Plot size, shape, and co-registration error determine expected overlap

J.W. Flewelling

1 Seattle Biometrics & Analysis, LLC, 9320 40th Ave. NE, Seattle, WA, USA, 98115-3715; 206-268-0728; flewelling@seattlebiometrics.com

Introduction

Forest inventories which rely upon remotely sensed data and field plot data often have a need to co-register field plot locations with the remotely sensed imagery. For some remotely sensed data sources, precise co-registration may be almost impossible. Cook (2000) notes that the positional error of Landsat pixels may not be better than 1 pixel (28.5 m), and that GPS errors for Forest Inventory and Analysis (FIA) plots often exceed 5 m. The effect of registration error upon an estimator for forest area in a simulation study with Landsat data was to increase variances by factors ranging up to five (Patterson and Williams 2003). In contrast, lidar data often has sub-meter accuracy, and has been used with field plots where the GPS accuracy is about 0.2 m (Næsset 2004). Tree and crown matching procedures can sometimes improve the co-registration; Dorigo et al. (2008) successfully applied such procedures to angle-count plots.

Lidar data can be summarized into pixels of any size and shape for area-based analyses. Area-based statistics summarize the vertical distribution of lidar returns over fixed-area pixels. The size and shape of ground plots can be chosen to exactly match the size and shape of the pixels. For area-based biomass estimators, such as those developed by Næsset (2004), the sampling principles are almost textbook perfect: the sample units are the lidar pixels, and the ground plots are coincident with the pixels subject to positional error that is negligible.

Size and shape of plots does have an effect upon the efficiency of an inventory. Experience with past inventories may suggest which sample designs and plot configurations have lead to acceptable variance. Alternatively, an understanding of marginal efficiencies for selected design factors may help guide the design process; this article provides some insight relevant to the design factors of plot shape and size in the presence of co-registration error.

One measure of the effectiveness of the sampling protocol is overlap between the ground plots and the corresponding plots or pixels in the remotely sensed data. If the distribution of the co-registration errors is known or assumed, the expected overlap can be computed. Results are shown for three matched equal-area plot shapes: circle, square and hexagon, and a hexagonal plot paired with an equal-area circular plot.

Methods

Consider a circular plot of radius $R$ superimposed on the remotely sensed data and a circular ground plot with the same radius, their centers being separated from one another by a distance $d$. That distance is the assumed absolute co-registration error. If $d$ is zero, the fractional overlap ($f$) is 1.0. If $d \geq 2 \times R$, the fractional overlap is zero. For $0 < d < 2 \times R$, the overlap is:

$$f(d) = \left[ \varphi - \sin(\varphi) \right] / \pi \quad \text{where } \varphi = 2 \times \arccos\left( \frac{d}{2 \times R} \right)$$ (1)

Extending Forest Inventory over Space and Time. May 19-22, 2009, Quebec City, Canada
and \(\text{arcos}\) is the inverse cosine function.

Next consider a square plot with sides of length \(L\) superimposed on the remotely sensed data, a square ground plot of the same size and orientation, the offset between the plots being a distance \(d\) at an angle \(\theta\) (Figure 1). If the offset distance is zero, the overlap is 1. For \(\theta\) restricted to the range \(0 < \theta \leq \pi/4\):

\[
f(d, \theta) = \left[ L - d \times \cos(\theta) \right] \times \left[ L - d \times \sin(\theta) \right] / L^2 \quad \text{if} \quad d \leq L \quad \text{or} \quad \theta > \text{arcos}(L/d) \quad (2)
\]

and zero otherwise.

![Figure 1. Square remotely sensed plot and corresponding ground plot showing co-registration error \(d\) at angle \(\theta\). The cross-hatched portion is the overlap of the two plots.](image)

The expectation of the foregoing function over a uniformly distributed \(\theta\) in the range \([0, \pi/4]\) is the integral of the function divided by \(\pi/4\). Expected overlaps are zero for \(d/L > \sqrt{2}\), and

\[
f(d) = 1 + (1 / \pi) \times (d/L) \times \left[ -4 + (d/L) \right] \quad \text{if} \quad 0 \leq d/L \leq 1
\]

\[
= 1 + (1 / \pi) \times \left\{ -(d/L)^2 + 4 [(d/L)^2 -1]^{0.5} -4 \times \text{arcos}(L/d) -2 \right\} \quad \text{if} \quad 1 < d/L < \sqrt{2} \quad (3)
\]

Now consider two identically sized and oriented hexagonal plots, with side length \(H\). Offset distance is \(d\) at angle \(\theta\). Due to symmetry, the only angles which need to be considered are from zero to \(\pi/6\). For any angle beyond this range, an equivalent within-range angle is the minimum of \([\text{mod}(\theta, \pi/3), \pi/3 - \text{mod}(\theta, \pi/3)]\). The area of a hexagonal shaped plot is \((3/2) \times (\sqrt{3}) \times H^2\). The apothem \((S)\) has length \((\sqrt{3}) \times (H/2)\). The horizontal and vertical offsets are:

\[
x = d \times \cos(\theta) \quad (5)
\]

\[
y = d \times \sin(\theta) \quad (6)
\]
For moderate offsets, including all with \( d < H \), the fractional overlap is:

\[
f(d, \theta) = \left\{ 3 \times H^2 + (\sqrt{3}) \times \left[ -y^2 + (4S - 2y) \times (H - x) \right] \right\} / \left[ 9 \times H^2 \right] \tag{7}
\]

As an alternative to the use of formulas, a square grid can be established with sufficient extent to encompass both the original plot and the offset plot. The distance between grid lines is set to \( H/400 \). Each grid point is evaluated as to whether or not it was within both the original plot and the offset plot. The dot count method is used to estimate the overlap area. Results are averaged over \( \theta \), uniformly distributed over the range \([0, \pi/6]\). This numeric procedure is used for the overlap of hexagons, and for the overlap of equal-area pairs of circles and hexagons. In the results section, the overlap functions, \( f(d) \), are evaluated for a range of offsets. In a subsequent step, the offsets distances are assumed to be the absolute values of normally distributed random variables with mean zero and specified variance; the expectations of the overlaps are computed. Numeric integration over the normal distribution is used to calculate expected plot overlaps.

**Results**

Expected overlaps for specified offsets are calculated using Eqn. 1 for circular plots, Eqn. 3 and 4 for square plots, and numerical integration for hexagonal plots and for circles paired with hexagons. The independent variable is the ratio of the offset \( d \) and the square root of plot area. Table 1 has selected results; Figure 1 compares the results for paired circular plots with those for a circle and hexagon pair. The effect of a random co-registration error is reported in Table 2.

![Figure 2](image-url).

**Figure 2.** Overlap fraction \( f \) as a function of the co-registration error \( d \) and plot area. The solid line is for a pair of equal-area circular plots; the dashed line is for a circular plot paired with a hexagonal plot of the same area.
Table 1. Expected plot overlap as a function of plot shape, fixed co-registration error \((d)\), and plot area.

<table>
<thead>
<tr>
<th>(d/\sqrt{\text{Area}})</th>
<th>CIRCLE</th>
<th>SQUARE</th>
<th>HEXAGON</th>
<th>CIRCLE - HEXAGON</th>
</tr>
</thead>
<tbody>
<tr>
<td>0.0</td>
<td>1.0000</td>
<td>1.0000</td>
<td>1.0000</td>
<td>0.9657</td>
</tr>
<tr>
<td>0.1</td>
<td>0.8873</td>
<td>0.8759</td>
<td>0.8838</td>
<td>0.8839</td>
</tr>
<tr>
<td>0.2</td>
<td>0.7755</td>
<td>0.7581</td>
<td>0.7709</td>
<td>0.7738</td>
</tr>
<tr>
<td>0.4</td>
<td>0.5583</td>
<td>0.5416</td>
<td>0.5565</td>
<td>0.5575</td>
</tr>
<tr>
<td>0.6</td>
<td>0.3564</td>
<td>0.3506</td>
<td>0.3572</td>
<td>0.3561</td>
</tr>
<tr>
<td>0.8</td>
<td>0.1800</td>
<td>0.1851</td>
<td>0.1791</td>
<td>0.1800</td>
</tr>
<tr>
<td>1.0</td>
<td>0.0453</td>
<td>0.0451</td>
<td>0.0472</td>
<td>0.0455</td>
</tr>
</tbody>
</table>

Table 2. Expected plot overlap as a function of plot shape, plot area, and standard deviation (\(std\)) of the co-registration error \((d)\), the latter assumed to be from a normal distribution with mean zero.

<table>
<thead>
<tr>
<th>(\text{std}(d)/\sqrt{\text{Area}})</th>
<th>CIRCLE</th>
<th>SQUARE</th>
<th>HEXAGON</th>
<th>CIRCLE - HEXAGON</th>
</tr>
</thead>
<tbody>
<tr>
<td>0.1</td>
<td>0.9104</td>
<td>0.9018</td>
<td>0.9079</td>
<td>0.9012</td>
</tr>
<tr>
<td>0.2</td>
<td>0.8223</td>
<td>0.8099</td>
<td>0.8191</td>
<td>0.8167</td>
</tr>
<tr>
<td>0.3</td>
<td>0.7370</td>
<td>0.7243</td>
<td>0.7342</td>
<td>0.7330</td>
</tr>
<tr>
<td>0.4</td>
<td>0.6567</td>
<td>0.6451</td>
<td>0.6544</td>
<td>0.6535</td>
</tr>
</tbody>
</table>

Discussion

Plot shape has almost no effect upon overlap; there are noticeable differences between shapes for very small co-registration errors. Gobakken and Næsset (2009) show an effect of co-registration error upon the variance of lidar-based estimators for yield. Parker and Evans (2007) took care to ensure that plot sizes and shapes for regression development exactly matched those for application. If lidar data are to be used for individual tree crown (ITC) recognition, sampling schemes may limit the utility of tree data near the edge of the field plots (Flewelling 2009); hence circular or hexagonal plots may be more efficient than square plots for ITC analyses.

Lewis and Hutchinson (2000) provide models for the effect of spatial error upon prediction accuracy under a paradigm of point sampling a surface whose spatial structure is modeled with a semivariogram. Their approach provides a general structure that would seem to be particularly adept at dealing with positional errors which are larger than the dimensions of a field plot. For positional errors of smaller magnitude, the simple concept of expected plot overlap may provide some guidance in choosing plot size and shape.

Extending Forest Inventory over Space and Time. May 19-22, 2009, Quebec City, Canada
Conclusion

Plot shape has little influence upon expected overlap of field plots and crown plots from remote imagery. Hence for area-based analyses, concerns for registration error need not influence choice of plot shape. However, if ITC analyses are planned, circular or hexagonal ground plots may be more efficient than square plots. Sampling schemes which tessellate the remotely sensed data are attractive for how well they fit into sampling theory; such schemes are easily implemented with square or hexagonal plots. Results here indicate that hexagonal crown plots could be combined with circular ground plots with little reduction in expected plot overlap.

The effects of compass error on plot layout for square or hexagonal plots have not been examined. No attempt been made to combine these geometry-based overlap results with any of the methods for characterizing spatial variance.

Literature Cited


Patterson, P. L. and M.S. Williams. 2003. Effects of registration errors between remotely sensed and ground data on estimators of forest area. For. Sci. 49(1): 110-118.