

# REPEATABILITY IN PHOTO-INTERPRETATION OF TREE CANOPY COVER AND ITS EFFECT ON PREDICTIVE MAPPING

Thomas A. Jackson, Gretchen G. Moisen, Paul L. Patterson, and John Tipton

## ABSTRACT

In this study, we explore repeatability in photo-interpreted imagery from the National Agriculture Imagery Program that was sampled as part of the National Land Cover Database 2011 Tree Canopy Cover pilot project. Data were collected in 5 diverse pilot areas in the US, including one each in Oregon, Utah, Kansas, Michigan and Georgia. Repeatability metrics. The intra-class correlation coefficient as well as repeatability graphics are explored for each pilot area. In addition, we conduct a simulation to illustrate the effect of varying tolerance and compliance rates on the predictive maps of tree canopy cover.

## INTRODUCTION

The National Land Cover Database (NLCD) includes a map the tree canopy cover of the entire United States. In support of developing a new 2011 Tree Canopy Cover product, a pilot test was recently launched to investigate numerous questions surrounding this map's development. The pilot used photo-interpreted data from large-scale aerial photography to give individual percent tree canopy estimates. Coupled with spatial predictor variables, these data were then used as the response variable in predictive models of tree canopy cover, addressing questions related to sampling intensity, value of specific predictor variables, modeling technique, and appropriate subpopulations (Tipton and others 2011, and Moisen and others 2011). The goal of this specific manuscript is to quantify the repeatability of the data and calculate the effects of fluctuation in individual interpreters on the mapping of the canopy cover.

## METHODS

### DATA

An intensive (approximately 1 km) grid of photo plots was established over five pilot areas the size of one Landsat scene, with one each in Georgia, Kansas, Michigan, Oregon, and Utah. On each photo plot, 105 dots were interpreted as landing on a tree or a non-tree using aerial photography acquired through the National Agriculture Imagery Program

(NAIP). To assess repeatability amongst photo interpreters, 208 plots were interpreted by every interpreter. This was a blind study in that the interpreters were unaware of which plots contributed to the repeatability data set. Percent tree canopy cover was obtained on each repeatability plot by averaging over the 105 dots per plot. This data was used to determine the repeatability of the photo interpretations. Also, for the five pilot areas, predictor variables were obtained including Landsat TM reflectance values and various GIS layers such as aspect, slope, digital elevation maps, and other measures important to tree growth. The response variable was the percent canopy cover discussed earlier. These predictors and response variables were used in modeling the effects of fluctuation in individual observers on the mapping of canopy cover.

### REPEATABILITY

Each of the five pilot areas contained between two and five observers to interpret the plots, with Georgia having four, Michigan having two, Utah having five, Kansas having five and Oregon having three. Repeatability among photo interpreters was first examined by constructing graphs of the paired canopy cover percentages for all pairs of photo interpreters within each pilot area. Agreement between photo interpreters was also explored by calculating the percent of observations for each pair of interpreters that fell within 10 percent of each other.

There are several different methods to test if data sets are different. A commonly used method for repeatability is the kappa function. Because the kappa function is primarily used for qualitative data and not continuous data as seen in this study, a new metric for the repeatability was investigated. The Interclass Correlation Coefficient (ICC, Equation 1) was first proposed by Pearson in 1901, but was only applicable for comparing two variables.

$$\rho_{ICC} = \frac{2 \sum_{i=1}^{k-1} \sum_{j=i+1}^k \sigma_{ij}}{(k-1) \sum_{i=1}^k \sigma_i^2 + \sum_{i=1}^{k-1} \sum_{j=i+1}^k (\mu_i - \mu_j)^2} \quad (1)$$

Thomas Jackson, Graduate Student, Colorado State University Department of Statistics, Fort Collins, CO, 80526  
Gretchen G. Moisen, Research Forester, US Forest Service Rocky Mountain Research Station, Ogden, UT, 84401  
Paul L. Patterson, Statistician, US Forest Service Rocky Mountain Research Station, Fort Collins, CO, 80526  
John Tipton, Graduate Student, Colorado State University Department of Statistics

Carrasco et al (2003) have expanded on his initial work to incorporate more than two variables, taking into account the mean, variance and covariance of each variable in question. The equation returns a score between 0 and 1, which can be interpreted as an  $R^2$  value; that is, the closer to 1, the better the repeatability between all observers. An ICC value was obtained for each of the five pilot areas.

### TOLERANCE AND COMPLIANCE RATES

Most measurement quality objectives (MQO) are stated in terms of a tolerance and a compliance rate (Pollard et al 2006, Westfall 2010). The tolerance limit is the amount of error we are willing to accept from each observer and is obtained based on the accuracy necessary for analytical purposes and the ability of the observer to repeat the findings. The compliance rate is the minimum percentage of observations that should be within tolerance. To calculate the effects of fluctuation in individual observers on the mapping of canopy cover, Random Forests (Breiman 2001) was employed to model tree canopy cover as a function of the predictor variables using data that were perturbed under varying tolerance and compliance rates. The changes in  $R^2$  values were then measured in terms of these perturbations. The predictors and response used in the Random Forests model were the same as those used in Tipton et al (2011). Initial  $R^2$  values were obtained for each of the five pilot areas using an independent test set using the definition of  $R^2$  of sum of squares regression divided by the sum of squares total. Then, 1,000 simulations were run at varying tolerances and varying compliance rates to investigate the effect on the mapping of the canopy cover. For each tolerance and compliance rate, an  $R^2$  value was obtained in the same manner as above. This was accomplished by first setting aside a portion of the original data, perturbing the remaining data based on the tolerance and compliance rates, and obtaining the predictor coefficients for the canopy cover using the Random Forests model on the perturbed data. Once this was accomplished, the portion of data set aside for validation was used to construct the  $R^2$  value.

## RESULTS AND DISCUSSION

### REPEATABILITY

Figure 1 illustrates the paired canopy cover percentages for all pairs of photo interpreters from the Kansas pilot area and is representative of the results from Michigan, Oregon, and Utah. We can see that the grouping of the points tends to lie on or closely around the one-to-one line ( $y$  equals  $x$  line), which indicates that there is strong agreement between observers. Another way to look at this, Table 1, illustrates results from the Oregon pilot area and shows that all three interpreters have at least 70 percent of their observations within 10 percent of each other, with better agreement between some observers than others. Figure 2, however, illustrates the lack of a pattern seen in some of the paired

canopy cover percentages for all pairs of photo interpreters from Georgia. Table 2, showing the percent of observations for each pair of interpreters that fell within 10 percent of each in Georgia, further illustrates this. Here we see that three of the observers (observers 1, 3 and 4) are in strong agreement with each other, with observer 2 not being in strong agreement with any of the other three. We can see that observer 2 does not have more than 50 percent of its observations within 10 percent of each other. Observer 2 tends to be the outlier in the Georgia pilot area, with the other three observers being in strong agreement with each other not only in the paired canopy cover plots but within percentages as well. For the other four pilot areas, the proportions are always above 63 percent, with the majority of the calls being above 70 percent, which indicates that the agreement between observers is very high. An ICC value was obtained for each of the five pilot areas and can be seen in Table 3. As discussed above, the agreement of the observers from the Georgia pilot area was shown not to be as strong as seen in the other four pilot areas. This is again seen in the low ICC value from the Georgia pilot area, but since observer 2 is already known to be the outlier from the Georgia pilot area, it can be treated as an outlier and removed from the dataset. With observer 2 removed, the ICC value from Georgia increases from 0.6491 to 0.9830. With the high ICC values returned, this not only echoes the findings from the one-to-one plots and the within percentages that there is strong repeatability across observers, but also gives us a way to quantify the strength of the repeatability.

### TOLERANCE AND COMPLIANCE RATES

The effect of the fluctuation was measured using the ratio of the perturbed  $R^2$  value over the original  $R^2$  value. Table 4 shows results for the Georgia pilot area. The tolerances are on the left hand side and indicated as decimals, and the compliance rates are across the bottom and indicated as percentages. The ratios from the other four pilot areas were all within 2 percent of each other. We can see from this table that to develop models that explain at least 80 percent of the variability that the model from unperturbed data provides, we have to have a minimum tolerance limit of 20 percent and a minimum compliance rate of 85 percent.

## CONCLUSIONS

With the inclusion of several different metrics for all five pilot areas, we were able to conclude that, with proper training and preprocessing, the repeatability of the study areas is high. Not only are the one-to-one plots and within percentages strong indicators of repeatability, but high ICC scores are able to quantify this fact. The repeated simulations for all pilot areas show that as long as strict guild lines are upheld, large proportions of the true predictive strength can be returned.

## LITERATURE CITED

**Breiman, L.** 2001. Random Forests. Machine learning. 45:5-32.

**Moisen, G.G.;** Coulston, J.W.; Wilson, B.T. [and others]. 2011. Choosing appropriate subpopulations for modeling tree canopy cover nationwide. In McWilliams, W., Moisen, G.G.; Roesch, F. (eds.). Forest Inventory and Analysis (FIA) Symposium 2010; October 5-8, 2010; Knoxville, TN. Proc. SRS-CD. Knoxville, TN: U.S. Department of Agriculture Forest Service, Southern Research Station. 1 CD.

**Carrasco, J L.;** Jover, J L I. 2003. The concordance correlation coefficient estimated through variance components. IX Conferencia Espanola de Biometria.

**Pollard, J.E.;** Westfall, J.A.; Patterson, P.L. [and others]. 2006. Forest Inventory and Analysis national data quality assessment report for 2000 to 2003. RMRS-GTR-181. U.S. Department of Agriculture Forest Service, Rocky Mountain Research Station.

**Tipton, J.;** Moisen, G.G.; Patterson, P.L. [and others]. 2011. Sampling intensity and normalizations: exploring cost-driving factors in nationwide mapping of tree canopy cover. In McWilliams, W., Moisen, G.G.; Roesch, F. (eds.). Forest Inventory and Analysis (FIA) Symposium 2010; October 5-8, 2010; Knoxville, TN. Proc. SRS-CD. Knoxville, TN: U.S. Department of Agriculture Forest Service, Southern Research Station. 1 CD.

**Westfall, J. A.** ed. 2010. FIA national assessment of data quality for forest health indicators. Gen. Tech. Rep. NRS-53. Newtown Square, PA: U.S. Department of Agriculture Forest Service, Northeastern Research Service. p.80

**Table 1—Within 10 percent proportions for Oregon Pilot Area**

|       | Obs 1 | Obs 2  | Obs 3  |
|-------|-------|--------|--------|
| Obs1  | 1.00  | 0.7588 | 0.7075 |
| Obs 2 |       | 1.00   | 0.8103 |
| Obs 3 |       |        | 1.00   |

**Table 3—Interclass correlation coefficient values for all five plot areas**

| Pilot Area | ICC Value |
|------------|-----------|
| Georgia    | 0.6491    |
| Kansas     | 0.9126    |
| Michigan   | 0.9424    |
| Oregon     | 0.9299    |
| Utah       | 0.8652    |

**Table 2—Within 10 percent proportions for Georgia Plot Area**

|       | Obs 1 | Obs 2  | Obs 3  | Obs 4  |
|-------|-------|--------|--------|--------|
| Obs 1 | 1.00  | 0.4615 | 0.9279 | 0.9087 |
| Obs 2 |       | 1.00   | 0.4327 | 0.4471 |
| Obs 3 |       |        | 1.00   | 0.9135 |
| Obs 4 |       |        |        | 1.00   |

**Table 4—R<sup>2</sup> ratios for Georgia Pilot Area tolerance and compliance limits**

| .3 | .7771 | .7530 | .7235 | .7004 | .6779 |
|----|-------|-------|-------|-------|-------|
| .2 | .8675 | .8468 | .8277 | .8101 | .7893 |
| .1 | .9472 | .9339 | .9195 | .9075 | .8933 |
|    | 100%  | 95%   | 90%   | 85%   | 80%   |

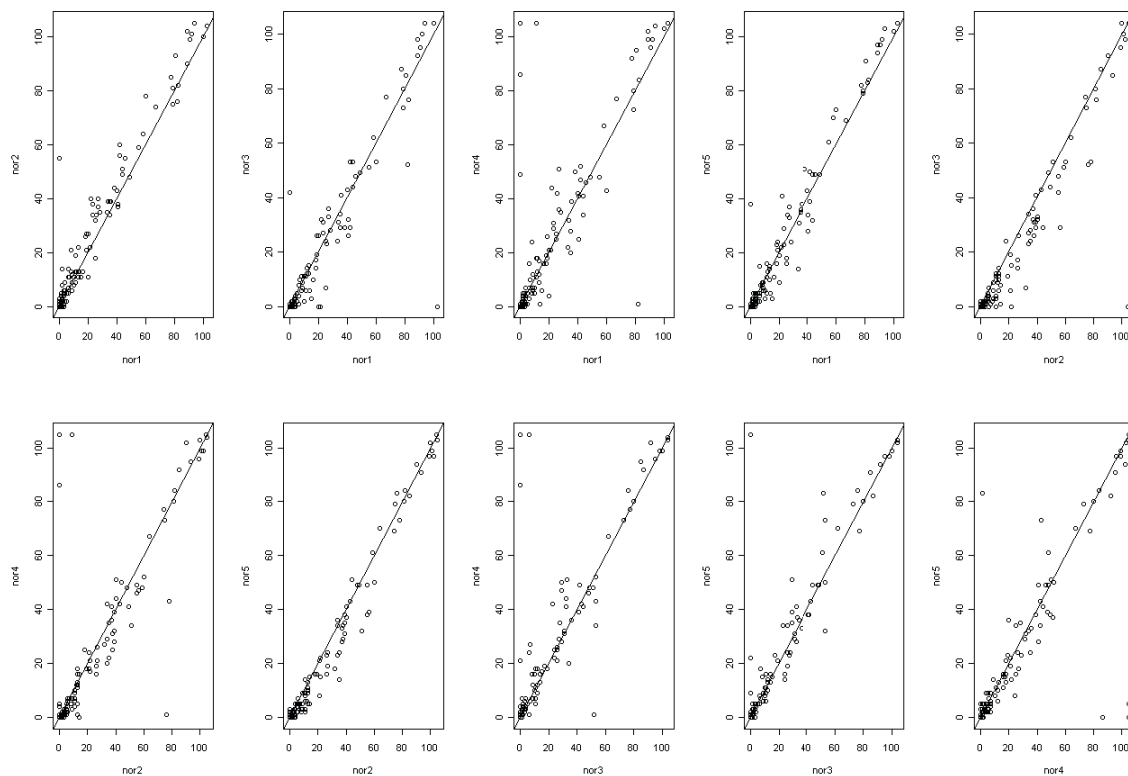


Figure 1—Paired Canopy Cover Plots for all Interpreters in the Kasas Pilot Area Including One-to-One Line

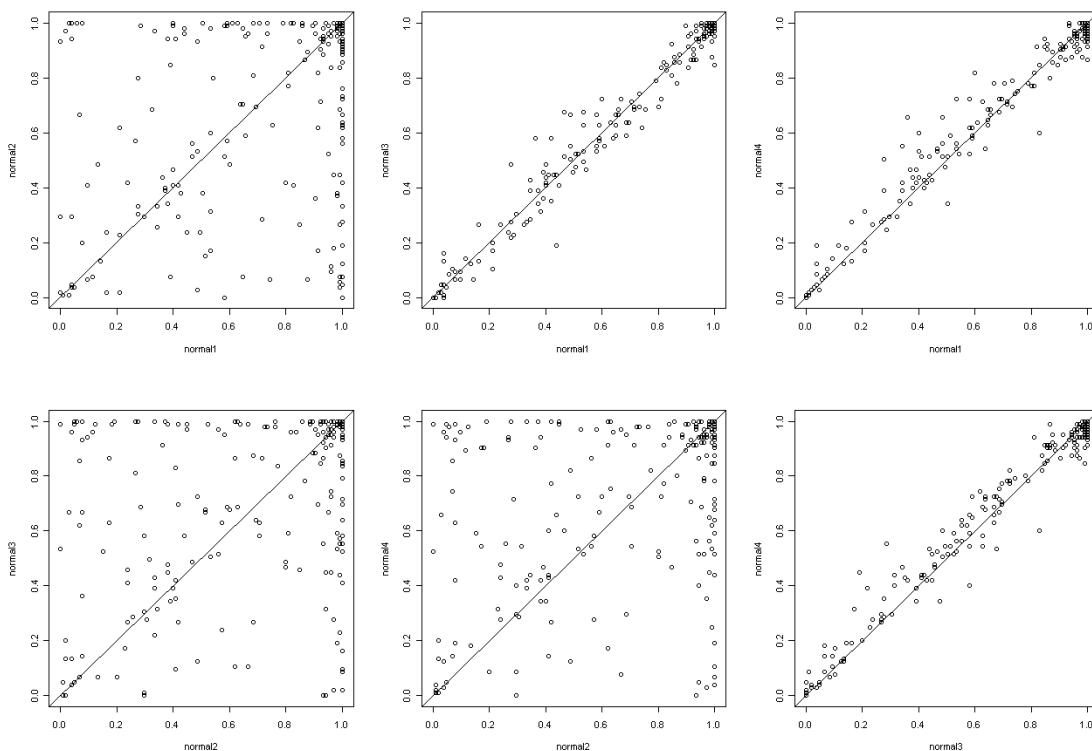


Figure 2—Paired Canopy Cover Plots for all Interpreters in the Georgia Pilot Area Including One-to-One Line