Validation of Geospatial Models Using Equivalence Tests

B. Tyler Wilson

NC Research Station

Numerous modeling efforts, both within and outside of the Forest Service, are underway to develop maps of forest attributes (e.g. area, volume, and growth) utilizing satellite imagery and other geospatial datasets. More rigorous statistical tools must be developed in order to evaluate these models and their resultant maps. This paper proposes a method for validating geospatial models and maps of forest attributes by using FIA plot data with equivalence tests. Unlike traditional significance testing for model validation, equivalence testing posits as the null hypothesis that the test statistics for the population of observations and predictions are different. With sufficient evidence the null hypothesis can be rejected and the model can be validated. This differs from traditional tests where a failure to reject the null hypothesis does not suggest that the model has been validated. The proposed methodology is applied to several geospatial models of forest area for illustration.

Validation of geospatial models using equivalence tests

B. Tyler Wilson, Mark H. Hansen, Ronald E. McRoberts
North Central Research Station, St. Paul, MN 55108
barrywilson@fs.fed.us

Evaluation of geospatial models

- FIA data used with numerous predictive models
- Not only “how much”, but “where”
- Geospatial data as input
  - Satellite imagery
  - Other raster data (e.g., DEM, climate)
  - Vector data (e.g. ecological regions, soils)
- Produce maps as output
- How good are resultant maps?

Primary reference

- Robinson, Andrew P., Remko A. Duursma, and John D. Marshall. “A regression-based equivalence test for model validation: shifting the burden of proof.” *Tree Physiology*, 25 (2005): 903-13.
- Derived from bioequivalence testing
- Used to compare efficacy of drugs

Hypothesis testing for models

- Compare two populations
  - observations
  - predictions
- Test statistic
  - e.g. mean difference between associated pairs

Traditional significance test

- Hypotheses
  - Null is that mean difference = 0
  - Alternative is that mean difference ≠ 0
- Specify α (region of rejection)
- Rejection of null hypothesis
  - acceptance of alternative hypothesis
- Failure to reject null hypothesis
  - not acceptance of null
  - simply lack of evidence
- Burden of proof misplaced for model validation

Equivalence test

- Hypotheses
  - Null is that mean difference ≠ 0
  - Alternative is that mean difference = 0
- Specify α and θ (region of equivalence)
- Rejection of null hypothesis
  - acceptance of alternative hypothesis
  - validates model
- Failure to reject null hypothesis
  - not acceptance of null
  - lack of evidence to validate model

Regression-based validation procedure

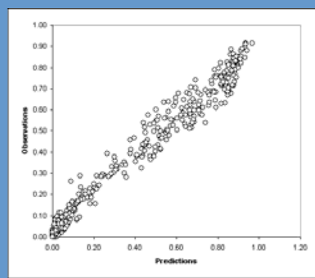
1. Tabulate observations and predictions
2. Subtract mean prediction from predictions
3. Define regions of equivalence
4. Fit linear regression
5. Test null hypotheses of dissimilarity

Interpreting the results

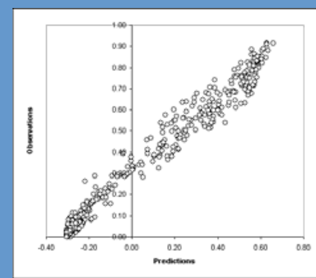
- Model validated if confidence interval for α within region of equivalence
- Separate tests for intercept and slope
- Alternative is report minimum θ where null hypothesis rejected

An illustration

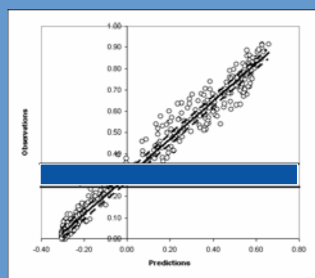
a.



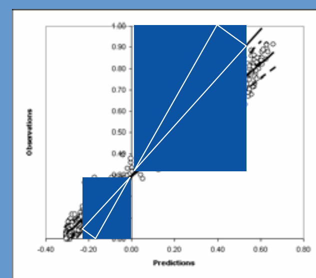
b.



c.



d.



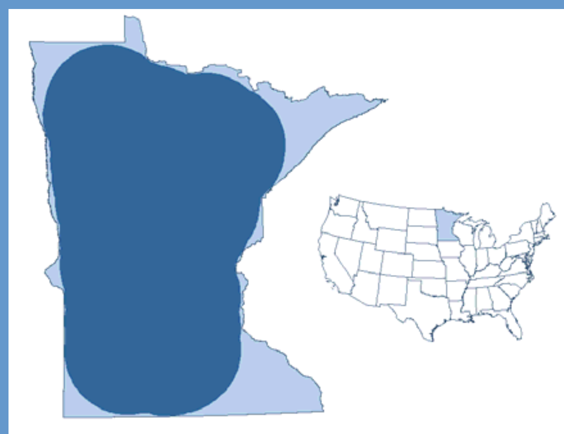
Modeling forest area

- Compare three geospatial models
- Produce prediction maps of forest area
- Observations estimated from FIA plots

Name	Technique	Imagery	Type
FIA-DT	Decision tree	250 m MODIS	Thematic
FIA-Logit	Logistic	250 m MODIS	Thematic
VCF	Regression tree	500 m MODIS	Continuous

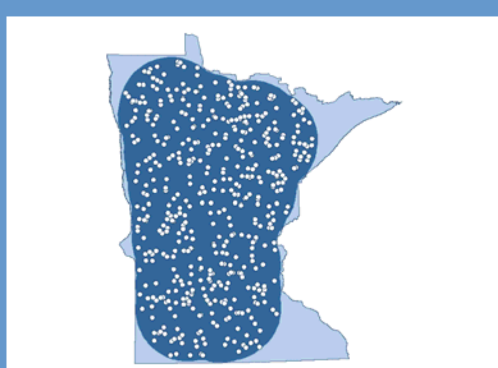
Evaluation area

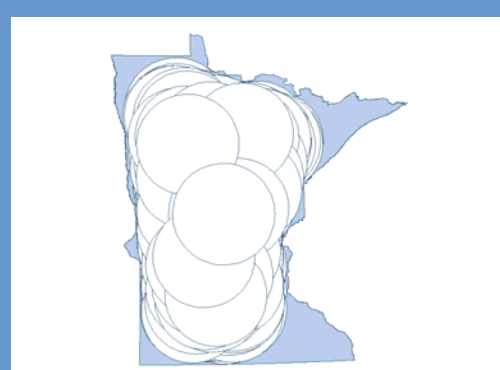
- Subset of Minnesota
- Easily applied to larger region



Circular estimation units

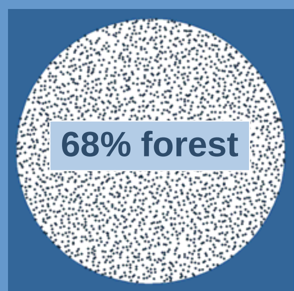
- Spatial mismatch between plots and pixels
- Use circular estimation units instead
- Random center points
- 500 circles each at radius of 5 - 100 km
- Each circle a data point in validation procedure





Two estimates for each circle

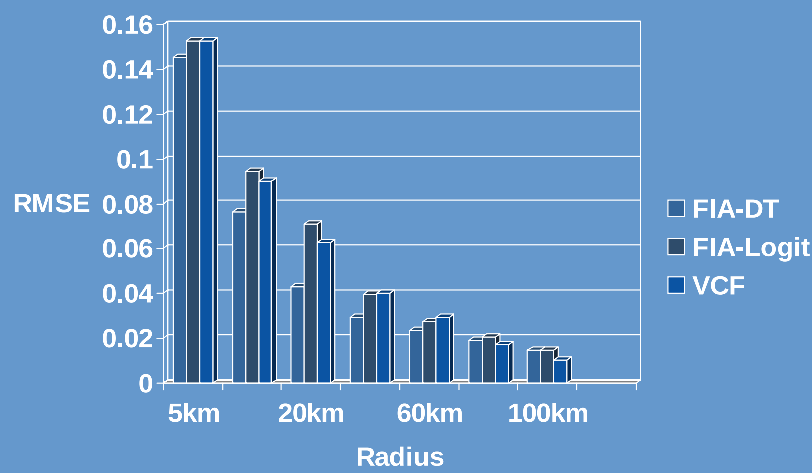
- Compute estimates from FIA plot observations and model predictions
- Model-based estimate average of all pixels within a circle
- Estimates are proportion forest



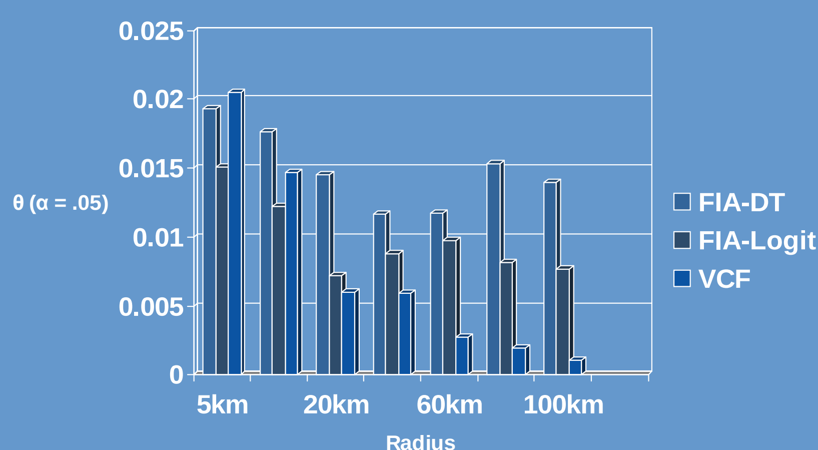
*Not true plot locations

Coefficient of determination

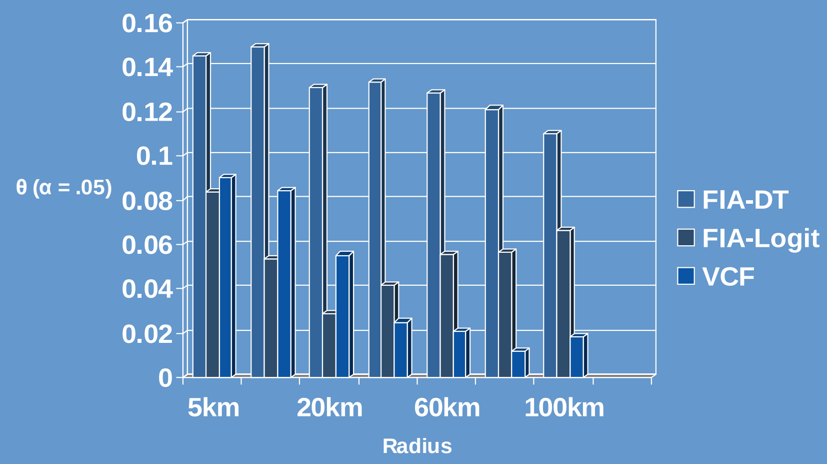
Standard error of regression



Equivalence test of intercept



Equivalence test of slope



Conclusions

- Readily applied at national scale
- Coarse or fine spatial resolution
- Relatively simple to implement