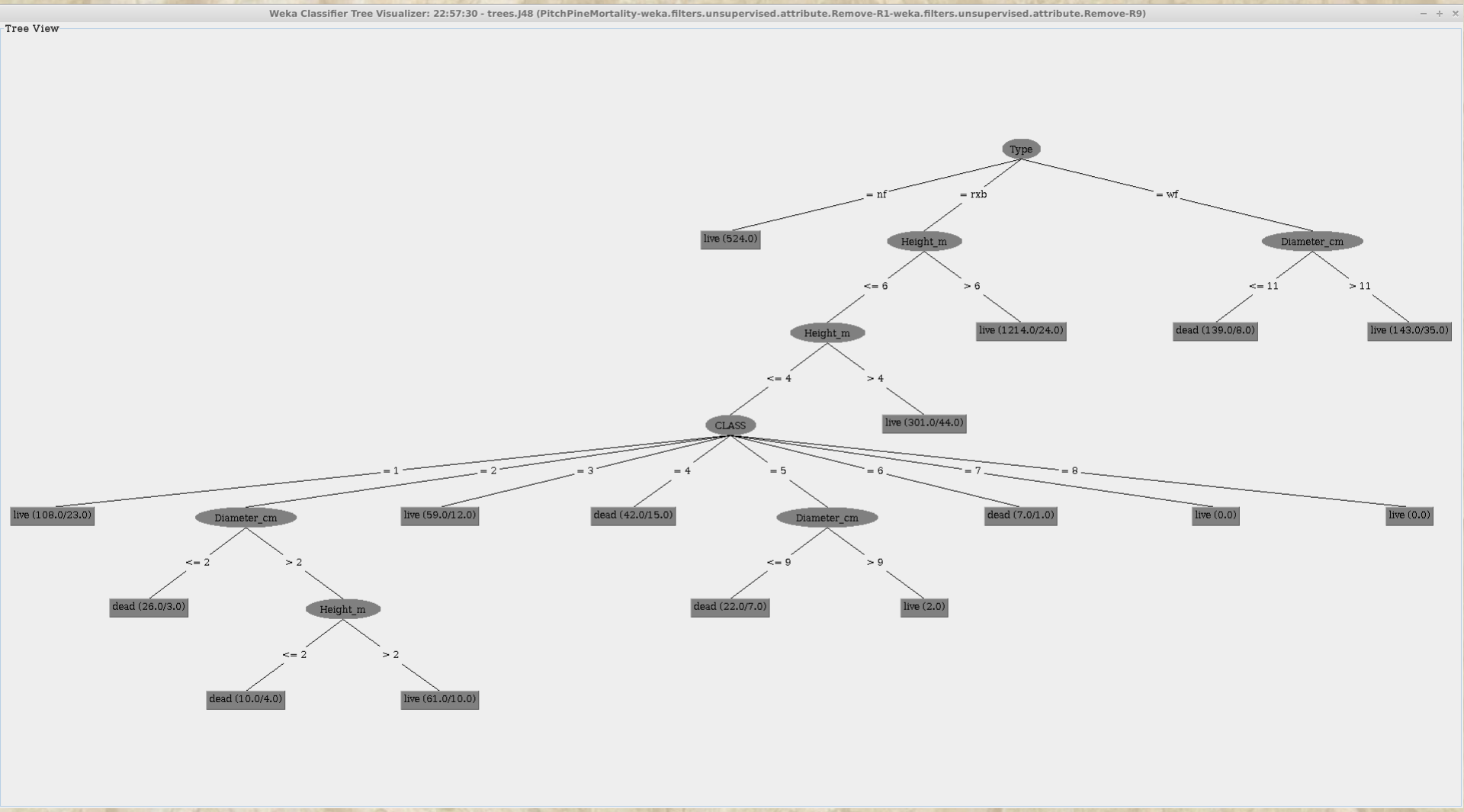
=== Detailed Accuracy By Class ===

	TP Rate	FP Rate	Precision	Recall	F-Measure	ROC Area	Class
	0.976	0.413	0.939	0.976	0.957	0.912	live
	0.587	0.024	0.789	0.587	0.673	0.912	dead
Wtd. Avg.	0.924	0.361	0.918	0.924	0.919	0.912	

=== Confusion Matrix ===

a	b	<-- classified as
2246	56	a = live
147	209	b = dead

J48 Decision Tree Example



Decision Support – Linear/Goal Programming

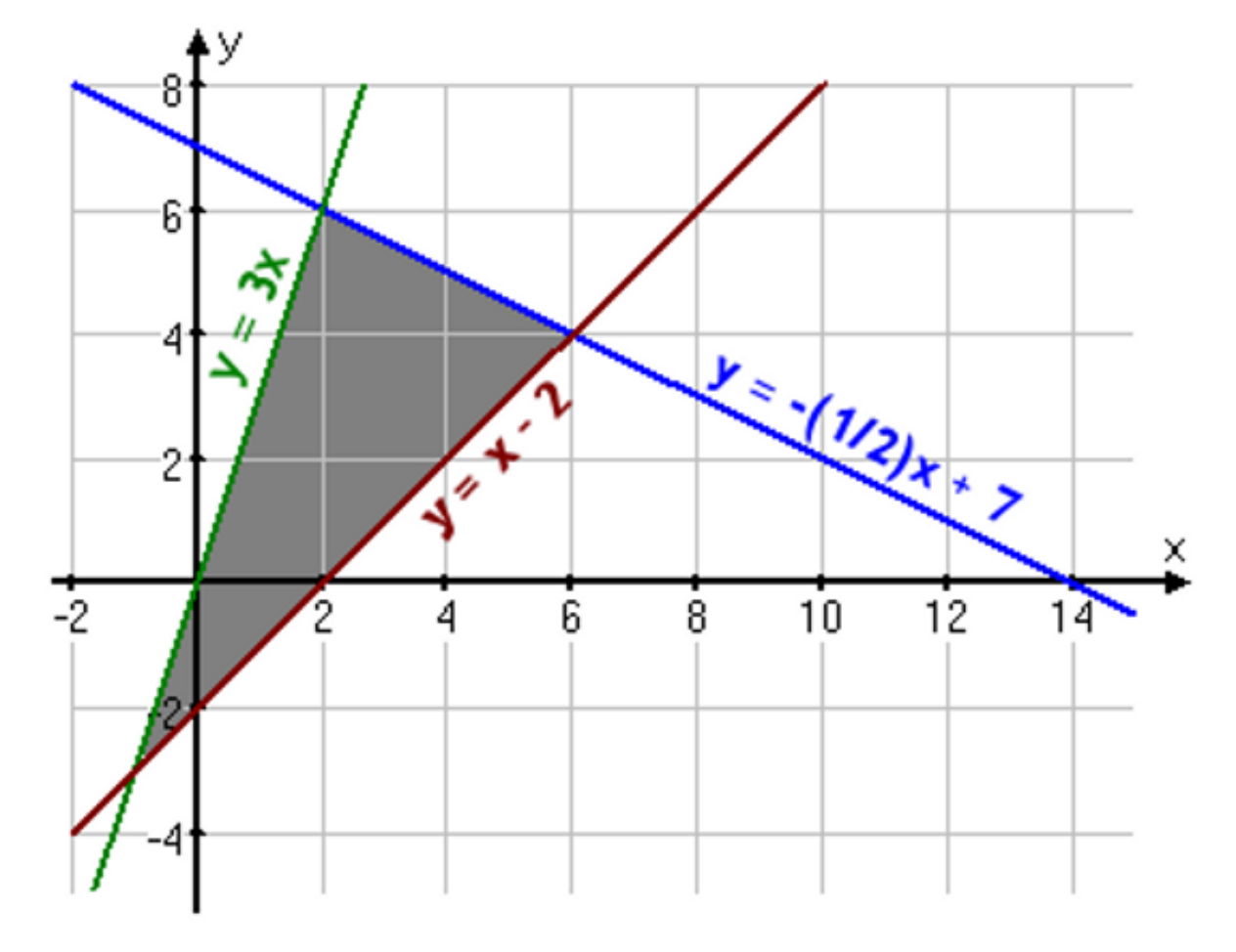
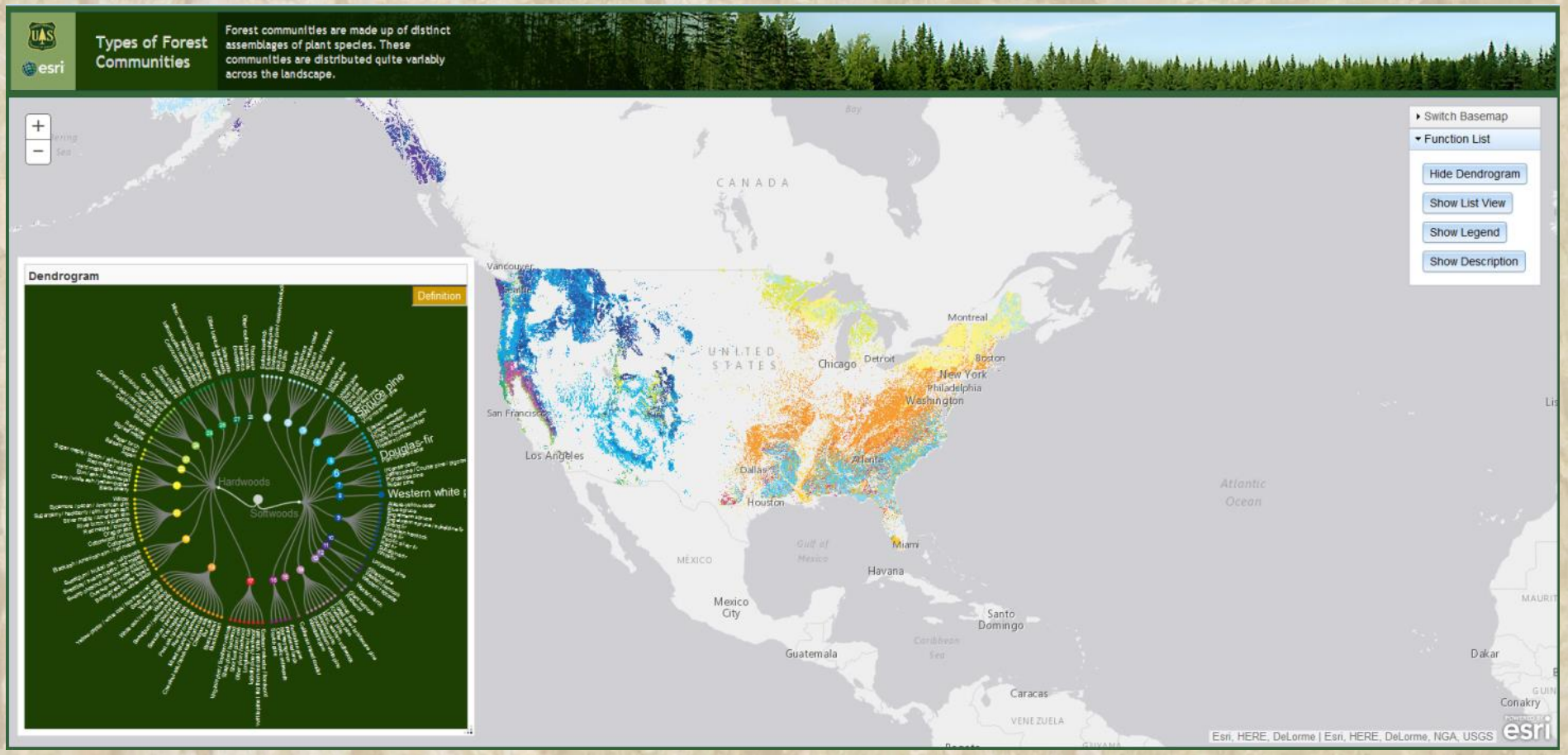
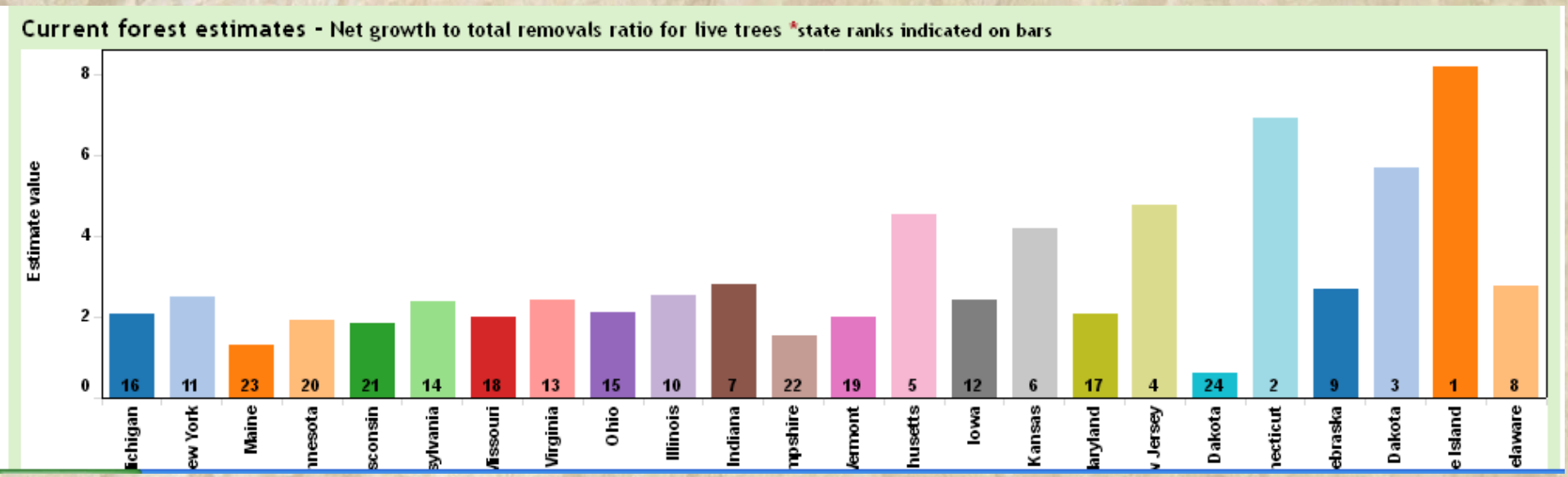


Image from Purplemath “Linear Programming: Introduction”
http://www.purplemath.com/modules/linprog.htm

Data Visualization



Data Dashboards



Other Data Visualization Ideas

http://thewestsidestory.net/blog

Questions?

William Zipse
NJ State Forest Service
william.zipse@dep.nj.gov
609-984-0815