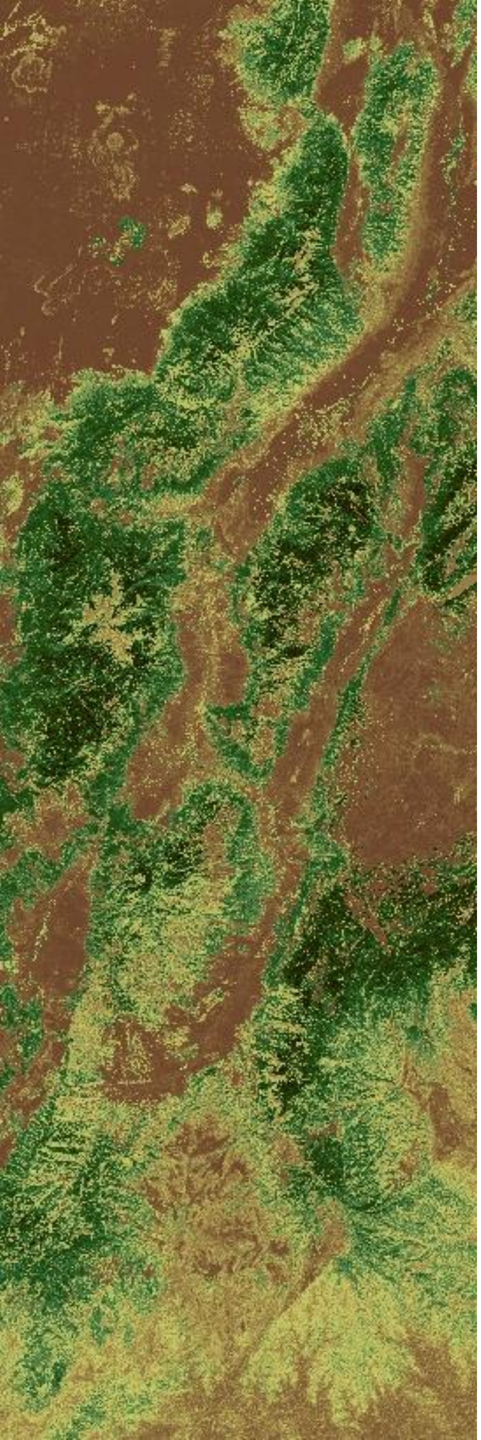


NLCD 2011 Tree Canopy Cover Pilot Project

Research & Development, Quantitative Sciences

Remote Sensing Applications Center

Rocky Mountain Research Station, FIA

Southern Research Station, FIA

Northern Research Station, FIA

Pacific Northwest Research Station, FIA

Colorado State University

State and Private Forestry, Forest Health Protection

USGS, EROS

Gretchen Moisen, John Coulston, Liz Freeman, Tracey Frescino, Chris Toney, Paul Patterson, John Tipton, Thomas Jackson, Mark Finco, Warren Cohen, Ty Wilson, Ken Brewer

The challenge…

Define the NLCD 2011 tree canopy cover mapping process

……in 8 months or less

What is the NLCD?

- NLCD is the National Land Cover Database:
  - Land cover classification layer, percent tree canopy cover layer, and a percent impervious surface layer. Primarily based on **LANDSAT** (30meter pixels) imagery and ancillary data.
  - Produced by the Multi-Resolution Land Cover (MRLC) consortium.
  - Available free from http://www.mrlc.gov/
- The MRLC is a consortium of the following agencies and programs: These are the clients of the NLCD

EPA NOAA USFS

USGS LANDFIRE BLM

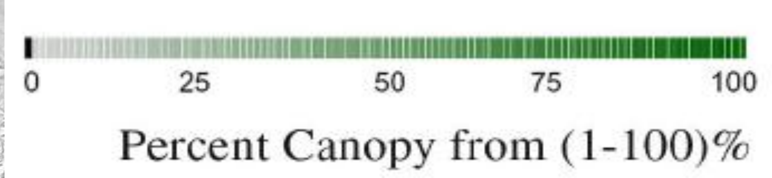
NRCS NPS NASA

USFWS OSM

Percent Tree Canopy Cover is Important !

(Example of the NLCD 2001 Percent Tree Canopy Layer)

- An integral part of both international and US forest land definitions
- Important both within forest land areas and in areas not traditionally considered forest.
- Irrespective of land use , it's an additional dimension of fragmentation
- Knowing where trees are (not just the forest) is an important first step in quantifying carbon and managing tree resources.



The Past: 2001 Percent Tree Canopy Cover

- The USGS lead the 2001 effort to map percent tree canopy cover for the United States at 30m resolution.
- This dataset serves as one of the inputs for Landfire modelling.
- The canopy cover layer is a relatively popular product. From FY07 through FY09 the canopy cover lay was downloaded from MRLC 400 times per month on average (not including ftp downloads).
- The US Forest Service examined these data for updating the 2000 assessment of urban tree cover as part of the Resource Planning Act Assessment

The Opportunity

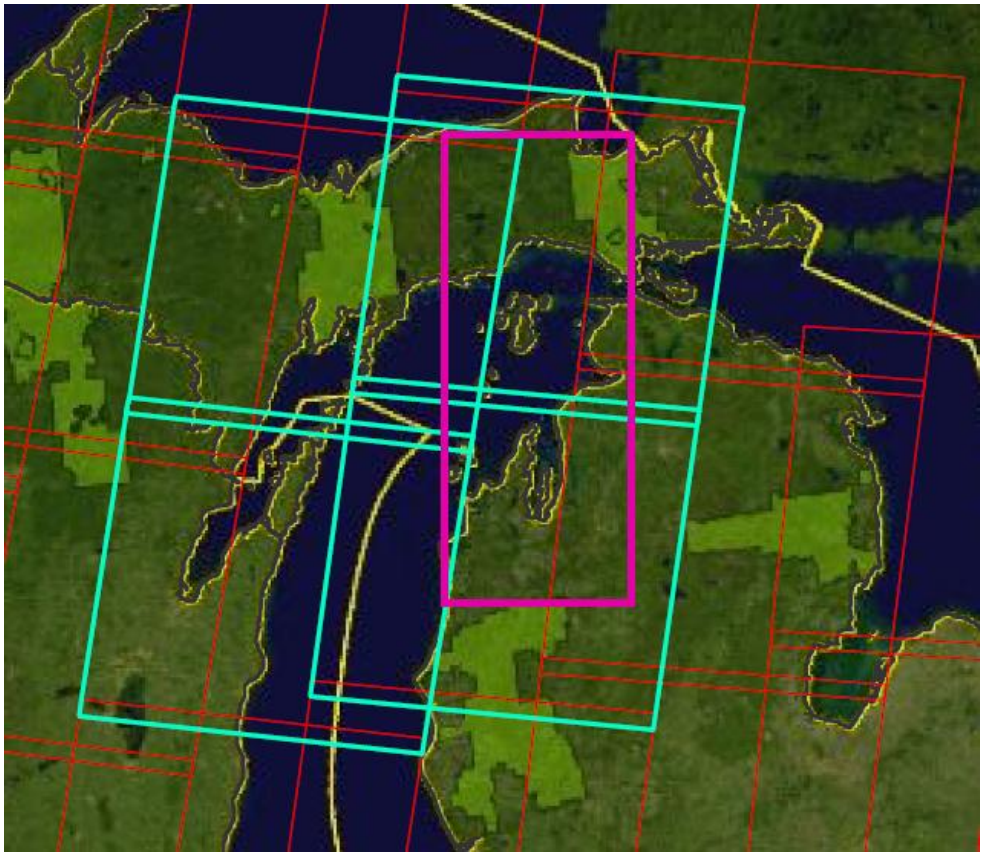
- Motivation for Forest Service and FIA Leadership
  - If it's related to trees, the Forest Service should be saying it
  - FIA is a fundamental component of Forest Service research.
  - FIA is a data rich program:
  - Consistency between map based and plot based estimates:
  - If needed, the FIA survey design is easily intensified
- How are we positioned ?
  - Implementation of tree canopy cover estimates on at all sampling locations.
  - Experience with national mapping
    - Biomass map
    - Forest type map update
    - Imputation approaches for the Atlas project

Developing the 2011 NLCD Canopy Cover Product: General Approach

- Pilot
  - Several, relatively small study areas to conduct research and inform the prototype
  - Used to identify initial prototype design
- Prototype
  - Scaling up of initial design from pilot
  - Test and verify initial design and make adjustments as necessary.
  - Finalize design components.
- Production
  - Design is fixed at this point
  - Hit enter – turn the crank – etc.
- Continued program research

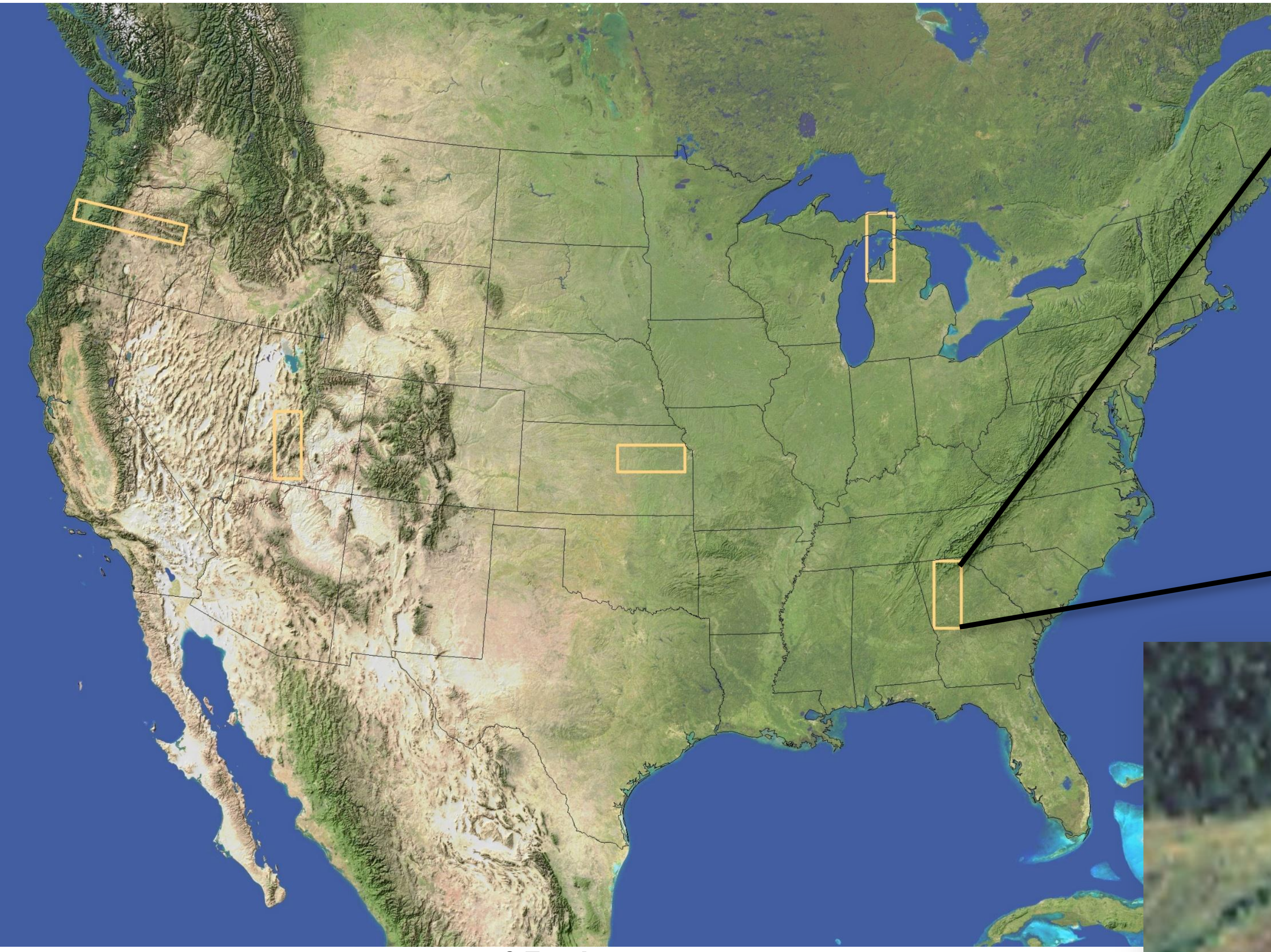

Pilot Areas

Pilot Area Characteristics



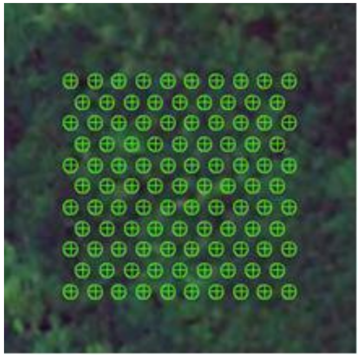
1. Approximately one Landsat scene in area
2. Cover multiple scenes
3. Cover multiple gradients
   I. urban area
   II. different vegetation types

NLCD Pilot Study Design



105 photo points to estimate
% tree canopy cover from NAIP
1m imagery

2011 General Modelling Approach



=

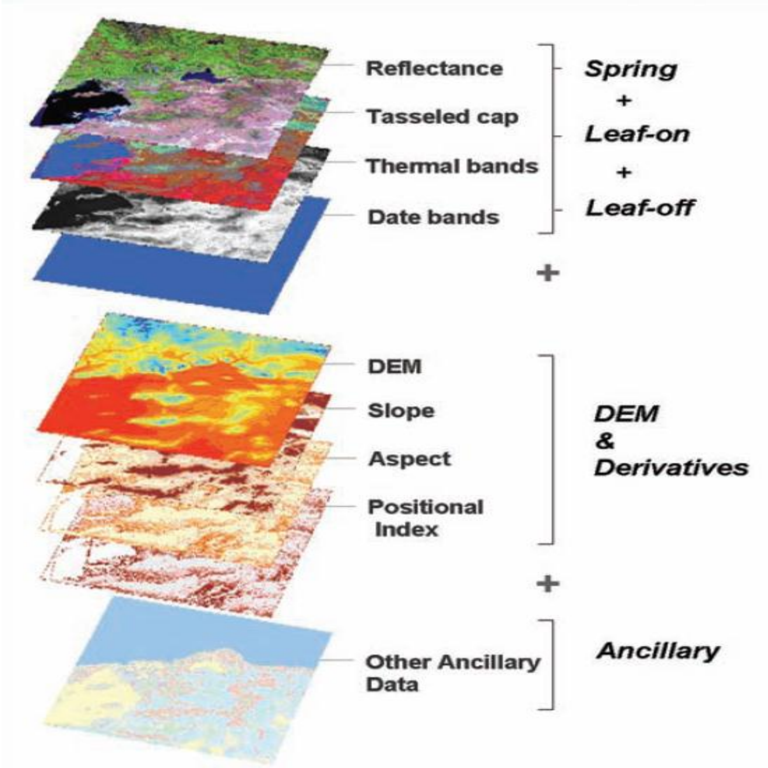


Fig. from Homer et al. 2007

Response developed by photo Interpreting Tree crown cover on NAIP Imagery for ~4160 8100m^2 sampling chips per Landsat scene.

Example modelling techniques

- Random forests
- Stochastic gradient boosting
- Support vector machines

Questions

Q1. Which modeling tool should we use?

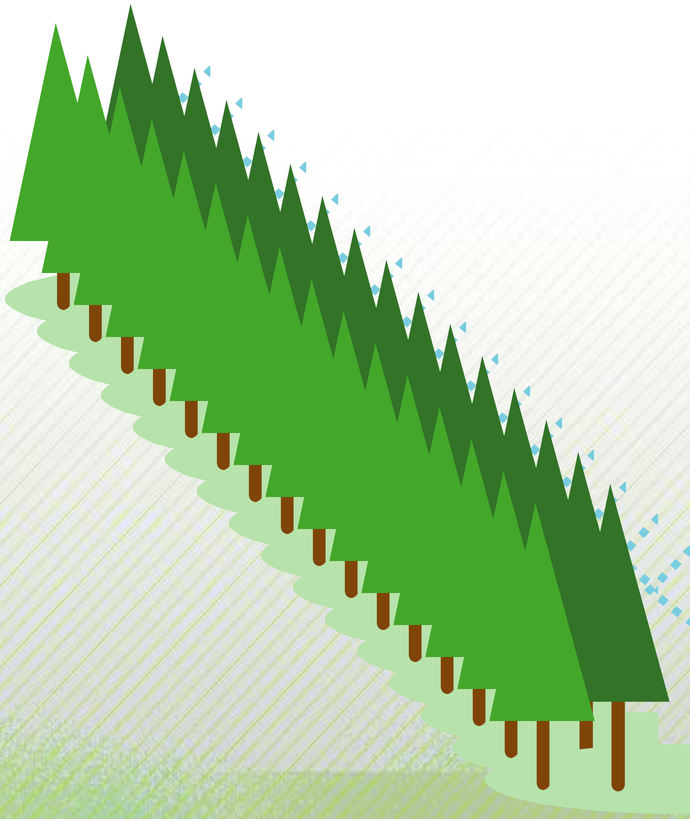
Q2. How big should our modeling zones be?

Q3. Would an initial masking model help?

Q4. How many sample plots/dots are needed?

Q5. What's the effect of repeatability?

Random Forests (Breiman, 2001)



Ensemble method using many classification and regression trees. One tree's construction:

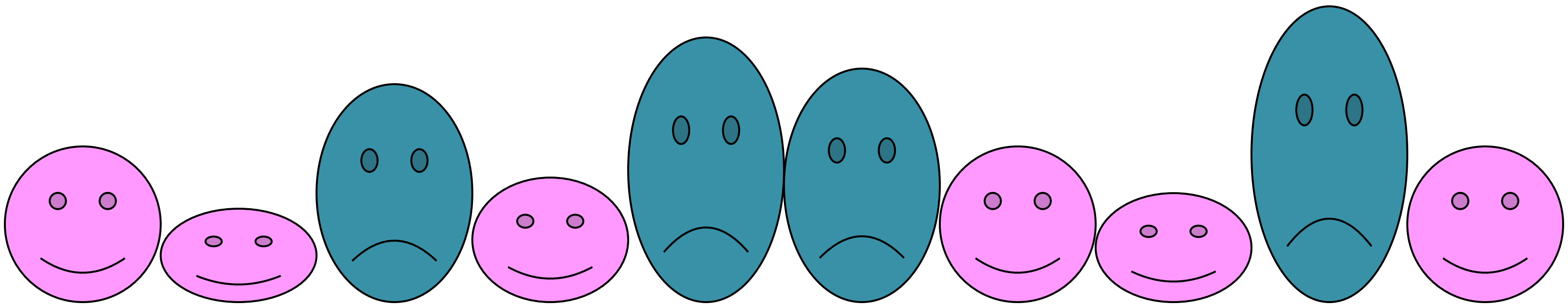
- Bootstrap sample from training data with replacement
- At each node a random sample of predictor variables is chosen and best split chosen
- Tree grown to maximum size (no pruning)
- Process repeated with many tree, each of which gets a vote, or is averaged for model predictions
- OOB error estimates

randomForest – Liaw & Wiener

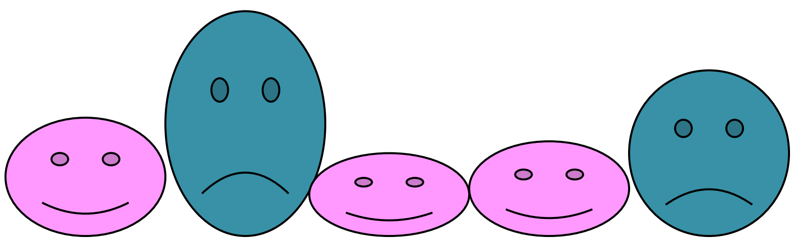
Stochastic Gradient Boosting (Freidman 2000, 2001, 2002)
Generalized Boosted Models (Ridgeway 1999)
Boosted Regression Trees (Elith et al. 2008)

1. Initialize the process using all the data

2. Compute "pseudo" residuals for all the data



3. Sample a subset of these pseudo-residuals without replacement



4. Fit a very small classification or regression tree to the residuals and create updates for each terminal node

5. Update the initial model

6. Repeat steps 2 -5 until convergence

gbm – Ridgeway

ModelMap

R package developed by Liz Freeman and Tracey Frescino, new version recently loaded on CRAN

model.build() : constructs predictive models of continuous or discrete responses using Random Forests or Stochastic Gradient Boosting

model.diagnostics(): validates these models with an independent test set, cross-validation, or OOB predictions on the training data and creates graphs and tables of the model validation

model.mapmake(): applies the models to GIS image of predictors to create prediction surfaces

model.interaction.plot(): provides a diagnostic plot useful in visualizing two-way interactions between predictor variables

Comparing RF and SGB

- All pilot areas: GA, MI, KS, UT, OR
- 25% Tuning, 50 % Training, 25% Test
- Response: TCC from photo interpretation
- Predictors

  Landsat (all bands but TIR)

  Topography (elevation, slope, aspect , CTI)

  Ecoregion

  Previous NLCD canopy cover and land cover class

Comparing RF and SGB

Usual accuracy metrics on independent test set

Tuning process
Density distributions of predictions on test set
Variable importance
Variable interactions
Maps (Utah)
Density distributions of maps (Utah)

Tuning Process

RF: **mtry -** number of explanatory randomly selected
(2, 4, 8)
ntrees- number of trees in the forest
(100, 200, …, 2500)
20 models for each value of mtry built,
and subtests run on increasing number of trees

SGB: **n.trees** – number of iterations
interaction.depth – tree complexity
shrinkage – learning rate
bag.fraction – training fraction
built 10 models of 6000 trees each for ranges shown above to optimize parameters

Tuning Results

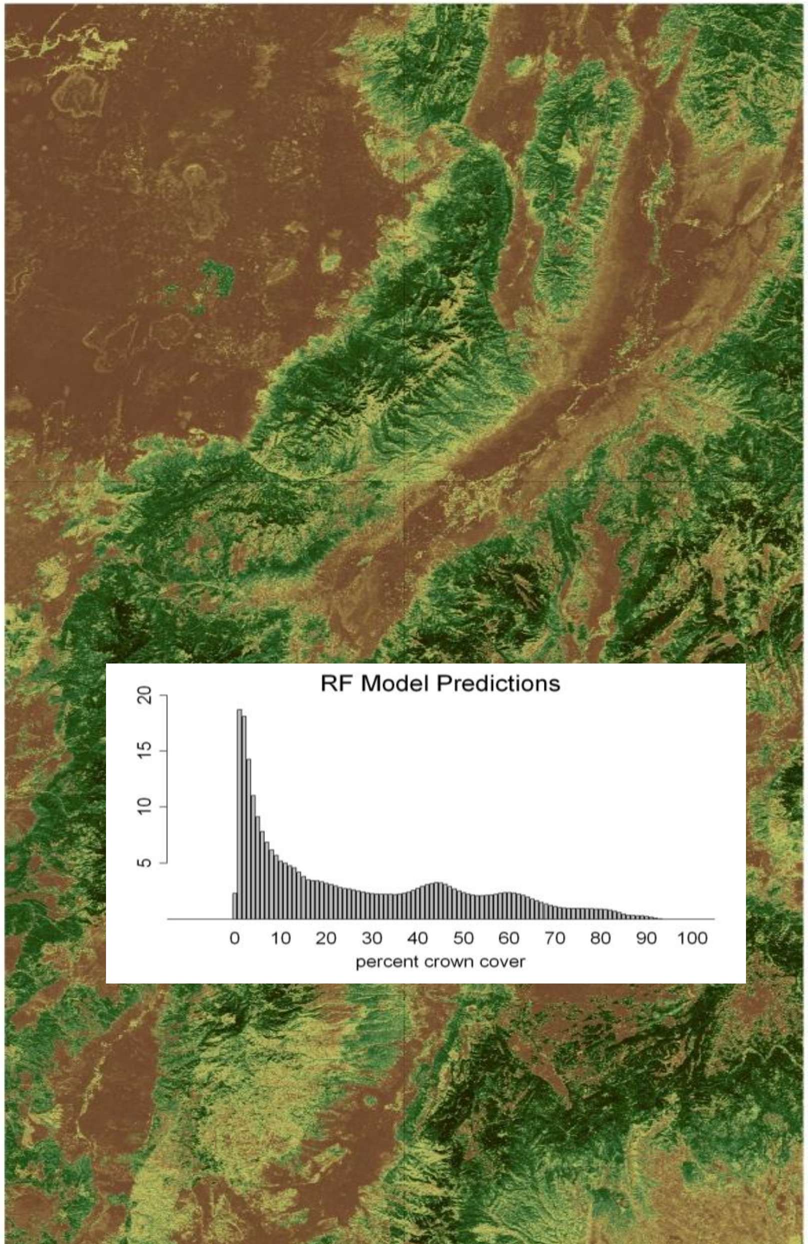
	Random Forest		Stochastic Gradient Boosting				
	ntrees	mtry	n.trees	best n.trees	Interaction Depth	Bag Fraction	Shrinkage
Georgia	2000	4	5000	3581	10	0.4	0.002
Kansas	2000	8	6000	4109	10	0.7	0.001
Michigan	2000	8	5000	4033	4	0.2	0.001
Oregon	2000	4	6000	4174	10	0.2	0.002
Utah	2000	8	6000	4490	10	0.5	0.002

Comparing RF and SGB

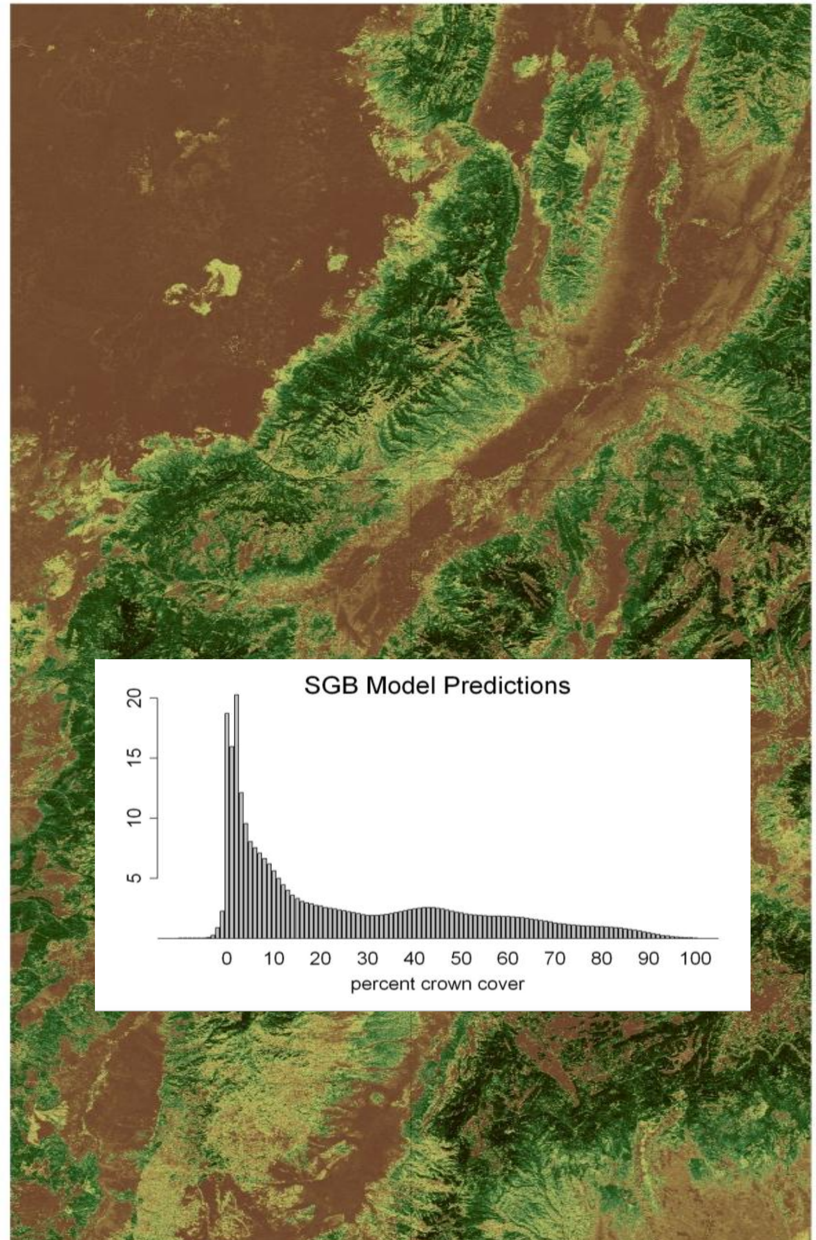
Stochastic Gradient Boosting								
state	Obs Mean	Pred Mean	Diff	pearson	spearman	MSE	slope	intercept
GA	0.65	0.66	-0.01	0.92	0.88	0.02	1.01	-0.01
KS	0.13	0.13	0.00	0.90	0.67	0.01	1.04	0.00
MI	0.45	0.44	0.01	0.90	0.85	0.04	1.04	0.00
OR	0.41	0.41	0.00	0.90	0.88	0.02	0.97	0.01
UT	0.30	0.29	0.01	0.83	0.85	0.03	1.00	0.01

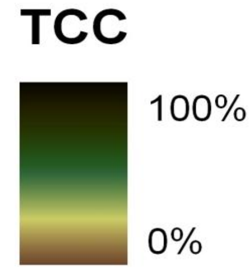
Random Forest								
state	Obs Mean	Pred Mean	Diff	pearson	spearman	MSE	slope	intercept
GA	0.65	0.66	0.00	0.92	0.88	0.02	1.03	-0.02
KS	0.13	0.13	0.00	0.90	0.66	0.01	1.04	0.00
MI	0.45	0.43	0.02	0.90	0.87	0.04	1.02	0.01
OR	0.41	0.41	0.00	0.90	0.88	0.02	1.00	0.00
UT	0.30	0.29	0.01	0.83	0.85	0.03	1.01	0.00

RF

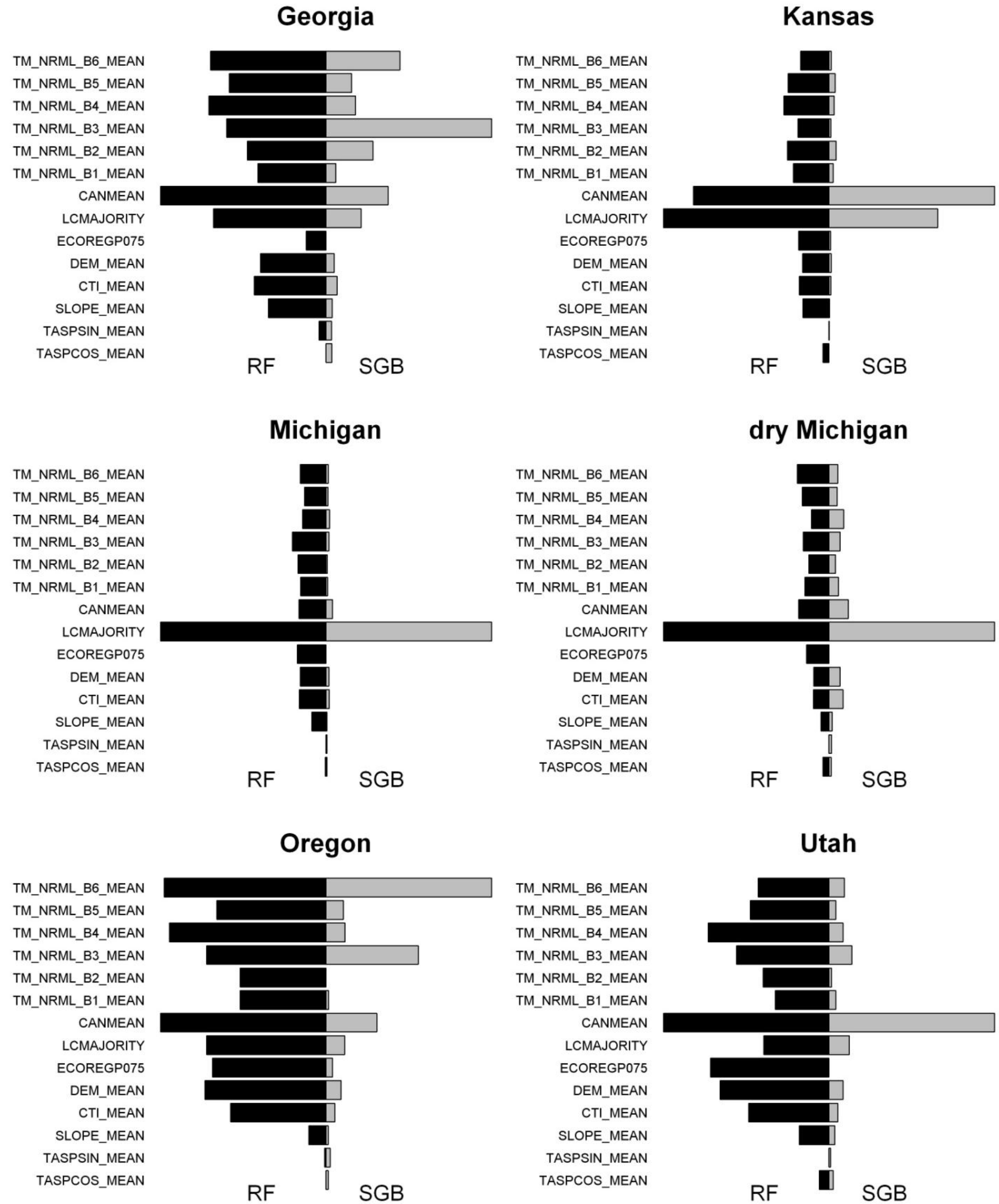

RF Model Predictions
20
15
10
5
0 10 20 30 40 50 60 70 80 90 100
percent crown cover

SGB

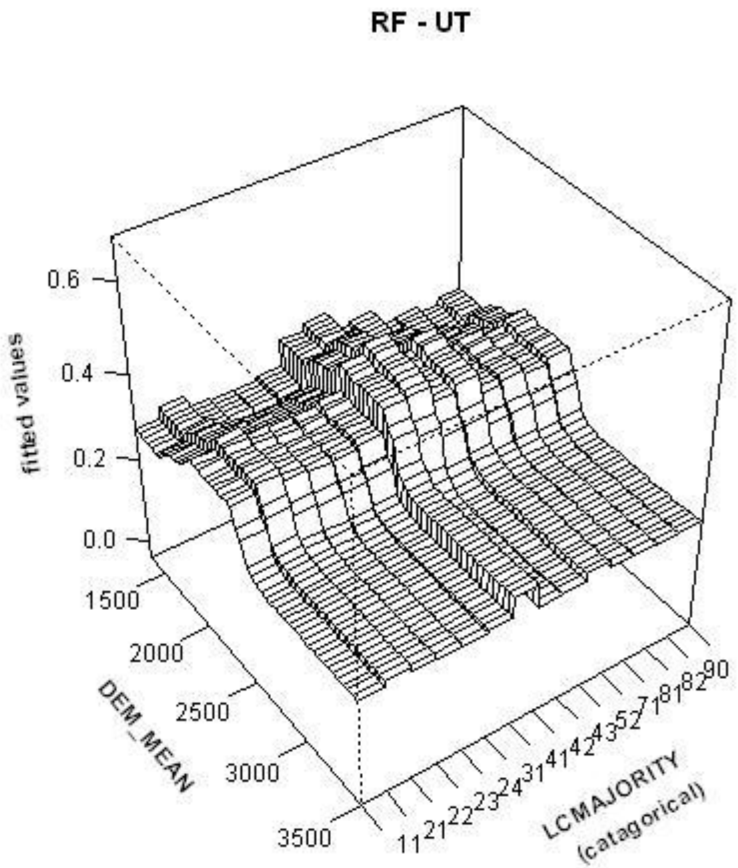

SGB Model Predictions
20
15
10
5
0 10 20 30 40 50 60 70 80 90 100
percent crown cover

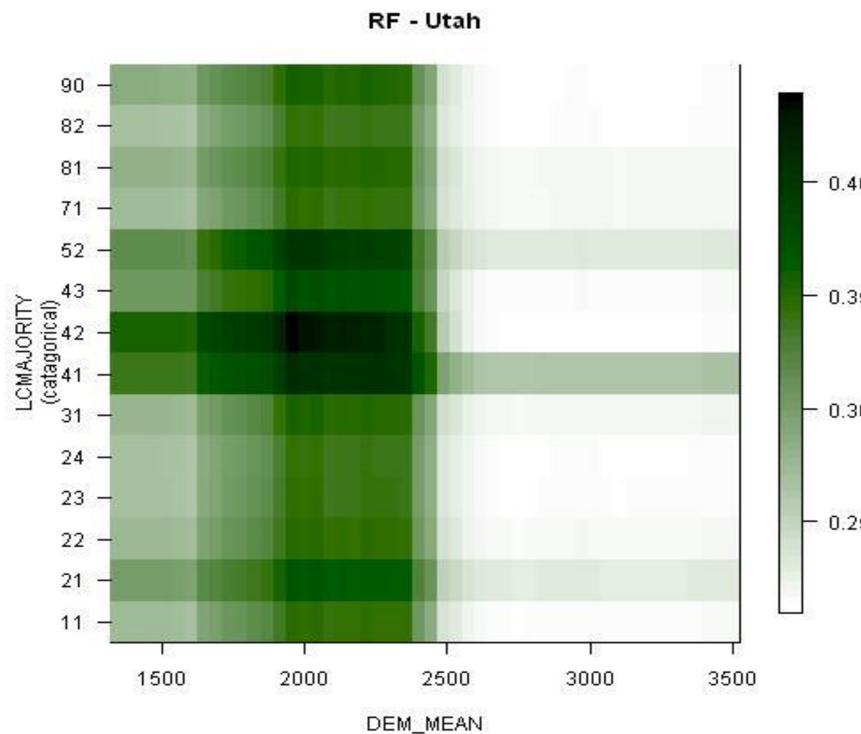

TCC
100%
0%

Variable importance



Interactions





Questions

Q1. Which modeling tool should we use?

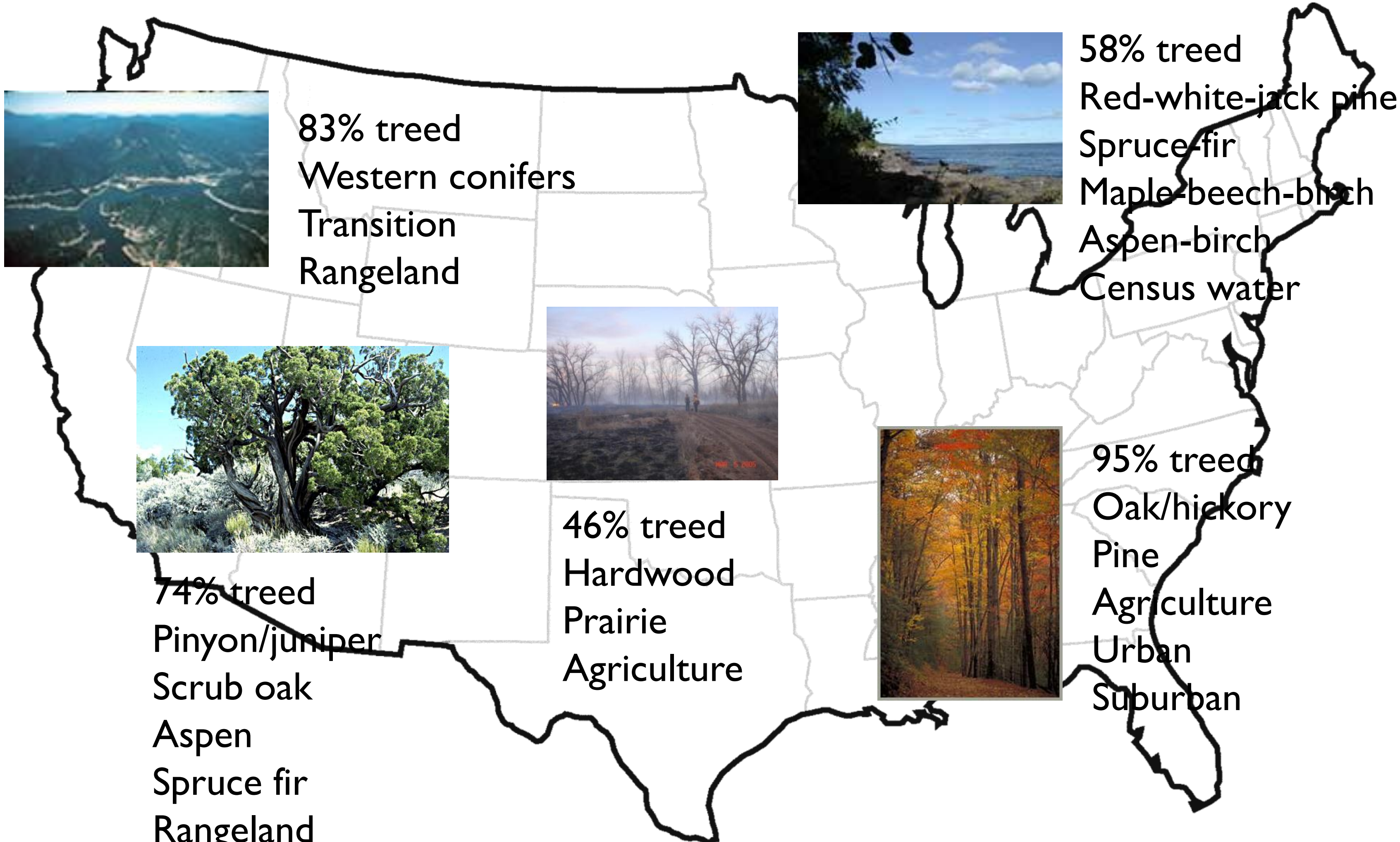
Q2. How big should our modeling zones be?

Q3. Would an initial masking model help?

Q4. How many sample plots/dots are needed?

Q5. What's the effect of repeatability?

Modeling TCC in diverse forests



Previous national mapping efforts

Numerous mapping zones and within-zone models

Small vs. large mapping zones

Using RF, 25% test, modeled TCC as functions of large set of the predictors

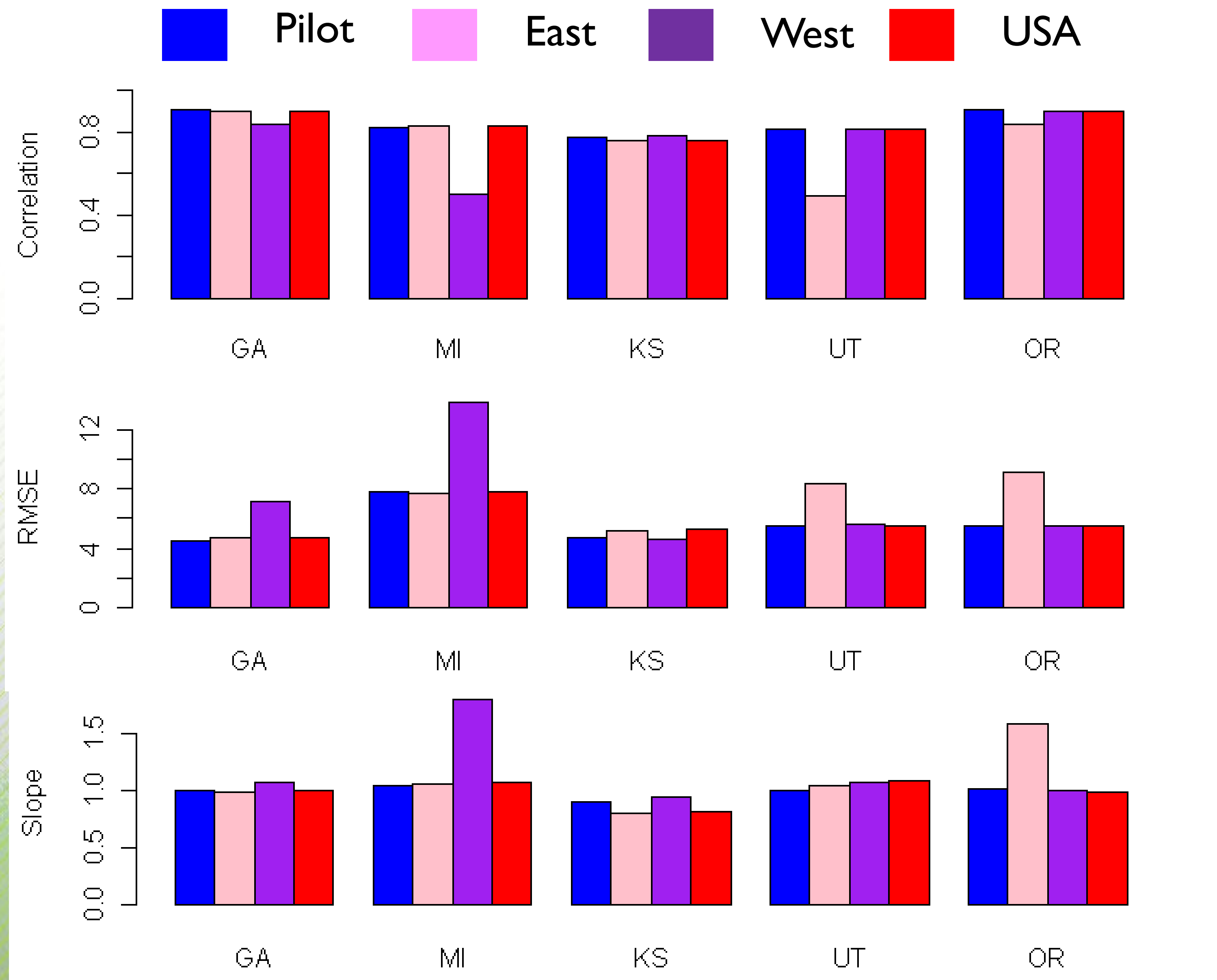
5 'Pilot' models: used data from respective pilot areas

1 'East' model: used data from GA, MI, KS

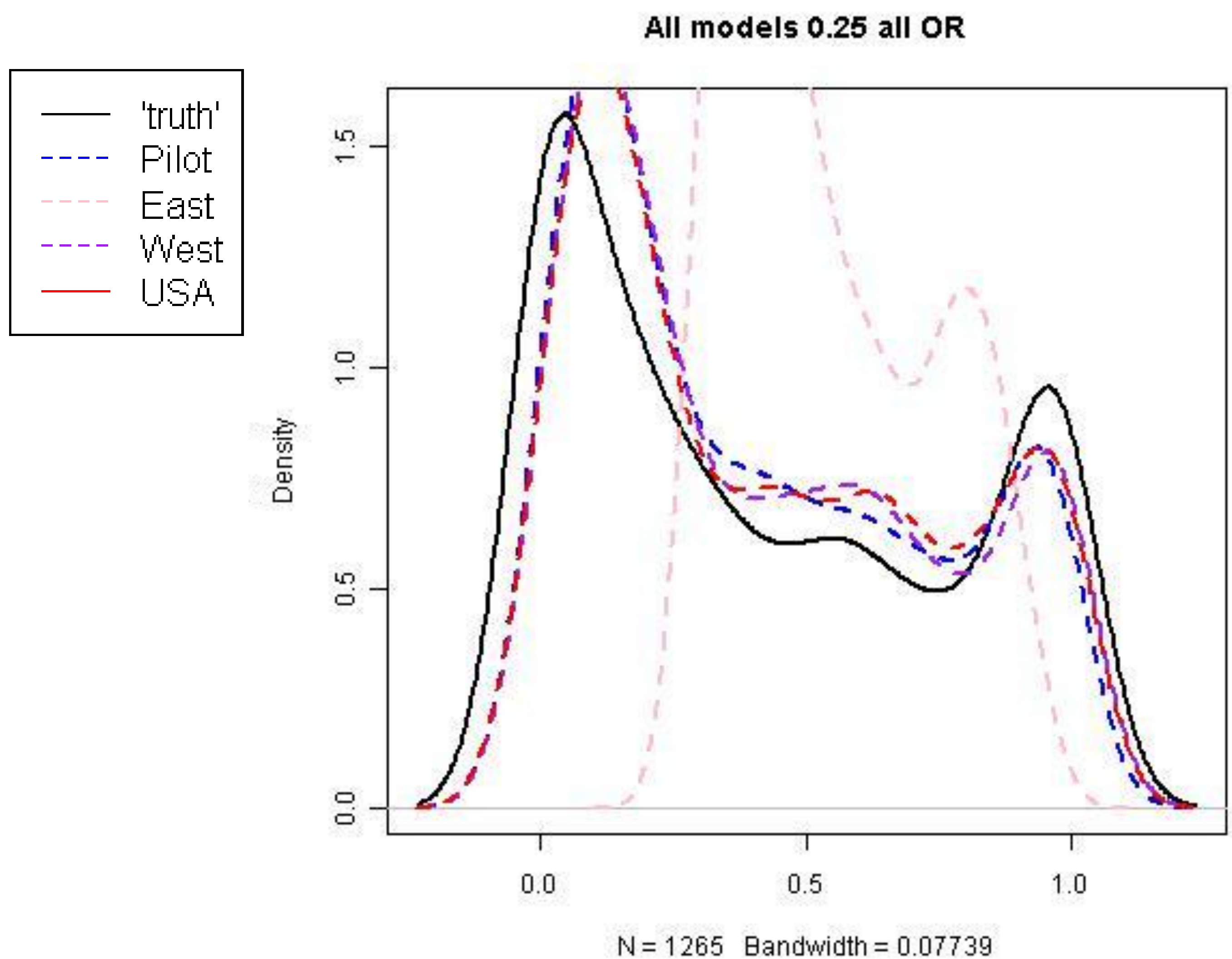
1 'West' model: used data from OR, UT, KS

1 'US' model: used data from all 5 pilot areas

Effect of larger models on accuracy metrics



Effect of larger models on densities



Questions

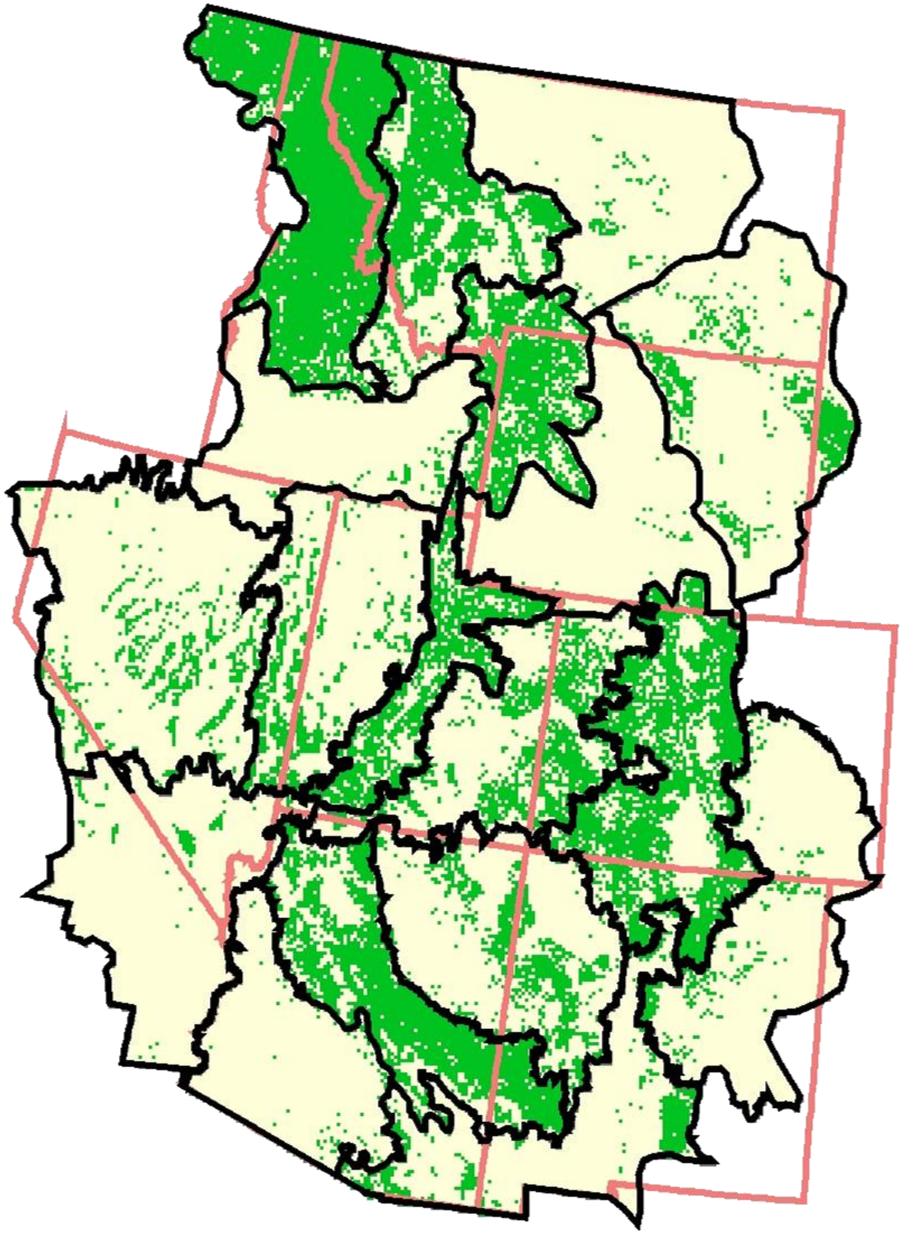
Q1. Which modeling tool should we use?

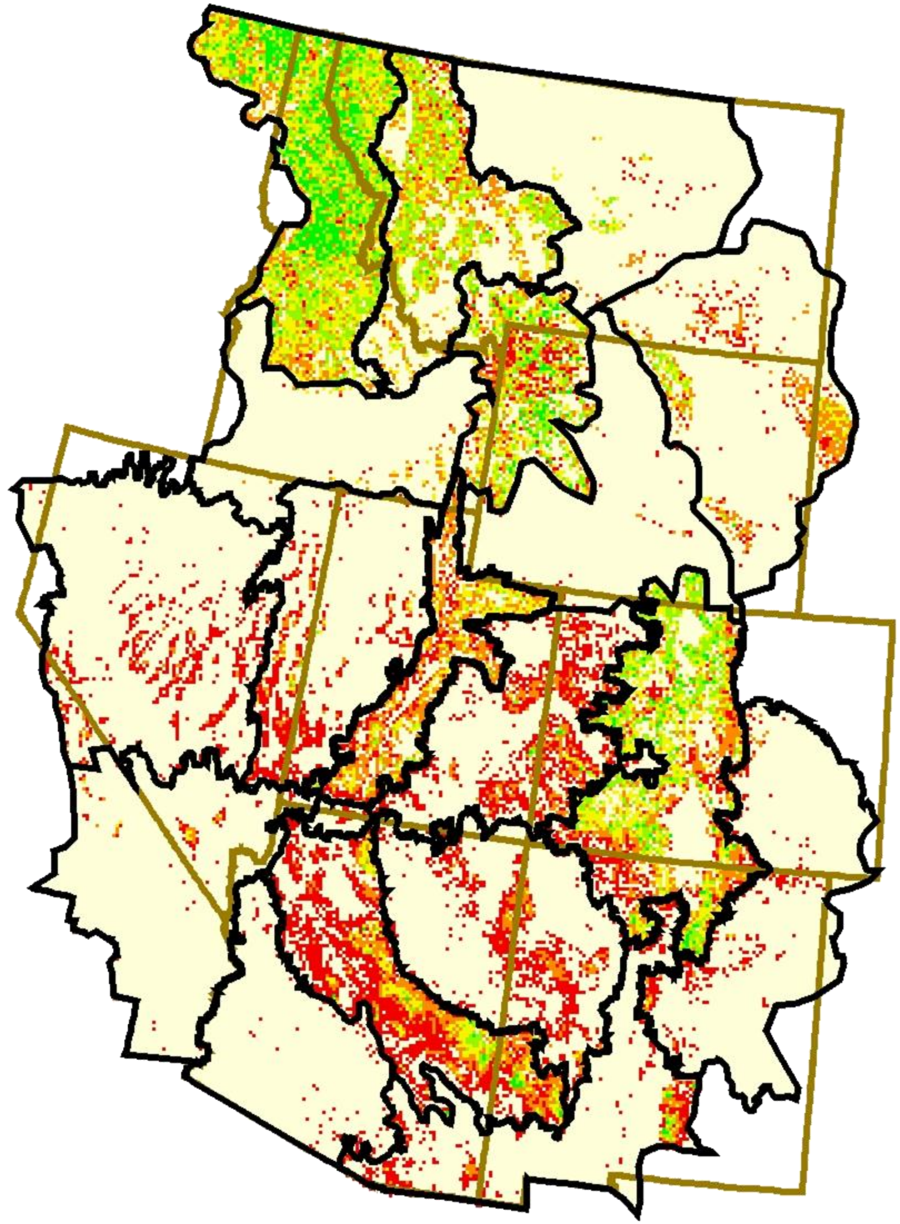
Q2. How big should our modeling zones be?

Q3. Would an initial masking model help?

Q4. How many sample plots/dots are needed?

Q5. What's the effect of repeatability?


Forest/nonforest


Biomass

To mask or not to mask

Using RF, 25% test set

- Model 1: tree presence/absence (TNT)
- Predicted TNT over trg and test, setting threshold: PredPrev=Obs
- Model 2: model TCC using only data where predicted TNT=1
- Predict TCC over test set from model 2
- Set TCC= 0 where predicted TNT = 0

PresenceAbsence

R package developed by Liz Freeman on the CRAN

Provides a set of functions useful when evaluating the results of presence-absence models.

Functions for calculating threshold dependant measures such as confusion matrices, PCC sensitivity, specificity, and Kappa

Produces plots of each measure as the threshold is varied

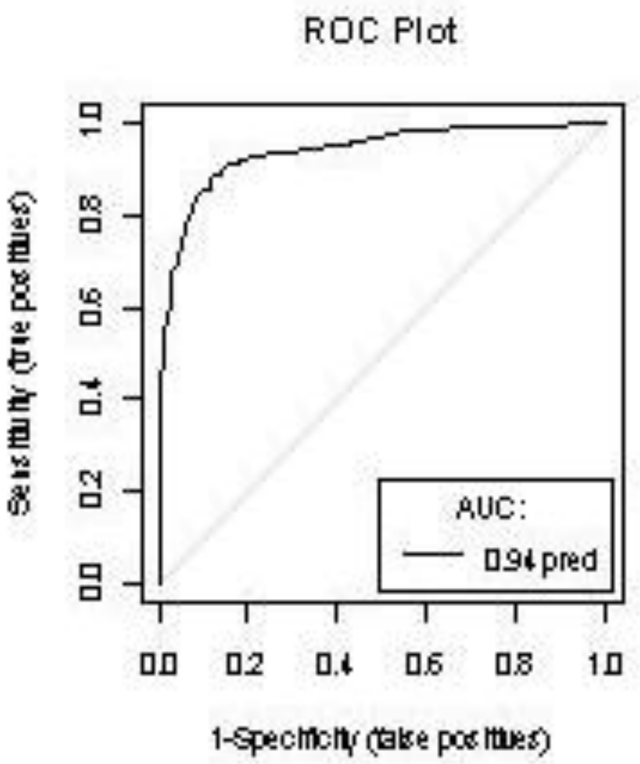
Plots threshold-independent ROC curves along with the associated AUC (area under the curve)

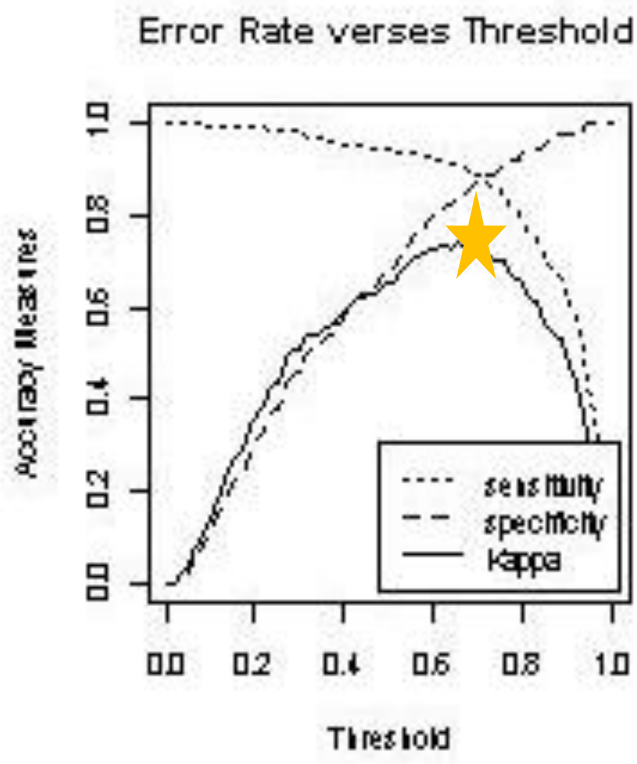
Thresholding options in PresenceAbsence

Default: threshold=0.5
Sens=Spec : sensitivity=specificity
MaxSens+Spec: maximizes (sensitivity+specificity)/2
MaxKappa: maximizes Kappa
MaxPCC: maximizes PCC
PredPrev=Obs : predicted prevalence=observed prevalence
ObsPrev: threshold=observed prevalence
MeanProb: mean predicted probability
MinROCdist: minimizes dist between ROC plot and (0,1)
ReqSens: user defined required sensitivity
ReqSpec: user defined required specificity
Cost: user defined relative costs ratio

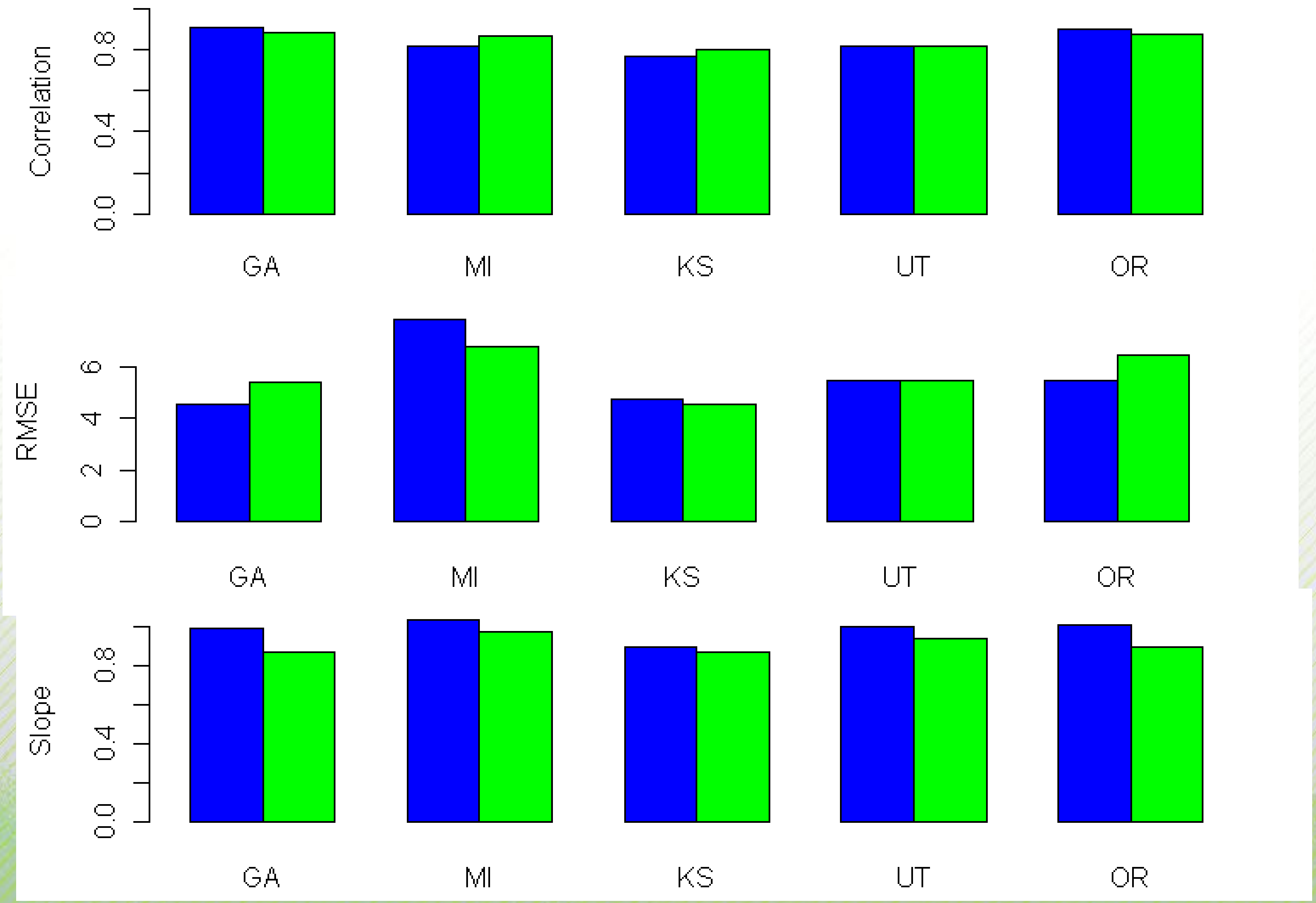
Example TNT model in UT

- Good model fit
- Threshold maximizing kappa ~ threshold at prevalence



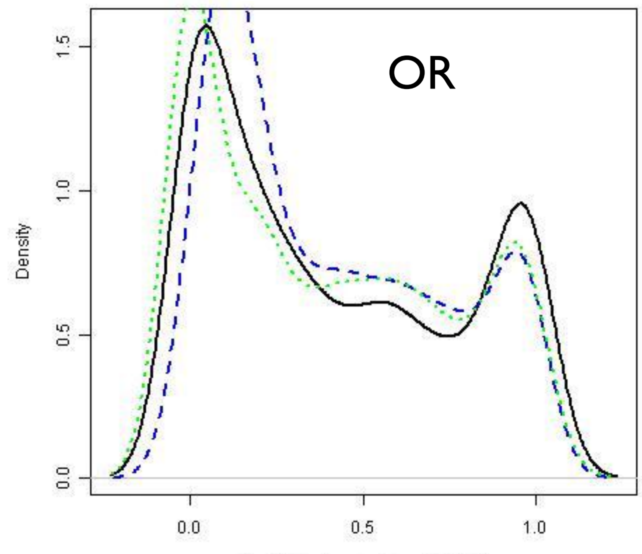


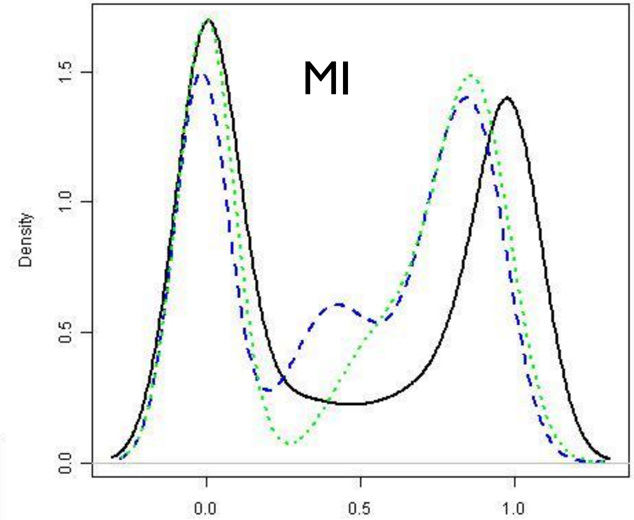
Effect of TNT mask on accuracy metrics

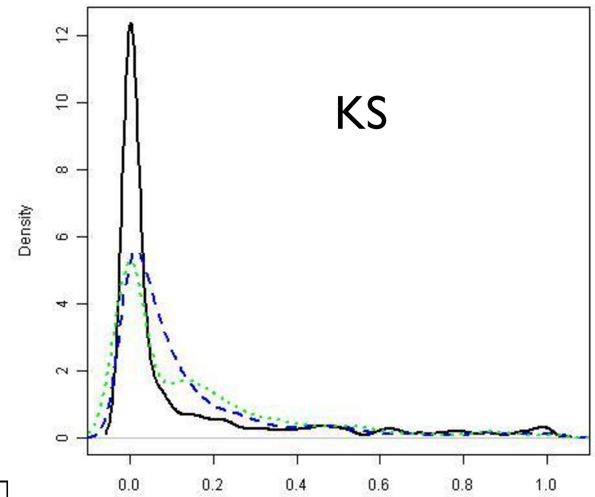


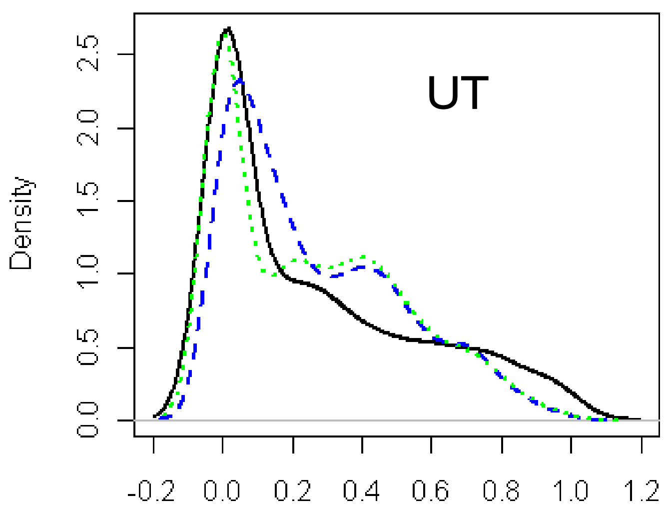
No mask TNT Mask

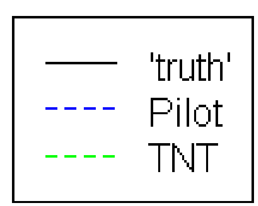
Effect of TNT models on densities

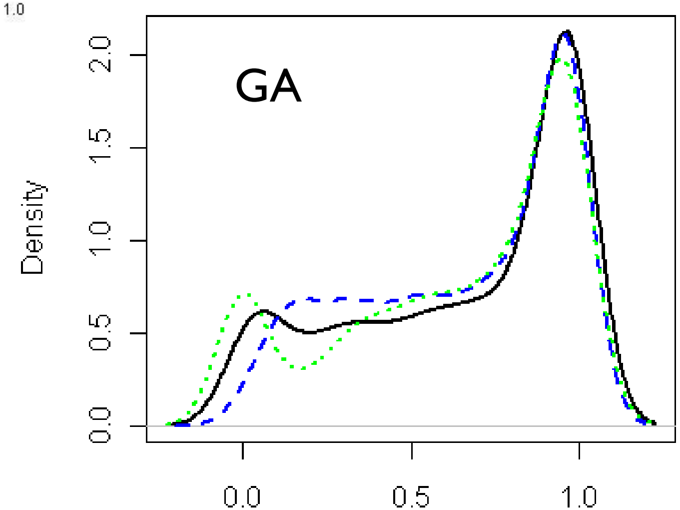


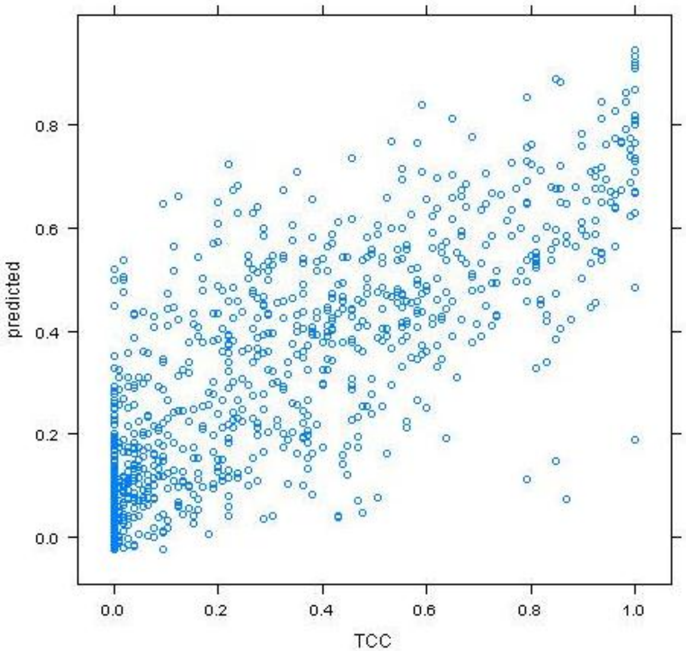


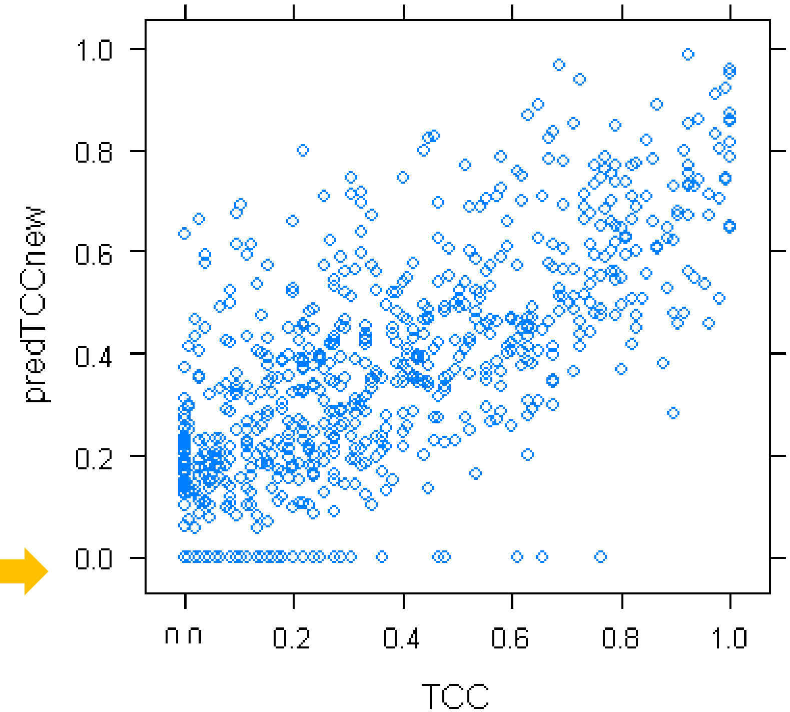


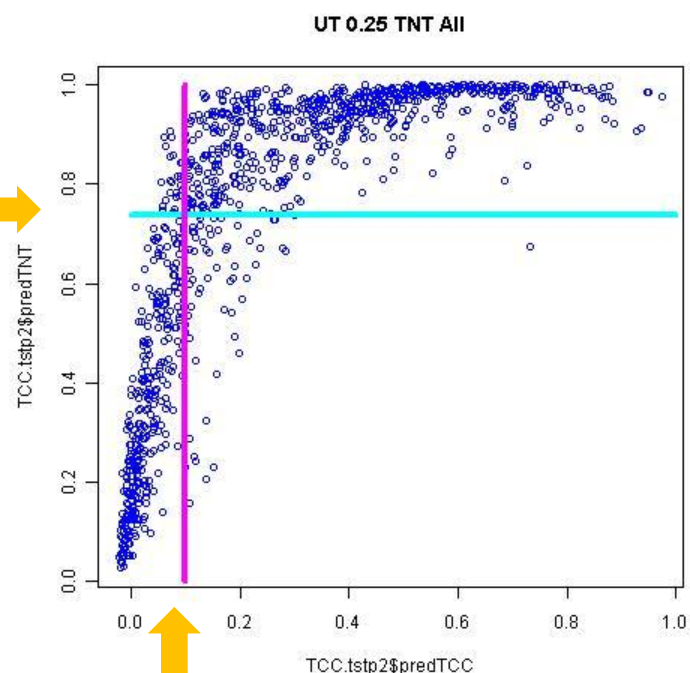






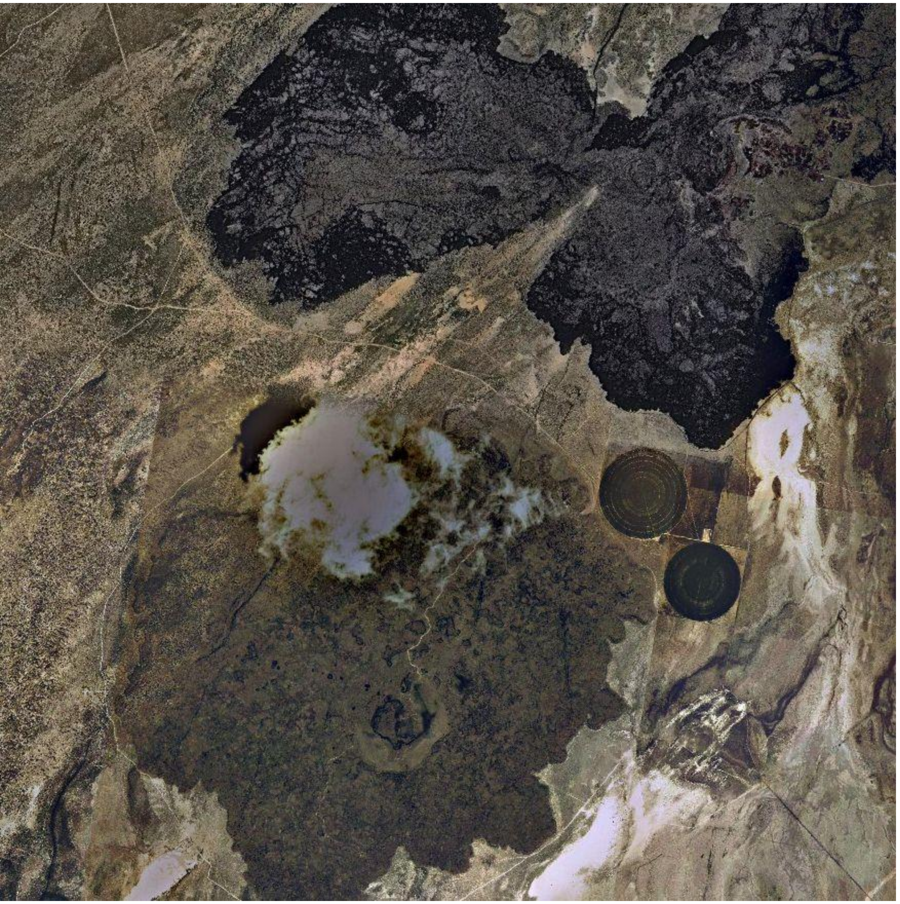

predicted
TCC
0.0
0.2
0.4
0.6
0.8
1.0

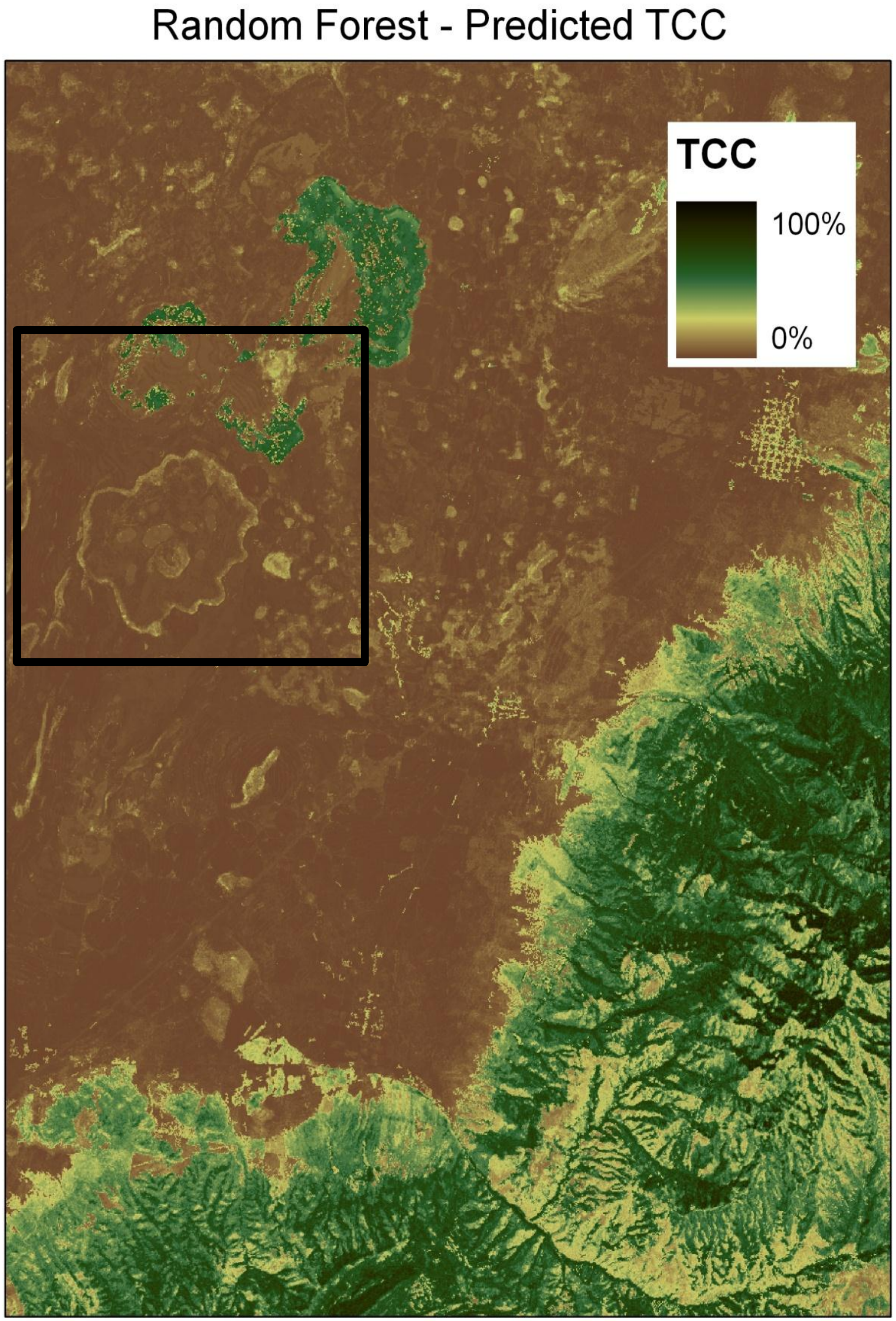

predTCCnew
TCC
0.0
0.2
0.4
0.6
0.8
1.0


UT 0.25 TNT All
TCC.tstp2$predTNT
TCC.tstp2$predTCC
0.0
0.2
0.4
0.6
0.8
1.0

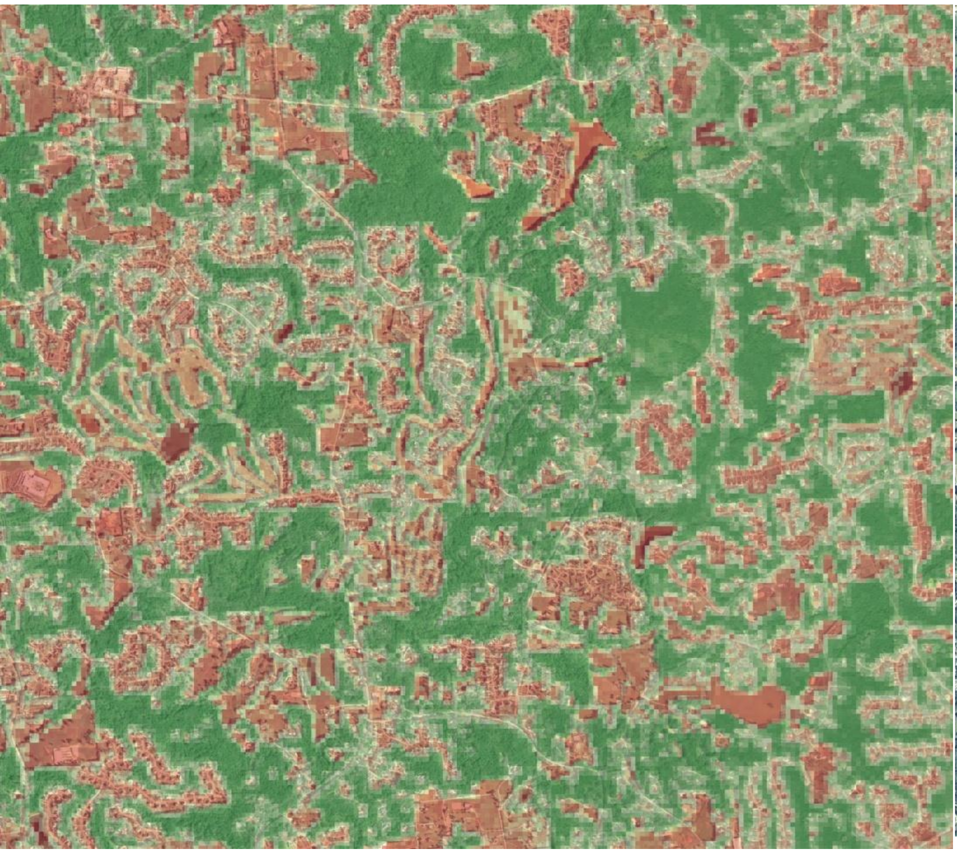
The reality….

Predicting trees where they're not





Random Forest model in GA



Predicted TCC

2009 NAIP

Questions

Q1. Which modeling tool should we use?

Q2. How big should our modeling zones be?

Q3. Would an initial masking model help?

Q4. How many sample plots/dots are needed?

Q5. What's the effect of repeatability?

FIA Grid and Intensification

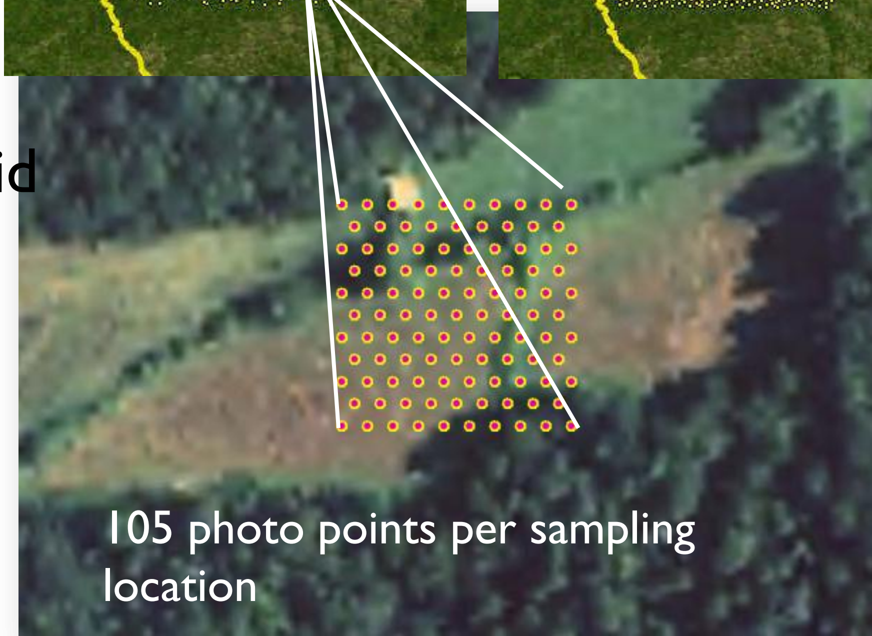
Standard Intensity FIA grid:

1. Modeled % canopy cover estimates
2. Photo Interpreted estimates

Intensified Grid: Photo Interpreted % canopy cover

<u>Definitions:</u>

- Sampling Grid
- Dot Grid

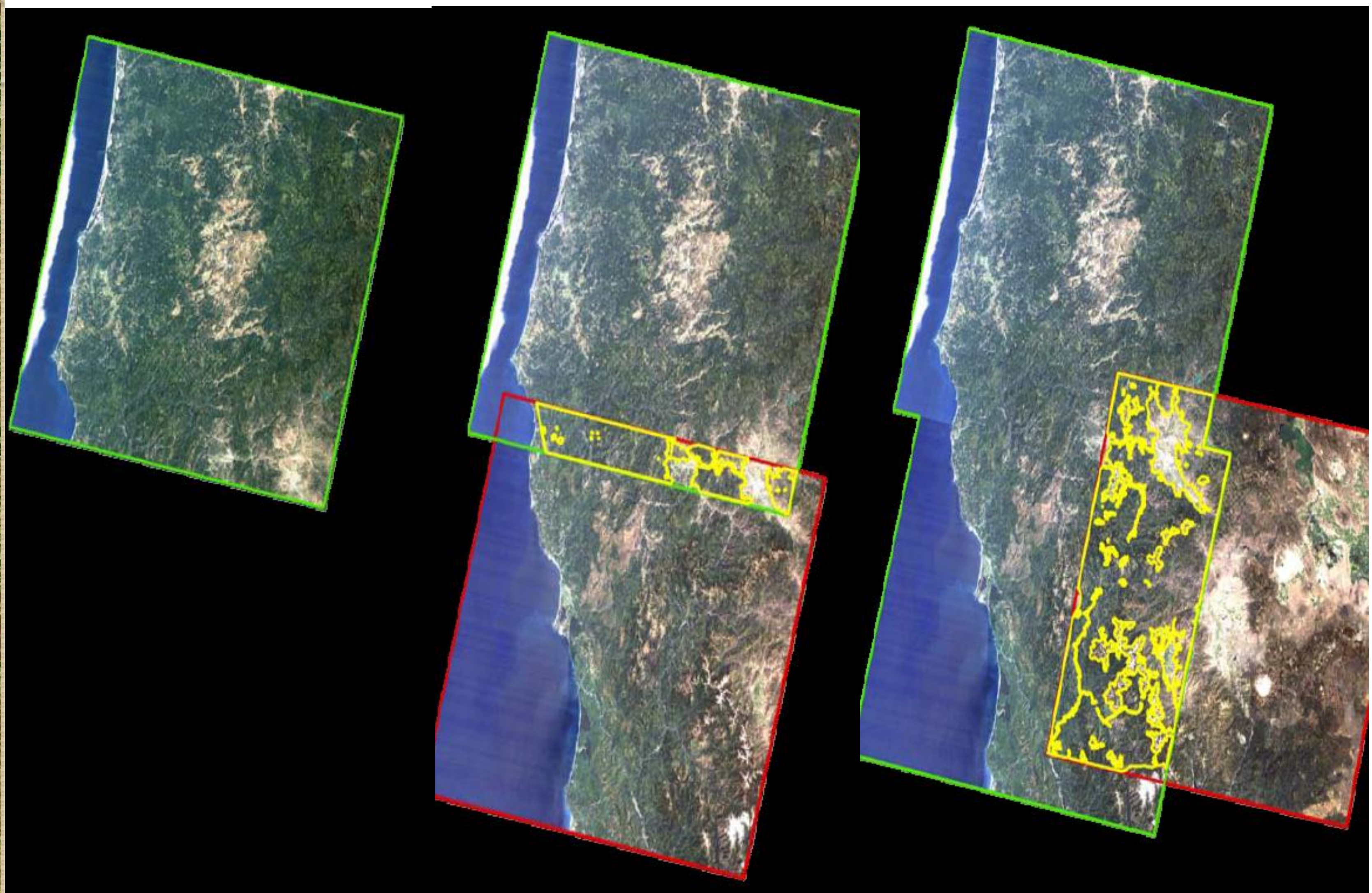


Plot Sample Intensity vs Model Fit

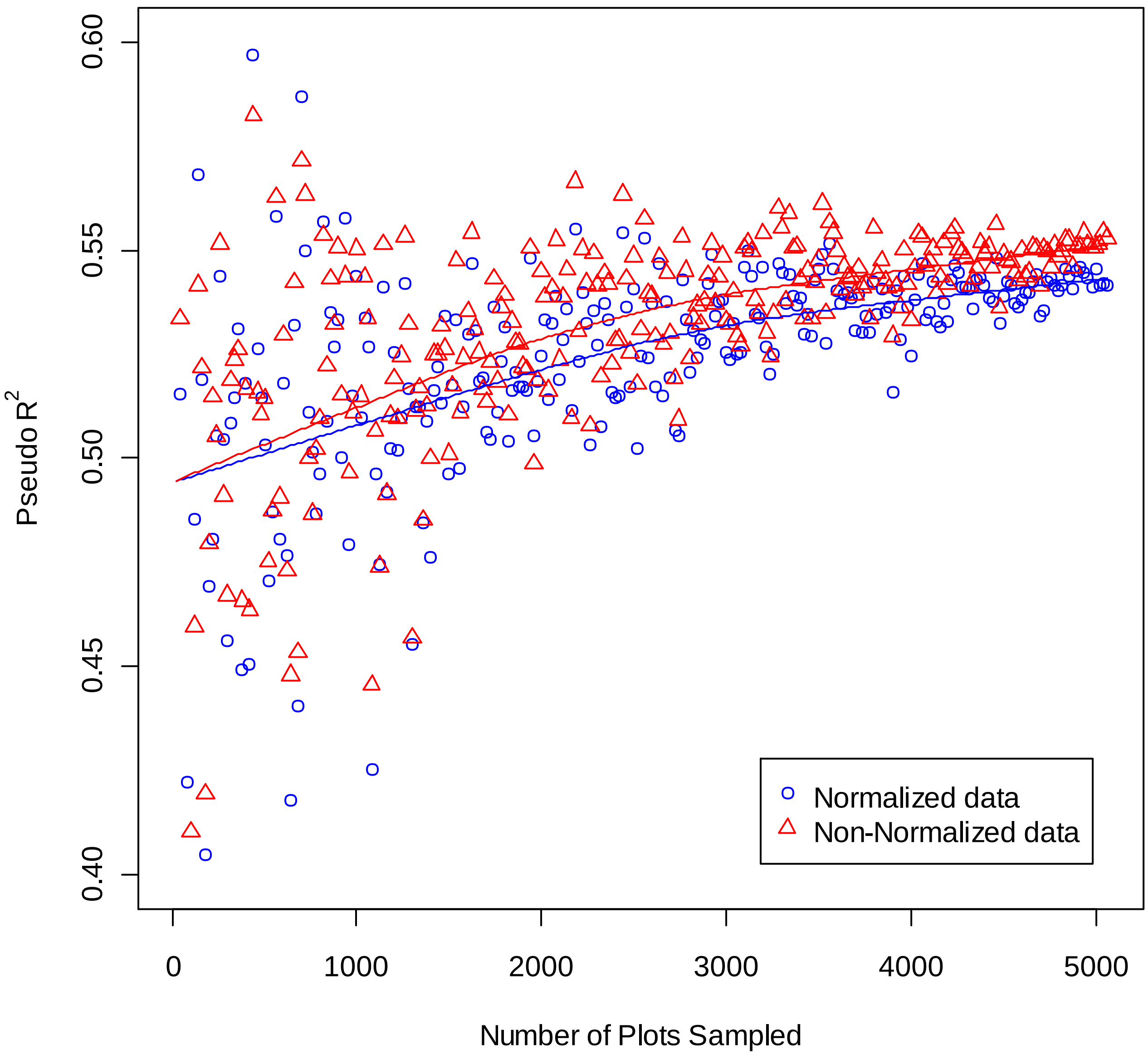
- What is the effect of the number of plots on model fit?
- What is the effect of the number of photo dots on model fit?

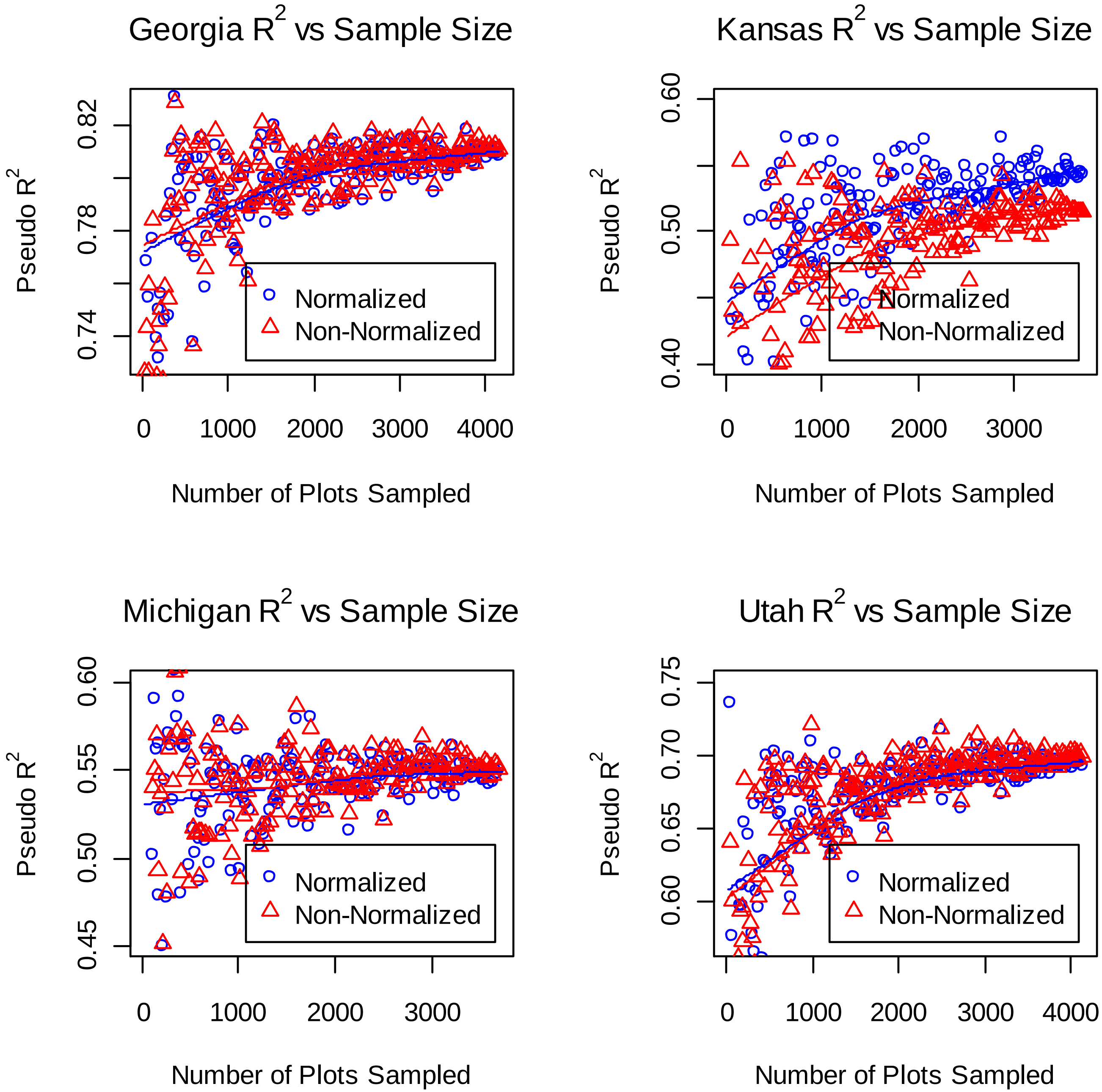
Bonus: For both normalized and non-normalized landsat-based predictors

Normalization?



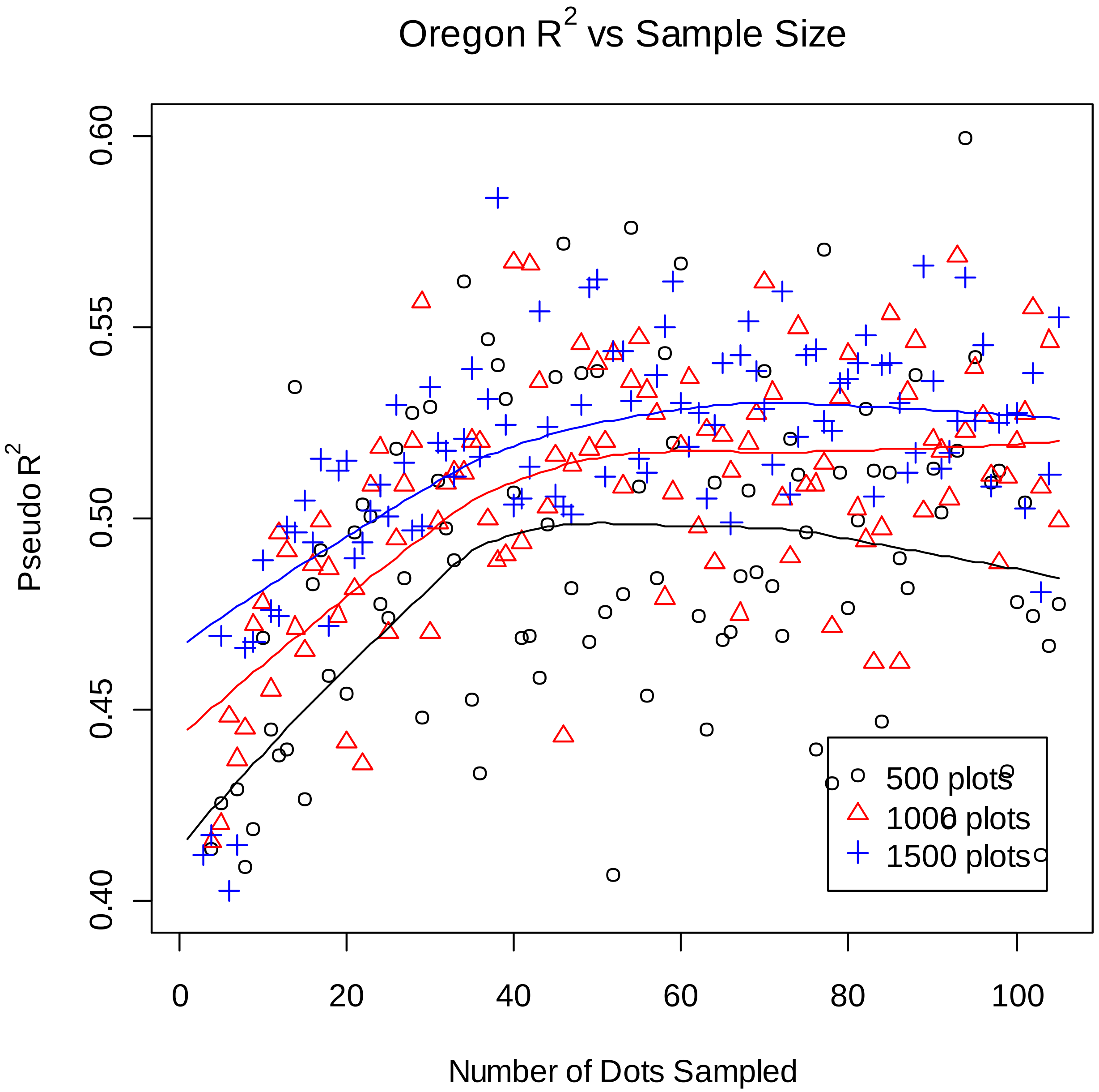
Oregon R^2 vs Sample Size



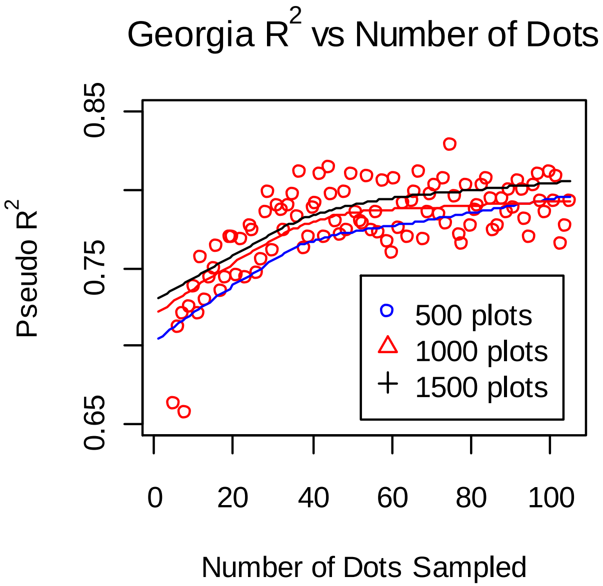

Georgia R2 vs Sample Size
Pseudo R2
0.74
0.78
0.82
0
1000
2000
3000
4000
Number of Plots Sampled
Normalized
Non-Normalized
Kansas R2 vs Sample Size
Pseudo R2
0.40
0.50
0.60
0
1000
2000
3000
Number of Plots Sampled
Normalized
Non-Normalized
Michigan R2 vs Sample Size
Pseudo R2
0.45
0.50
0.55
0.60
0
1000
2000
3000
Number of Plots Sampled
Normalized
Non-Normalized
Utah R2 vs Sample Size
Pseudo R2
0.60
0.65
0.70
0.75
0
1000
2000
3000
4000
Number of Plots Sampled
Normalized
Non-Normalized

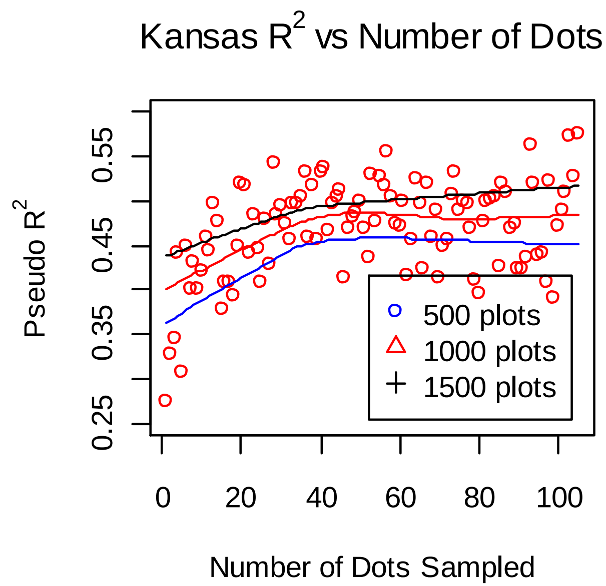
Dot Grid Sampling vs Model Fit

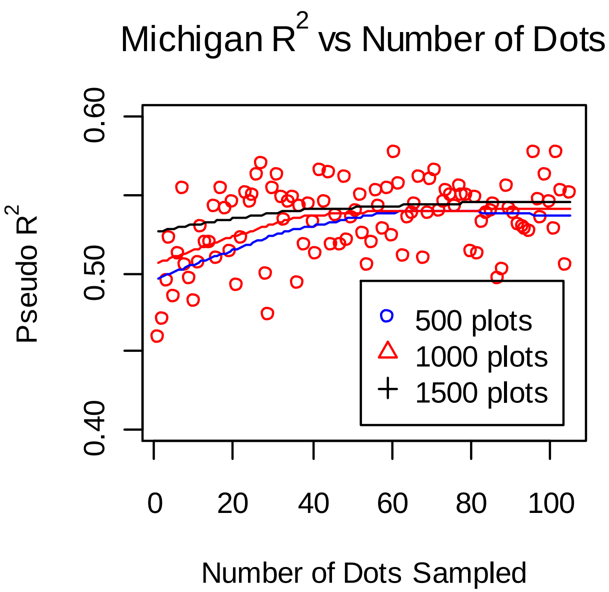
- What is the effect of the number of dots on model fit?
- For 500, 1000, and 1500 sample points we plot the number of dots vs model fit

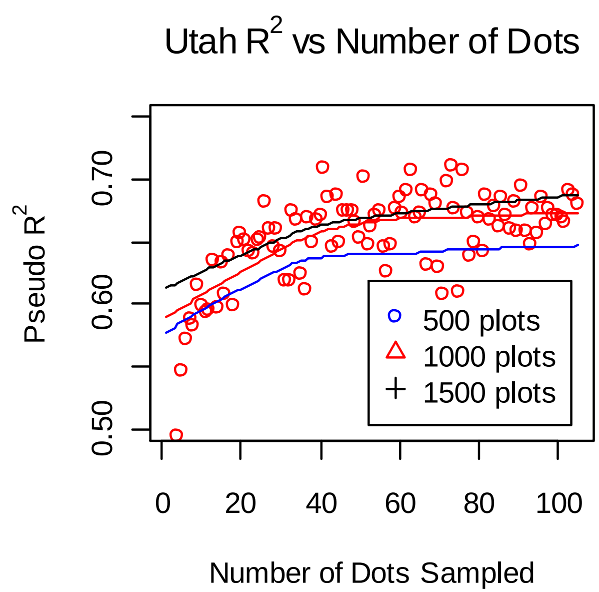


Oregon R2 vs Sample Size
Pseudo R2
0.60
0.55
0.50
0.45
0.40
0
20
40
60
80
100
Number of Dots Sampled
500 plots
1000 plots
1500 plots


Georgia R2 vs Number of Dots
Pseudo R2
0.65
0.75
0.85
0
20
40
60
80
100
Number of Dots Sampled
500 plots
1000 plots
1500 plots


Kansas R2 vs Number of Dots
Pseudo R2
0.25
0.35
0.45
0.55
0
20
40
60
80
100
Number of Dots Sampled
500 plots
1000 plots
1500 plots


Michigan R2 vs Number of Dots
Pseudo R2
0.40
0.50
0.60
0
20
40
60
80
100
Number of Dots Sampled
500 plots
1000 plots
1500 plots


Utah R2 vs Number of Dots
Pseudo R2
0.50
0.60
0.70
0
20
40
60
80
100
Number of Dots Sampled
500 plots
1000 plots
1500 plots

Questions

Q1. Which modeling tool should we use?

Q2. How big should our modeling zones be?

Q3. Would an initial masking model help?

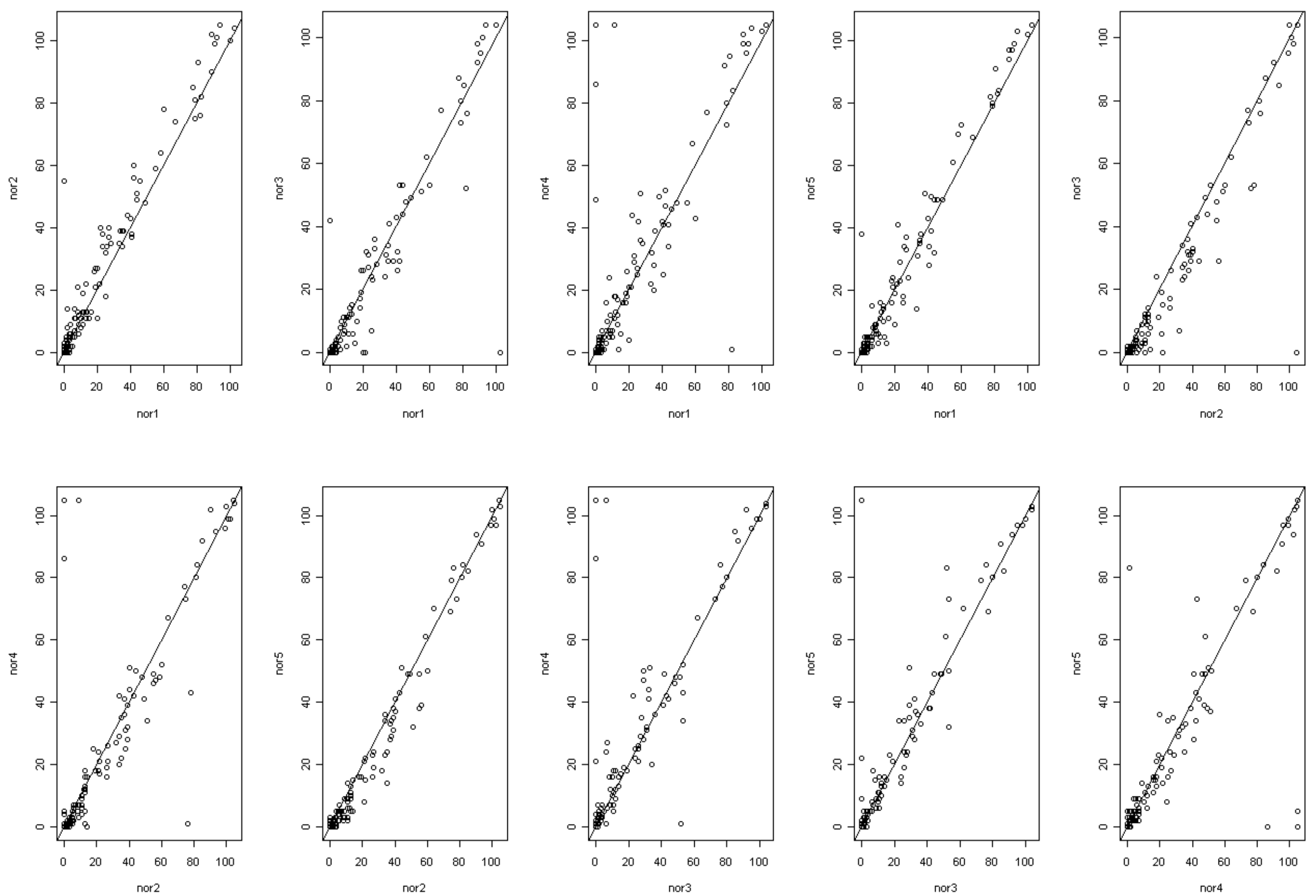
Q4. How many sample plots/dots are needed?

Q5. What's the effect of repeatability?

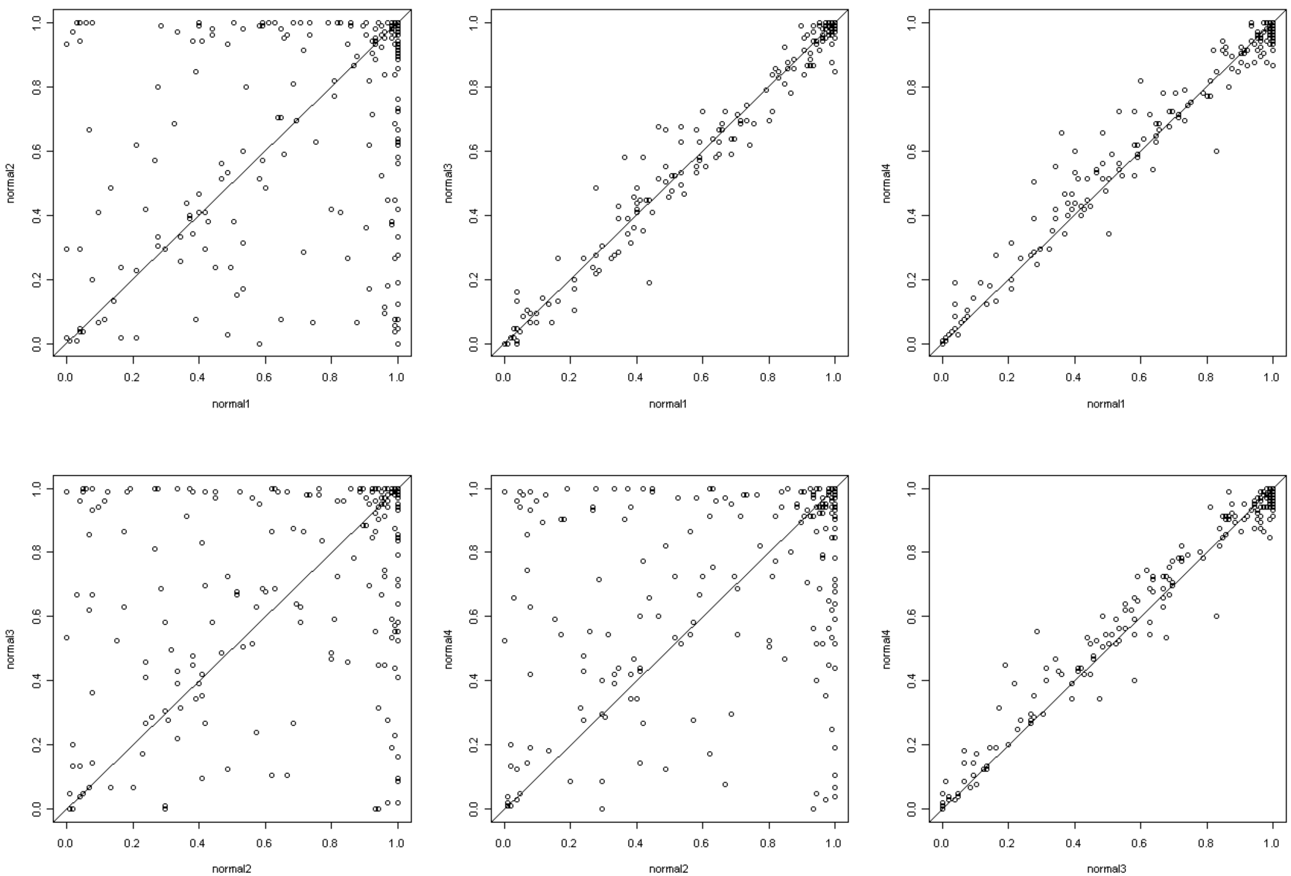
Repeatability data

- Each pilot area interpreted by between two and five interpreters
- 208 photo plots were identified in each pilot area as blind repeatability plots and interpreted by each interpreter

Kansas: Pairwise plots of TCC for 5 photo-interpreters



Georgia Percent Canopy Cover



ICC Values

- The Interclass Correlation Coefficient (ICC) is a measure of the agreement of two variable taking into account the variance, covariance and mean of each variable in question
- First proposed in 1901 by Pearson, Carrasco et al have expanded what was originally conceived to extend to more than two observers
- The ICC value is a score between 0 and 1, read like an R Squared

Georgia:	0.65
Kansas:	0.91
Michigan:	0.94
Oregon:	0.93
Utah:	0.87

Tolerance and Compliance

- Tolerance is the amount by which we are willing to let the observed data vary from the true data
- The Compliance percentage is the amount of the observed data that we are willing to let fall within the Tolerance amount

Sensitivity of models to varying levels of tolerance and compliance Results of 1,000 Simulations

Ratio of Perturbed R^2 to True R^2

30%	0.7832	0.7586	0.7085	0.6884	0.6470
20%	0.9078	0.8716	0.8400	0.8131	0.7984
10%	0.9744	0.9585	0.9562	0.9259	0.9234
	100%	**95%**	**90%**	**85%**	**80%**

Repeatability Results

- The ICC values for measuring repeatability in the data range from 0.65 to 0.94
- The Georgia Data set emphasizes the importance of consistent training
- Compliance rates for 10% tolerance ranged from 96% to 64%
- A tolerance of 20% with 85% compliance could result in 80% model perfomance

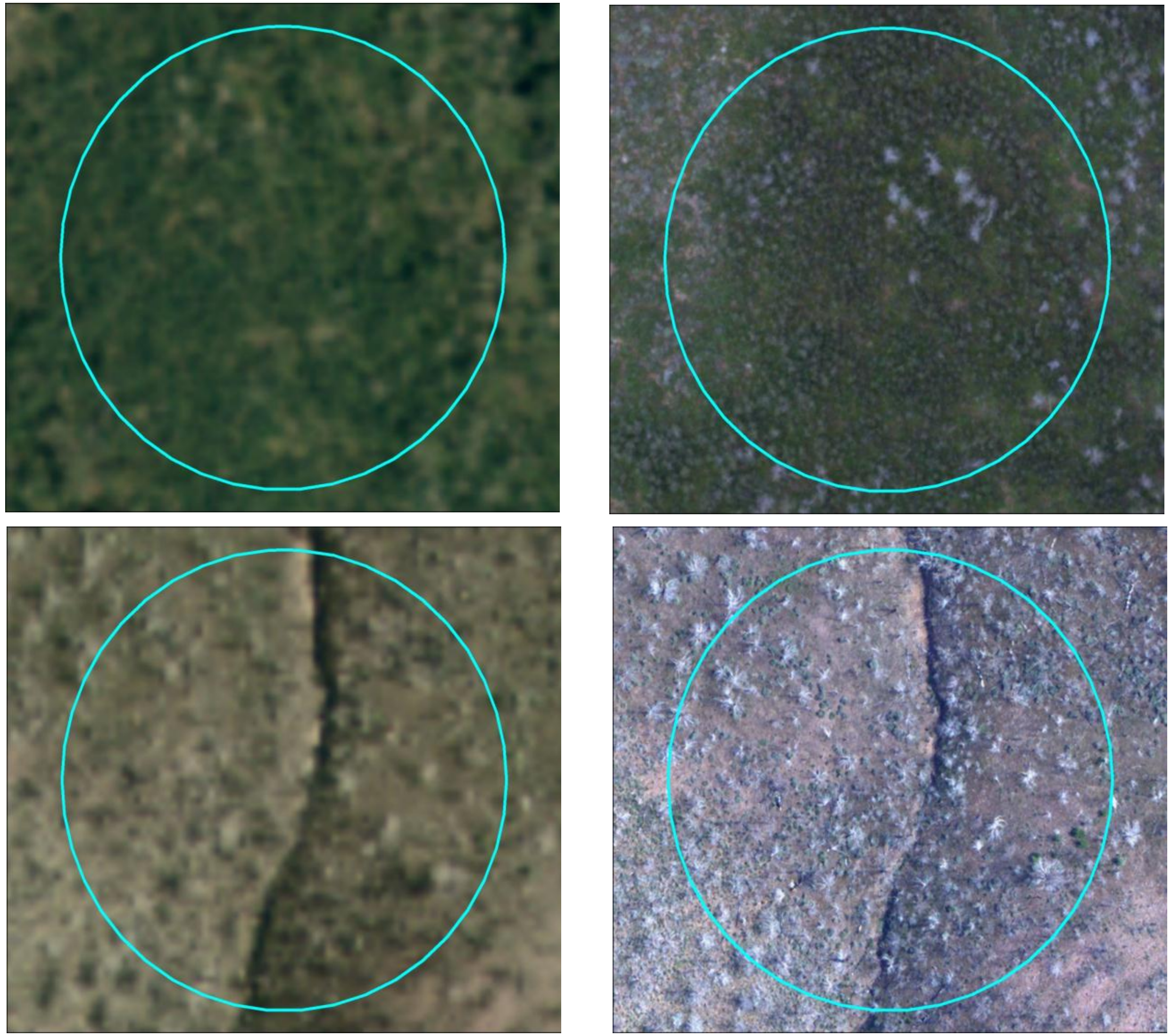
Many other issues "measuring" tree canopy cover

Quality of the map is dependent on the quality of the data used to train the models

Different methods for estimating live tree canopy cover can yield very different results

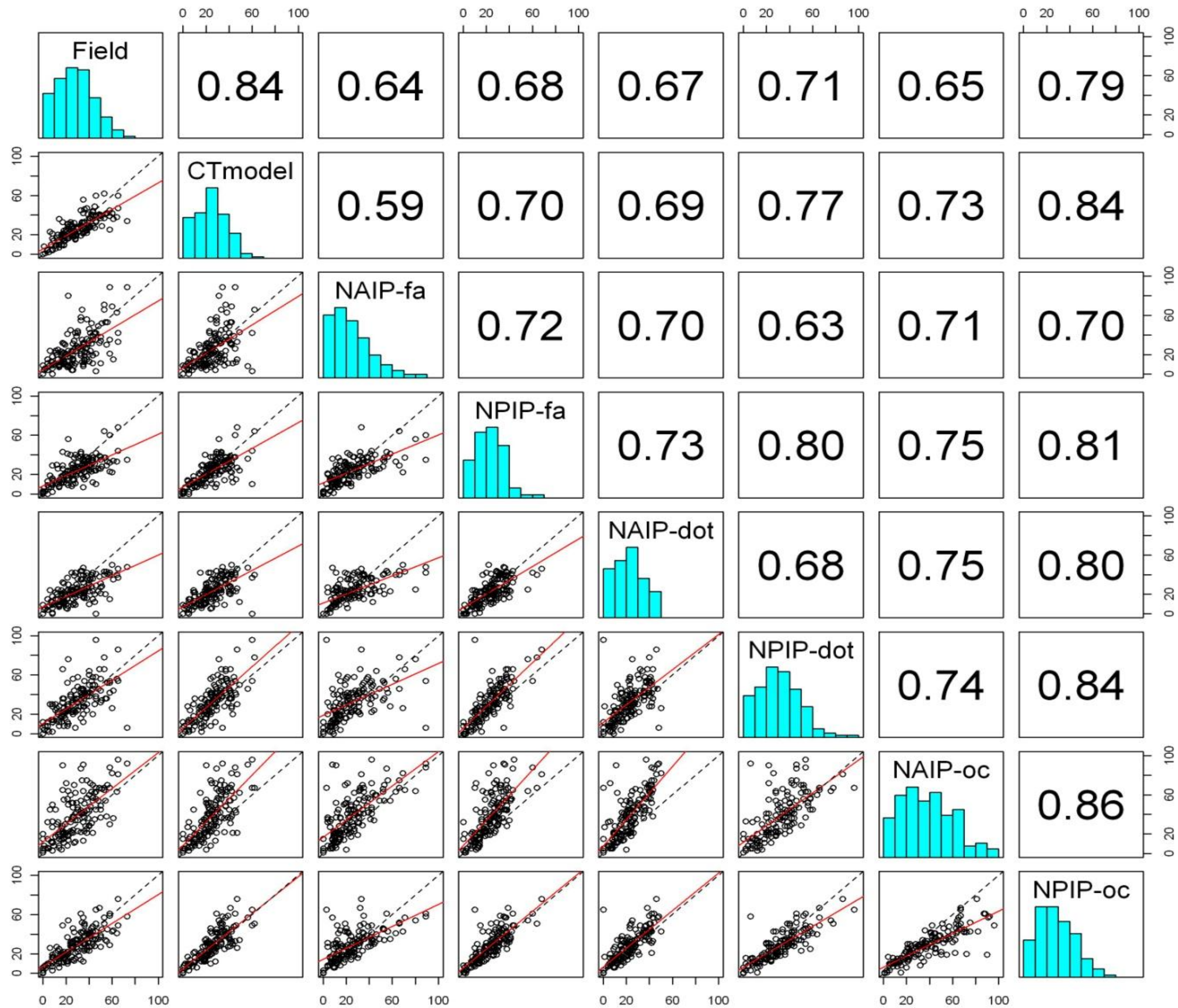
Compared ocular, image segmentation, and dot count assessments from digital aerial photography, as well as field-based measurements (Toney et al. In Press; Frescino and Moisen, In Press)

NAIP 6 in



Field transects, field crown models, segmentation, dot counts

Comparing 8 measures



Summary

Q1: RF and SGB are both powerful predictive modeling tools producing very similar results. RF has a number of practical advantages.

Q2: Larger modeling units, potentially at the E-W scale would be both logistically sensible and could provide reasonable model fits.

Q3: Empirical tree/no tree masks may not be worth the effort, but accurate water and other masks will be needed.

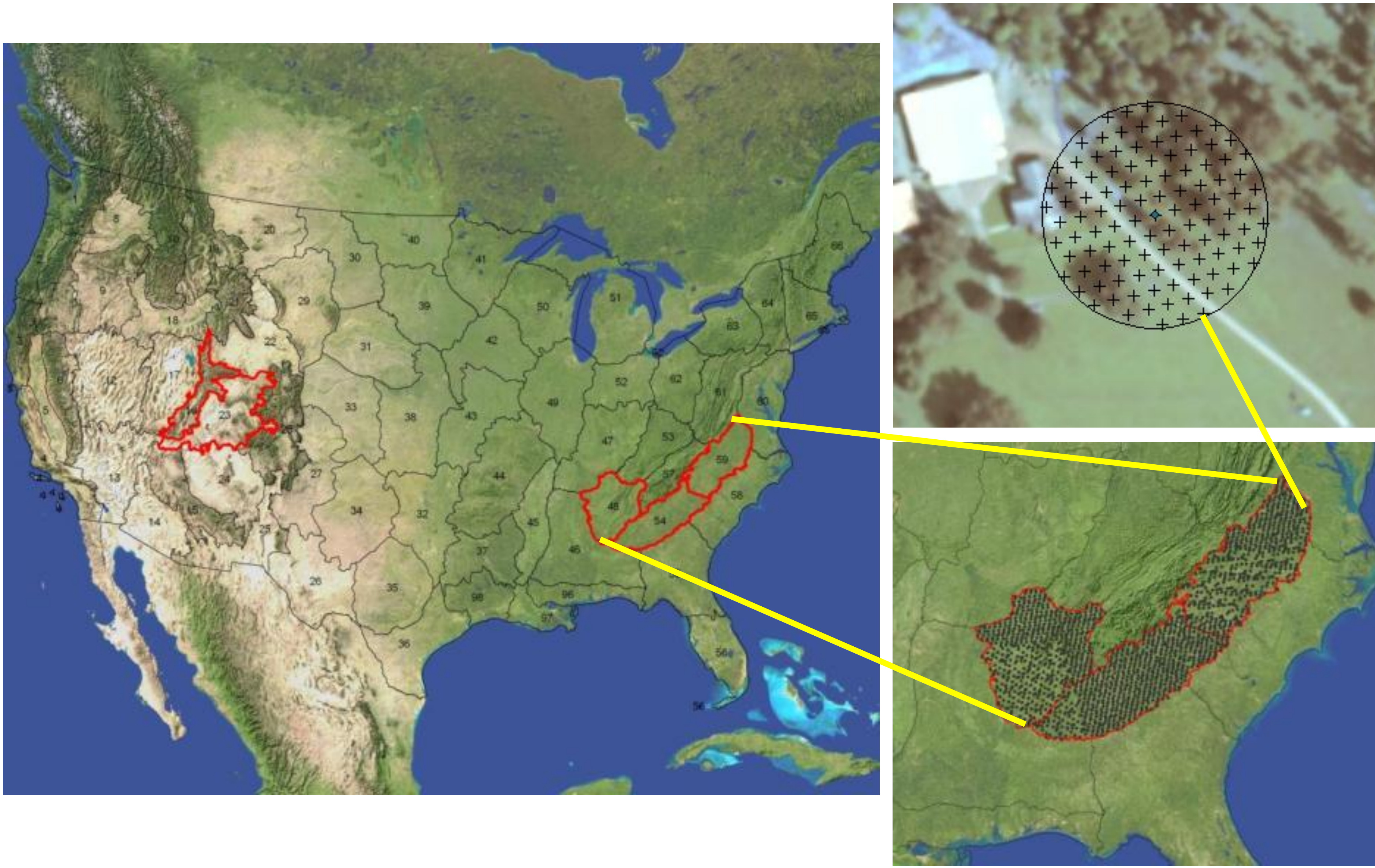
Q4: Photo sampling at the intensity of the FIA grid with less than 100 dots or less may be sufficient.

Q5: Tolerance/compliance rates can be derived to constrain adverse affects on models, but always need to bear in mind "truth?"

Pilot Phase - Key Questions

- Research on alternative pixel-level modelling techniques, alternative stratification/grouping strategies, using ordinal data for developing model, and model stability under different sampling intensification levels. *(Moisen et al.,Tipton et al., Coulston et al.)*
- Research on the impact of scale of observation on tree canopy cover estimates. Relationship among plot based, PI based, and modeled estimates (Toney et al. 2009) at multiple scales. *(Toney et al., Frescino et al., Gatziolis et al.)*
- Research on the impact of data normalization in the response variables. *(Tipton et al.)*
- Assessment and recommendations on photo interpretation repeatability *(Jackson et al.)*
- Research on modelling approaches for unique landscapes *(Sen et al.)*
- Synthesis *(Coulston et al.)*

Prototype Phase – Study Design



Prototype Phase – Key Questions

- How large an area can a single modeling unit encompass?
- How many samples (photo interpreted plots) are needed per modeling unit?
- Are normalized Landsat mosaic images required by the models or the maps?
- What is the minimum set of predictor layers needed by the models?

Timeline

2010	1Q	
	2Q	
	3Q	Pilot Complete
	4Q	
2011	1Q	
	2Q	Production Process Defined
	3Q	Production Begins
	4Q	
2012	1Q	
	2Q	
	3Q	
	4Q	
2013	1Q	
	2Q	
	3Q	
	4Q	CONUS Complete
2014	1Q	
	2Q	
	3Q	
	4Q	Coastal Alaska Complete
2015	1Q	
	2Q	
	3Q	
	4Q	HI, PR, VI Compelete

2010	Aug	Prototype Kickoff
	Sept	
	Oct	Pilot Complete
	Nov	
	Dec	
2011	Jan	
	Feb	Prototype PI data available
	Mar	
	Apr	Prototype ancillary data available
	May	
	Jun	Prototype analyses complete
	Jul	
	Aug	
	Sept	Production Begins

QUESTIONS ?