

# Agreement Between Measurements of Shrub Cover Using Ground-Based Methods and Very Large Scale Aerial Imagery

Corey A. Moffet

Author is Research Rangeland Scientist, US Department of Agriculture–Agricultural Research Service, US Sheep Experiment Station, Dubois, ID 83423, USA.

## Abstract

Very large scale aerial (VLSA) photography is a remote sensing method, which is collected and analyzed more efficiently than ground-based measurement methods, but agreement with ground-based measurements needs to be quantified. In this study, agreement between ground- and image-measured cover and precision, and accuracy of image locations and scale, were assessed. True image locations were determined by georeferencing images and conducting a ground search. Accuracy and precision of planned, aircraft, and georeferenced locations were evaluated by comparison with true image locations. Shrub cover was measured at true image locations using ground-based line-intercept and on the image using point-intercept. Sagebrush (*Artemisia* spp. L.), antelope bitterbrush (*Purshia tridentata* [Pursh] DC.), and spineless horsebrush (*Tetradymia canescens* DC.) were distinguished in the imagery. Agreement between ground- and image-based measurements was quantified using limit-of-agreement analysis. True ground locations of the VLSA images were within a 41-m radius of the aircraft location at the time of image acquisition, with 95% confidence. Using a panchromatic image from the QuickBird satellite (0.6-m pixel resolution) as a base map, 90% of true ground locations were within a 5-m radius of the location estimated from georeferencing the VLSA image to the base map. VLSA image-measured cover was, in general, unbiased with mean absolute differences between VLSA- and ground-based methods less than 1.3%. The degree of agreement and absence of bias between VLSA image-measured and ground-measured cover is sufficient to recommend using VLSA imagery to measure shrub cover.

## Resumen

Las imágenes aéreas a larga escala (VLSA) es una metodología de sensores remotos que se recolecta y analiza más eficientemente que las mediciones basadas en métodos terrestres, pero al igual, necesitan cuantificarse con las mediciones terrestres. En este estudio, la precisión, exactitud, localización y escala de la cobertura, medida con imágenes, se evaluó comparando con medidas de cobertura a nivel del suelo. La verdadera localización de la imagen se determinó con imágenes georeferenciadas y mediante búsquedas dirigidas en tierra. La exactitud y precisión de localizaciones aéreas y georeferencias se evaluaron por comparación con verdaderas localizaciones de imagen. La cobertura de arbustos se midió con localización de imágenes verdaderas, usando mediciones terrestres basadas en el método de intercepción de línea y con imágenes utilizando el punto de intercepción. Sagebrush (*Artemisia spp.*, L.), antilope bitterbrush (*Prussia tridentata* [Pursh] DC.), y horsebrush sin espinas (*Tetradymia canescens* DC.) se identificaron en las imágenes, la concordancia entre medidas terrestres y de imagen aérea se cuantificaron utilizando el análisis de limite de concordancia. Las locaciones terrestres de las imágenes aéreas a larga escala (VLSA) estuvieron entre un radio de 41 metros de la localidad aérea al tiempo de adquisición de la imagen, con un 95% de confianza. Utilizando una imagen panorámica del satélite QuickBird (a una resolución de 0.6 mega píxeles) 90% de las localidades verdaderas estuvieron dentro de un radio de 5 metros de las ubicaciones gereferenciando la imagen de VLSA en el mapa base. Las mediciones de cobertura de las imágenes VLSA fueron en general no tendenciosas con una diferencia absoluta promedio entre las imágenes de VLSA y los métodos terrestres de 1.3%. El grado de concordancia y la ausencia de tendencia entre las imágenes VLSA y las medidas terrestres de cobertura son satisfactorios para recomendar el uso de las imágenes VLSA para medir cobertura de arbustos.

**Key Words:** antelope bitterbrush, georeferencing, limit-of-agreement analysis, mountain big sagebrush, spineless horsebrush

## INTRODUCTION

For decades, managers have lacked statistically adequate data to properly guide rangeland management because of the high cost

---

The use of trade, firm, or corporation names in this publication is for the information and convenience of the reader. Such use does not constitute an official endorsement or approval by the US Department of Agriculture or the Agricultural Research Service of any product or service to the exclusion of others that may be suitable.

Correspondence: Corey Moffet, US Dept of Agriculture–Agricultural Research Service, US Sheep Experiment Station, 19 Office Loop, Dubois, ID 83423, USA. Email: cmoffet@pw.ars.usda.gov

Manuscript received 12 December 2008; manuscript accepted 25 March 2009.

of conventional data collection (West 1999). Very large scale aerial (VLSA) imagery has been used for experimentally measuring attributes such as invasive weed cover, vegetation cover, and bare ground (Seefeldt and Booth 2006, Blumenthal et al. 2007, Booth and Cox 2008). VLSA imagery has the advantage that numerous samples can be collected over large land areas in a short period of time at less cost than conventional methods. The cost of obtaining VLSA imagery-based vegetation cover measurements in a landscape-scale sampling was significantly less than even ocular measurements due, in large part, to travel costs between sample locations (Seefeldt and Booth 2006).

Before VLSA imagery can be recommended as a replacement for ground-based cover measurements, agreement between ground- and image-based cover measurements must be quantified. Agreement must be assessed with paired measurements made on the same experimental units, so that the only source of variation is the method. To date, this type of cover validation study for VLSA imagery has not been reported. Other authors have compared ground- and image-based measurement methods (Seefeldt and Booth 2006, Booth et al 2008). Seefeldt and Booth (2006) acknowledge that accuracy was not assessed, but compared four methods of measuring vegetation cover using 14 different management units as experimental units. They found, for example, that in nearly half the comparisons between laser point-frame and VLSA, the methods gave different results. Booth et al. (2008) compared ground- and VLSA-measured vegetation cover, among other characteristics, and reported that ground-measured vegetation cover and measures taken using laser point-frame, and point-intercept were poorly correlated with the VLSA-measured vegetation cover. The evaluation of accuracy and precision for cover measurements requires that the true image locations be determined. Because of inaccuracies of locations determined from a single global positioning system (GPS) reading and roll, pitch, and yaw of the aircraft, there are significant differences between the true image location on the ground and the location of the aircraft when the image was acquired.

Therefore, the objectives were to 1) characterize the accuracy and precision of aircraft-based positions and imagery ground-sample distance (GSD), 2) determine whether the precision of an image location was improved by using a proposed method for georeferencing VLSA images, 3) characterize agreement between ground- and VLSA image-based cover measurements, and 4) determine the minimum sample size required with different sampling schemes for assessing whether or not the image-based method is biased.

## MATERIALS AND METHODS

### Study Area

The study area was on the headquarters property of the US Sheep Experiment Station (USSES) near Dubois, Idaho (lat 44°17'N, long 112°7'W, approximately 1800 m elevation). Median annual precipitation was 319 mm, and mean January and July temperatures were -7°C and 20°C, respectively (National Oceanic and Atmospheric Administration 2003). The soils were all complexes of the Akbash (fine-loamy, mixed, superactive, frigid Calcic Pachic Haploxerolls), Meremba (fine-loamy, mixed, superactive, frigid Calcic Pachic Haploxerolls), and Pyrenees (loamy-skeletal, mixed, superactive, frigid Typic Calcixerolls) soils (US Department of Agriculture-Natural Resources Conservation Service, unpublished data, 1995).

Vegetation on the study site was a sagebrush-grass community dominated by mountain big sagebrush (*Artemisia tridentata* Nutt. subsp. *vaseyana* [Rydb.] Beetle). Subdominant shrub species included antelope bitterbrush (*Purshia tridentata* [Pursh] DC.), yellow rabbitbrush (*Chrysothamnus viscidiflorus* [Hook.] Nutt.), threetip sagebrush (*Artemisia tripartita* Rydb.), and spineless horsebrush (*Tetradymia canescens* DC.). Dominant grass and grass-like species present included Sandberg

bluegrass (*Poa secunda* J. Presl), bluebunch wheatgrass (*Pseudoroegneria spicata* [Pursh] A. Löve), sedge (*Carex* spp. L.), and Idaho fescue (*Festuca idahoensis* Elmer). Dominant forbs were parsnipflower buckwheat (*Eriogonum heracleoides* Nutt.), northwestern Indian paintbrush (*Castilleja angustifolia* [Nutt.] G. Don), longleaf fleabane (*Erigeron corymbosus* Nutt.), and littleleaf pussytoes (*Antennaria microphylla* Rydb.).

### Imagery

Color VLSA photographs were collected across the USSES headquarters property on the mornings of 16 and 17 June 2006, at an altitude of approximately 250 m above ground level (AGL) at a 23.2 m · s<sup>-1</sup> average ground speed. The aircraft was fitted with two digital cameras that could be triggered simultaneously. A Canon EOS 1Ds 11.1-megapixel (MP; 4064 × 2704 pixels) camera (Canon USA, Lake Success, NY) was configured with a 100-mm focal-length lens to acquire images with a large GSD (24 mm) and field of view (FOV; 0.62 ha) for georeferencing. The other camera, a 16.7-MP Canon EOS 1Ds Mark II (Canon USA, 4992 × 3328 pixels) was configured with an 840-mm focal-length lens that captured images with small GSD (2.3 mm) and small FOV (88 m<sup>2</sup>). Shutter speeds were manually set on each camera for 1/4000th of a second with safety shift enabled to allow slower shutter speeds when light was inadequate. During the total 5.2-h flight time, images were acquired with both cameras at 1134 locations across the USSES headquarters property. Of the 1134 locations where imagery was acquired, 457 were within the mountain big sagebrush community type. Seven areas, all in mountain big sagebrush communities, that differed in number of years since last burn were defined and three images within each of these seven areas were selected at random, for a total of 21 images used for validation of shrub cover. All shrub cover measurements were made on the 16-MP photographs.

The camera triggering system was electronically linked with the navigation system. The navigation system was powered by Tracker software (Track'Air B.V., Oldenzaal, The Netherlands). A preflight plan was made that defined flight paths 300 m apart and waypoints located 300 m apart along each path. Because the pilot could not consistently maneuver the aircraft directly over the preplanned point, cameras were simultaneously triggered when the aircraft met the following two conditions: 1) laterally within ± 100 m of the preplanned path and 2) on the line passing through the preplanned waypoint and perpendicular to the flight path. At the time the cameras were triggered, the aircraft position, determined by a DGPS Max (CSI Wireless, Calgary, Canada) differentially corrected GPS receiver with submeter accuracy, was logged and the image was associated with the planned image location. Laser altimeter data were also logged with a time stamp so that the aircraft's altitude AGL at the time the cameras were triggered could be precisely determined. For a more complete description of the VLSA imaging system see Booth and Cox (2006).

QuickBird imagery, used as the base for georeferencing, was a panchromatic scene captured 14 October 2002. The QuickBird imagery has a spatial resolution of 0.6 m, and at this resolution, reflectance patterns associated with roads, rock outcrop, bare soil, and varying shrub density are clear. The imagery is not of sufficient resolution for determining, for

example, what shrub species are producing an observed reflectance pattern.

### Georeferencing VLSA images

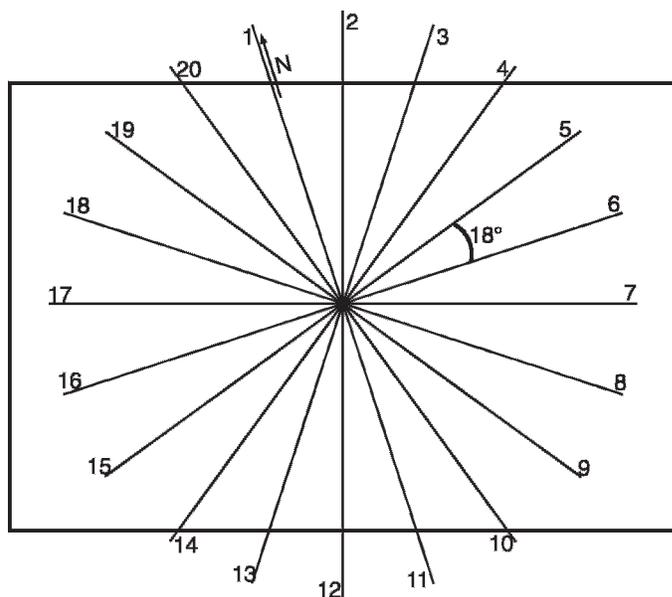
The VLSA images were georeferenced to reduce the search area when locating the true 16-MP image position. At each planned waypoint, one VLSA image was captured from each camera, with the 16-MP image nested within the 11-MP image. The 11-MP image, with its large FOV, exhibited patterns that could be discerned in the panchromatic QuickBird scene, but patterns in the 16-MP image lacked sufficient spatial context to recognize the location on the QuickBird image.

The process of georeferencing the small-FOV image required several steps. The first step was to use aircraft height, position, and orientation data to position, rotate, and scale the 11-MP image in a geographic information system (GIS; ArcMap 9.1; ESRI, Redlands, CA). The 11-MP image was then compared with the QuickBird image and five to seven ground control points were found and used to georeference the VLSA image (residual mean square error [RMSE] = 0.4 m). The 16-MP image was then loaded into the GIS and positioned, rotated, and scaled so that it was near the 11-MP image center. Patterns were compared between 16- and 11-MP images and five ground control points were selected (one near each corner and one near the center) to georeference the 16-MP image (RMSE = 0.01 m). Finally, the 16-MP image was printed and carried to the field along with coordinates of the georeferenced image center and corner locations. A Trimble GeoXT GPS receiver (Trimble Navigation Limited, Sunnyvale, CA) was used to navigate to the points. Ground and image patterns were compared to find the true locations of one center point and four corner points. Wood stakes were driven into the ground at these points and the coordinates of these locations were recorded with the same GPS receiver (average of 50 position fixes) as used to navigate to the plot. The type of GPS receiver used for this study has an accuracy of 1 m with wide area augmentation system enabled (Serr et al. 2006). The distance between corner stakes was measured with a cloth tape and used to compute the true image GSD. Thus, the four methods for defining image position were 1) preplanned waypoints, 2) aircraft GPS location, 3) georeferenced to the QuickBird image, and 4) true plot center point as determined by the GPS receiver.

### Shrub Cover Estimation

Sagebrush (*Artemisia* L.), spineless horsebrush, and antelope bitterbrush cover were classified in 21 VLSA images (16 MP). The 21 images used were from sites that spanned a >70-yr range since last burn. The variability in years since last burn was stratified into seven ranges (1–4, 7–8, 11–16, 25, 27–32, 59–69, and >70 years), and three images were selected at random from within each stratum.

One hundred pixels were selected at random from each image, and each pixel was characterized as one of the three shrub taxa, other, or unknown. Cover was measured as the proportion of sampled pixels from each class. Bitterbrush regions were also delineated and counted in each image. The number of pixels within bitterbrush regions was divided by the total number of pixels in the image to get fraction of bitterbrush cover. ImageJ software (Rasband 2006) was used



**Figure 1.** Diagram showing the relationship between the image plot boundary and the numbered line transects used for measuring shrub cover on the ground. The rectangular box shows the 8-m by 11-m area of the very large scale aerial image.

for image measurements. Conceptually, the classification method was the same as is done with the SamplePoint software (Booth et al. 2006a). However, ImageJ software was chosen because it is a more flexible image toolkit, which can be extended with macros and plug-ins to accomplish specific tasks. For example, an ImageJ plug-in was developed at the USSES to randomly select pixels for classification, provide a user interface for recording the pixel coordinates and the user's pixel classification, and record characteristics of delineated regions such as size, location, and classification.

Ground-based shrub cover measurements were made in a 5-m radius region centered on the image center stake (Fig. 1). Cover was measured along 20 lines radiating from the center point. Each line was 5 m and was numbered sequentially from 1 to 20, beginning with the north line and rotating 18° clockwise between each line. For repeatability analysis, the cover measured from the 10 even-numbered lines and the 10 odd-numbered lines were compared. Cover for each shrub species was taken as the sum of the shrub interval lengths for each shrub species divided by the total sampled length (100 m). If an interval had two or more shrub species, then the interval length was divided by the number of species, and that length was assigned to each species on that interval. For the ground measurement of bitterbrush numbers, bitterbrush plants with some portion inside the staked plot boundary were counted.

### Statistical Analysis for Location and GSD Data

Differences between preplanned waypoints, aircraft locations, georeferenced locations, and true GPS positions were characterized by considering each dimension of a coordinate system independently and fitting those data to the following linear model:

$$A_i - B_i = \beta_0 + \varepsilon_i \quad [1]$$

where  $B_i$  is  $i$ th coordinate (of the respective dimension: easting, northing, forward, or left) from position source  $B$  (planned, aircraft, georeferenced, or true),  $A_i$  is the  $i$ th coordinate from the alternative position source,  $\beta_0$  is the bias term, and  $\varepsilon_i$  is the residual error. This model was used to determine whether location was significantly biased. Empirical cumulative distribution functions were developed for values of uncorrected and bias-corrected deviations to determine 95% search distance along each dimension. The 95th percentile search area (SA) was computed as an ellipse:

$$SA = \pi r_1 r_2 \quad [2]$$

where  $r_1$  is the 95th percentile error in one dimension and  $r_2$  is the 95th percentile error in the orthogonal dimension. These search areas were computed for both measured and bias-corrected locations, using  $A_i - B_i$  and  $\varepsilon_i$  error distributions, respectively. For comparisons between preplanned waypoint and aircraft locations, data from all 1134 acquisitions were used in the analysis. All other comparisons used data from the 21 validation plots. In all analyses for this study, unless otherwise specified, the type I error rate ( $\alpha$ ) was 0.05.

### Statistical Analysis for Cover and Shrub Numbers

A paired  $t$ -test was used to determine whether the mean response was different between the VLSA and ground methods. Simple linear regression analysis was performed to characterize the relationship between an image-measured attribute ( $B$ ) and the corresponding ground-measured attribute ( $A$ ). In cases where the regression was significant, the measured image attributes were rescaled to the calibrated image measurements ( $B'$ ):

$$B' = \beta_0 + \beta_1 B \quad [3]$$

These adjustments assured that the relationship between  $A$  and  $B'$  had a 0 intercept and unity slope, but this linear scaling had no effect on the RMSE for the regression. Bland and Altman's (1986, 1999) limit-of-agreement (LoA) method was used to assess agreement. The LoA method is used to plot the difference between paired methods against the mean of the paired methods. The LoA was computed for uncalibrated relationships when neither the slope or intercept were significant and for calibrated relationships if slope, intercept, or both were significant. LoA is expressed as upper and lower limits, but for purposes of comparison with the repeatability coefficient, half the limit of agreement range is reported (i.e.,  $0.5 \text{ LoA} = [\text{upper limit} - \text{lower limit}]/2$ ). Repeatability was determined for the ground-measured cover by comparing measurements of the 10 even-numbered transect lines with measurements of the 10 odd-numbered transect lines. The 95% repeatability coefficient was calculated as  $2.77s_p$ ,

where  $s_p$  was the within-plot standard deviation, pooled across all plots (Bland and Altman 1999):

$$s_p = \sqrt{\text{var}(A - B)} \quad [4]$$

where  $A - B$  is the set of differences between measurements.

Georeferencing images and conducting the ground search to find true VLSA image locations costs time. An analysis was made to determine how many fewer samples were needed with each method to determine whether or not they gave different results for plot-level pairs (i.e., measurements made on the same plot, as in Fig. 1) or management unit-level pairs (i.e., measurements made on different plots, but with both plots being samples of areas with similar management). The minimum sample size was defined as the sample size required to detect a 4% difference in cover (or in the case of bitterbrush numbers, a difference of four plants) between methods (5% significance level and 90% power) when measurements were paired on the same plot and when measurements were paired on independent samples within the same management unit. This approach determines the minimum number of samples needed to detect bias between methods, but does not address differences in the methods' repeatability or their limits of agreement. The 21 plots used for this analysis were composed of seven management units that differed in the number of years since last burned, and had three observations per management unit. For measurements paired at the management unit level, ground- and image-measured samples are considered independent and include both measurement and within-management unit variation, so the standard deviation of a difference between two methods paired at the management unit level was calculated as the following:

$$s_m = \sqrt{\text{var}(A') + \text{var}(B')} \quad [5]$$

where  $A'$  and  $B'$  are the sets of within-group errors for ground- and image-based measurements, respectively. Because ground and image measurements were treated as independent within the management unit, the covariance term is 0 and therefore excluded. Minimum sample size was calculated viewing the samples as paired at either the plot level, with  $s_p$  standard deviation (Equation 4), or at the management unit level, with  $s_m$  standard deviation (Equation 5). The "power.t.test" function, in R (version 2.7.1), was used to calculate the minimum sample size, setting the significance level to 0.05, power of test to 0.90, type to paired, and alternative hypothesis to two-sided (R Development Core Team 2008). The resulting minimum sample size is the number of pairs required to detect a 4% difference in cover (or a four-plant difference) between the methods.

## RESULTS

### Location Accuracy and Precision

On average, there was little bias in aircraft locations relative to preplanned waypoints. The bias in the aircraft locations was on average 0.09 m from the preplanned waypoints. However, the

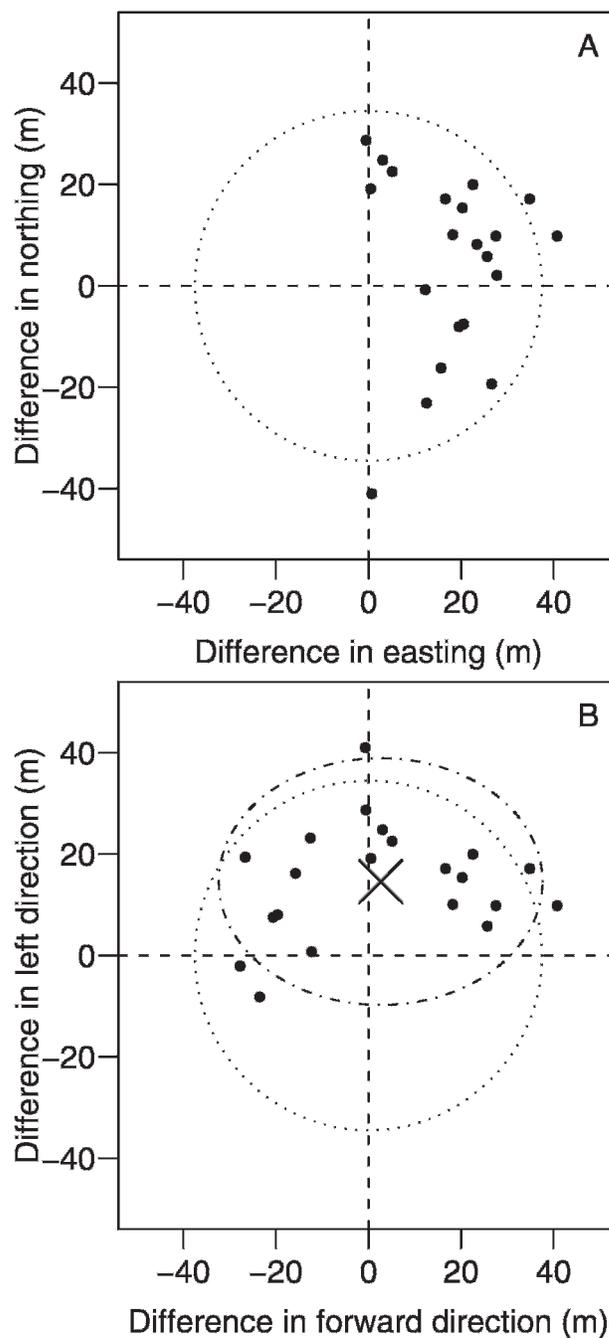
deviations between aircraft and preplanned locations were large; the maximum deviation was 49.01 m. The 95th percentile search area for a preplanned location from an aircraft location was 333 m<sup>2</sup>.

Characteristic patterns in the images could be matched with the pattern on the ground so as to be certain that ground locations were the true locations represented in the image. The true ground locations deviated significantly from the location of the aircraft at the time of image acquisition. The median deviation was 25.5 m, and 5% of distances were  $\geq 41.4$  m (Fig. 2A). Twelve VLSA images were collected while the aircraft was flying east and nine while the aircraft was flying west. Because camera alignment is a potential source of bias that would be canceled out and manifest as random error with equal numbers of images collected going east and west, a forward and left coordinate system was used to account for this. The GPS positions were, on average, shifted 14.6 m left of the aircraft location. After correcting for this bias, the random errors forward and left of the aircraft were still large (Fig. 2B). Error in the forward dimension between true GPS and corrected aircraft-based location was greater than the left position error. The median forward error from the corrected position was 18.6 m, and 5% of the locations had a forward error greater than 35.0 m. The median error in the left dimension was only 6.4 m, and 5% of the locations had a left dimension error greater than 24.4 m. The 95th percentile search area for true positions, beginning at an uncorrected aircraft easting and northing, was 4060 m<sup>2</sup>. The 95th percentile search area for true positions, when beginning at a bias corrected aircraft location, was still too large (2675 m<sup>2</sup>) to efficiently find matching image points from a ground search.

The positions obtained from georeferencing an image to QuickBird through an intermediate-scale VLSA image were near the true positions as determined on the ground using a GPS. There was no significant difference in the mean difference of either easting or northing between the georeferenced positions and the true ground location. Half the deviations were within 2.1 m, and 80% were within 4.6 m. Of the 21 images used, two were more than 19.9 m from the georectified location. In both cases, the image was of a location that had been burned within the last 4 yr, and the satellite scene did not show sufficient vegetation pattern to confidently select ground control points. Excluding these two images, 90% of the true locations were within 5.1 m (82 m<sup>2</sup>), and 80% were within 3.7 m (Fig. 3). A portion of this error in location is due to the GPS receiver error.

### Aircraft and Ground Based GSD

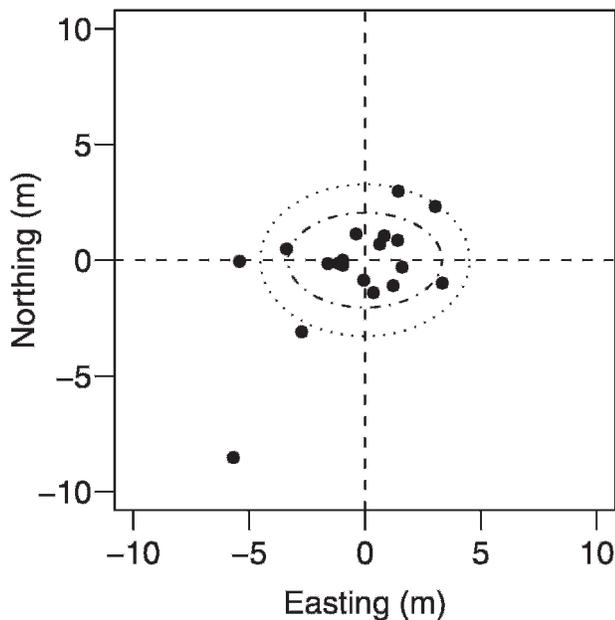
The mean aircraft-based GSD for 18 images averaged 2.3 mm  $\pm$  0.1 SD. The mean ground-based GSD for the 18 validation images was 2.4 mm  $\pm$  0.1 SD, which by paired *t*-test was significantly different from the 18 aircraft-based values. For subsequent analyses, aircraft-based GSD values were centered, that is to say the mean was subtracted from each value. For the 18 validation images, the relationship between ground and centered aircraft-based GSD had a 2.4-mm intercept and 0.92 slope. The intercept was significantly different from the mean aircraft GSD, which is consistent with the paired *t*-test, but the slope was not different from 1. This



**Figure 2.** Differences in position between the true image center point locations ( $n=21$ ) and the aircraft-based location in dimensions of (A) easting and northing and (B) forward and left. The dotted ellipses delineated the area in which 95% of the true locations were found without correcting for bias. The dash-dot ellipse, in panel B, delineates the search area in which 95% of the true locations were found after correction for bias in the aircraft location at the time of image acquisition.

means that the aircraft GSD underestimated the ground GSD by 0.1 mm or 5.0% and the area of a pixel by 0.58 mm<sup>2</sup> or 9.8%. A 1-m length on the ground would be estimated as 0.95 m, and a 1-m<sup>2</sup> area would be estimated 0.90 m<sup>2</sup>.

The one-half LoA width between aircraft- and ground-based GSD measurements was 0.1 mm or 3.2%. The mean plot area measured on the ground and calculated as mean width  $\times$  mean

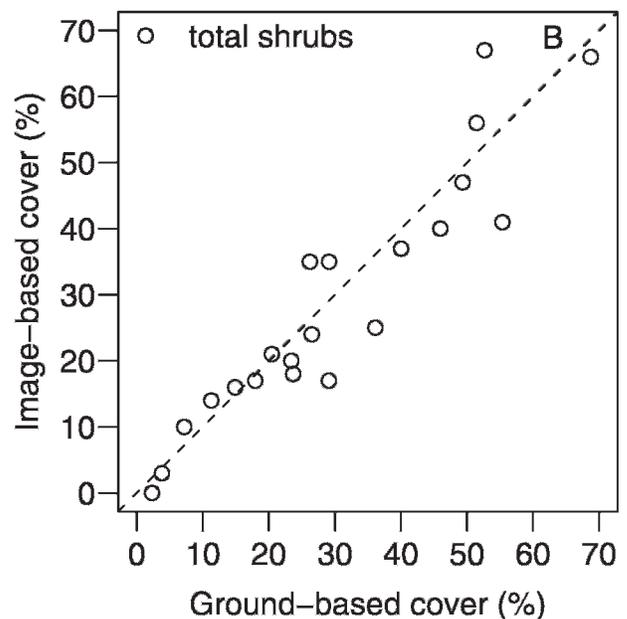
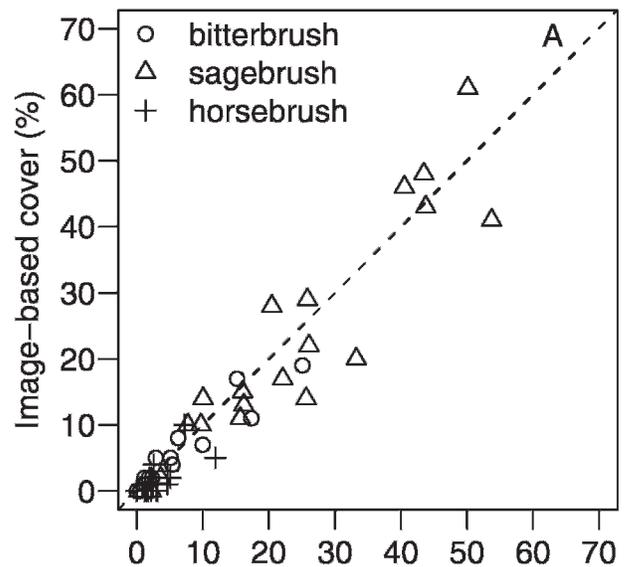


**Figure 3.** Difference in position between the true image center point locations ( $n=19$ ) and the georeferenced location in dimensions of easting and northing. Two images lacked sufficient pattern to georeference the image and were excluded. The dotted ellipses delineated the area in which 90% of the true locations were found. The dash-dot ellipse delineates the search area in which 80% of the true locations were found.

length was  $100 \text{ m}^2$ ; calculated as the product of aircraft-based GSD squared, pixel width, and pixel length, it was  $88.4 \text{ m}^2$ . The mean plot area calculated as the product of the bias-corrected aircraft GSD (i.e., aircraft GSD +  $0.12 \text{ mm}$ ) squared, pixel width, and pixel length was  $98.0 \text{ m}^2$ .

### Shrub Canopy Cover and Shrub Numbers

Table 1 shows statistics for ground-measured shrub canopy cover by shrub taxa using the line-intercept method from 21 plots used in the validation. These data are plotted in Figure 4 along with the paired measurements made from the VLSA imagery using the point-intercept method. Total shrub cover spanned the range from 2.3% to 68.7%, with a mean of 30.3%. Repeatability depended on the cover range. In general, the repeatability coefficient increased as mean shrub cover increased, meaning that line-intercept estimates are not as repeatable when shrub cover is at the upper end of the measured range.



**Figure 4.** Comparison of very large scale aerial image-based and ground-based cover for (A) sagebrush, bitterbrush, and horsebrush, and (B) the total shrub cover ( $n=21$ ). Image-based cover is measured from point-intercept and ground-based cover is based on line-intercept. The 1:1 dashed line is shown to aid comparison.

**Table 1.** Shrub canopy cover from line-intercept and bitterbrush count statistics measured in a mountain big sagebrush community ( $n=21$ ).

Shrub type	Ground measure	Mean	SE	Minimum	Maximum	Repeatability coefficient
Bitterbrush	Cover	4.6%	1.5%	0.0%	25.1%	3.83% <sup>1</sup>
Bitterbrush	Plants	9	2.1	0	30	—
Sagebrush	Cover	22.2%	3.6%	0.3%	53.7%	11.82%
Horsebrush	Cover	2.7%	0.6%	0.0%	11.9%	2.25% <sup>2</sup>
All shrubs	Cover	30.3%	4.0%	2.3%	68.7%	11.84%

<sup>1</sup>Repeatability coefficient was calculated from the plots where bitterbrush cover was greater than 0 for both even- and odd-numbered transect lines ( $n=15$ ).

<sup>2</sup>Repeatability coefficient was calculated from the plots where horsebrush cover was greater than 0 for both even- and odd-numbered transect lines ( $n=17$ ).

**Table 2.** Comparison of (A) ground- and (B) very large scale aerial (VLSA) image-measured shrub cover and bitterbrush plant counts ( $n = 21$ ). All ground-measured cover is by line-intercept (100 m total transect length) and image-measured cover is by point-intercept (100 random points).

Shrub type	VLSA measure <sup>1</sup>	Mean $A - B$ (%)	Model: $A = B' + \varepsilon = \beta_0 + \beta_1 B + \varepsilon$			
			$\beta_0$ (%)	$\beta_1$	$R^2$	Half LoA width (%)
Bitterbrush	DR	-0.81	0.48	0.76 <sup>2</sup>	0.94	3.18 <sup>3</sup>
Bitterbrush	DRC	2.00 <sup>4</sup>	1.69	1.04	0.89	6.16
Bitterbrush	PI	0.67	0.03	1.16	0.92	3.56
Sagebrush	PI	1.10	3.59	0.88	0.87	12.16
Horsebrush	PI	1.20 <sup>4</sup>	1.44 <sup>4</sup>	0.84	0.51	3.55 <sup>3</sup>
Total shrubs	PI	1.28	3.85	0.91	0.87	13.19

<sup>1</sup>VLSA measures are DR, delineated region; DRC, delineated region count; and PI, point-intercept.

<sup>2</sup>Statistically different from 1 ( $\alpha = 0.05$ ).

<sup>3</sup>Half limit-of-agreement (LoA) is for calibrated relationship ( $A$  and  $B'$ ) when the regression relationship was significant.

<sup>4</sup>Statistically different from 0 ( $\alpha = 0.05$ ).

**Bitterbrush.** Mean bitterbrush cover from the 21 locations was similar among the three measurement methods with 4.6% measured on the ground with line intercept, 5.4% measured on the image with delineated regions, and 4.0% measured on the image using point-intercept (Table 1; Fig. 4). The range in cover measured on the ground was 0% to 25.1%. The repeatability coefficient for the ground-measured bitterbrush canopy cover was 3.8% (Table 1). The slope of the relationship between ground-measured and delineated region image-measured cover was significantly less than 1, but the intercept was not different from 0 (Table 2). The mean difference in ground-measured and delineated region image-measured cover was not significantly different from 0, which suggests that small bitterbrush plants were not missed in the VLSA imagery (Table 2). Agreement between ground-measured and calibrated delineated region-measured cover was good with a one-half LoA width of 3.2%, which was nearly the same as the repeatability coefficient of the ground measurement method (Tables 1 and 2).

The relationship between ground- and image-measured cover (point-intercept) was not different from a 1:1 relationship and the mean difference between ground and VLSA measurements was not different from 0 (Table 2). Half the LoA width between ground- and image-measured cover was 3.6%, which is a slight reduction in agreement compared with the calibrated delineated regions method (Table 2). The point-intercept image-measured one-half LoA width was comparable to the ground-measured cover repeatability coefficient (Tables 1 and 2).

The number of bitterbrush plants counted in plots on the ground averaged nine and ranged between zero and 30 (Table 1). The number of bitterbrush plants measured in the VLSA images (i.e., delineated regions count) did not agree well with ground measurements. On average there were two fewer plants counted in the imagery than on the ground (Table 2). The relationship between the two methods had an intercept of 1.69 and a slope of 1.04 (Table 2), but the intercept was not significantly different from 0 and the slope was not different from 1. The one-half LoA width between VLSA and ground counts was  $\pm 6.30$  plants (Table 2).

**Sagebrush.** Threetip sagebrush and mountain big sagebrush could not be confidently distinguished from one another in the VLSA imagery, so for both VLSA and ground measurements

these two species were combined into sagebrush. The sagebrush cover ranged between 0.3% and 53.7%, with a mean of 22.2% as measured with ground methods (Table 1). The VLSA-measured cover for sagebrush ranged between 0% and 61% with a mean of 21.1% and the two methods resulted in similar measurements (Table 2; Fig. 4). The mean difference between ground- and VLSA-measured cover was 1.1%, which was not significantly different from 0 (Table 2). The intercept was not different from 0, and the slope was not different from 1. Half LoA width was 12.2%, which is not as narrow as that obtained for bitterbrush using either VLSA method, but is only slightly greater than the repeatability coefficient.

**Horsebrush.** Horsebrush cover, measured on the ground, had the smallest range of the three shrub taxa (0% to 11.9%) and the lowest mean (2.7%, Table 1). The repeatability coefficient for ground-measured horsebrush cover was 2.3% (Table 1). Mean VLSA-measured cover was 1.5%, on average 1.2% less than the ground measure (Table 2), and ranged between 0% and 10%. The relationship between ground and VLSA point cover had a 1.4 intercept, which was significantly different from 0, but the 0.84 slope was not different from 1 (Table 2). Half the LoA width between the ground and VLSA point cover was 3.55%, which is slightly larger than the repeatability coefficient of the ground method (Table 2).

**All Shrubs.** Total shrub cover, as measured on the ground, averaged 30.3% and ranged between 2.3% and 68.7% (Table 2). The 95% repeatability coefficient for ground-measured shrub cover, irrespective of shrub species, was 11.8% (Table 2). The VLSA image-measured cover ranged between 0% and 67% with a mean of 29%, which averaged 1.3% less than the ground measure (Table 2). The intercept was not different from 0, and the slope was not different from 1 (Table 2). Half LoA width was 13.2%, which is slightly larger than then the repeatability coefficient (Tables 1 and 2).

### Minimum Sample Size

The increase in sampling efficiency was variable among the measures, but on average, the number of pairs needed to detect a 4% difference in cover between methods paired at the management unit level was five times the number required when paired at the plot level (Table 3). Horsebrush was the

**Table 3.** Comparison of the sample size required for two sampling schemes to detect a 4% difference in cover between methods at 5% significance level and 90% power. The first scheme uses paired measurements at the plot level and the other used paired measurements within strata grouped into management unit levels. The standard deviations  $s_p$  and  $s_m$  are the pooled within paired standard deviation for pairs at the plot and management unit levels, respectively.

Shrub type	VLSA measure <sup>1</sup>	Plot level pairs		Management unit level pairs	
		$s_p$	Minimum sample size	$s_m$	Minimum sample size
Bitterbrush	DR	1.85	5	9.29	59
Bitterbrush	DRC	3.14	9	10.83	79
Bitterbrush	PI	1.82	5	7.56	40
Sagebrush	PI	6.20	28	13.83	128
Horsebrush	PI	2.00	5	2.84	8
Total shrubs	PI	6.73	32	12.32	102

<sup>1</sup>Very large scale aerial (VLSA) measures are DR, delineated region; DRC, delineated region count; and PI, point-intercept.

exception, and only twice as many management unit-level samples were needed compared with the plot-level samples. Among the point-intercept measures, pairing measurements at the plot level was most advantageous for bitterbrush, requiring only one-eighth as many pairs as required for the management-unit pairing.

## DISCUSSION

### Location Accuracy and Precision

With the current VLSA system, the true ground location of an image is dependent on where the aircraft was and the direction that the camera was pointed when the camera was triggered. The direction that the camera was pointed varies due to aircraft pitch and roll, and camera misalignment. If the camera was always pointed nadir at the time of triggering, then the true position would only depend on the location of the aircraft and the accuracy of the aircraft GPS system, and there would be little need for georeferencing the VLSA imagery. This warrants consideration of a modification to the VLSA system to use an inertial measurement unit so that error due to pitch, roll, and yaw of the aircraft can be removed. Theoretically, this type of technology would be helpful for users that need better location accuracy and precision. However, this solution would not improve the ability of the system to place an image any closer to planned locations than the current system and would still be unsatisfactory for repeat photography of small (< 100 m<sup>2</sup>) plots.

To validate the shrub cover estimates made with VLSA images, true locations of the VLSA images were needed so that ground measurements could be made on the same location. For all 21 images, the true locations were confidently determined. In two cases, when only aircraft locations at the time of image acquisition were available, search time to find the true image location exceeded 1 h. With bias-corrected aircraft locations as starting points, paired method validation data would be difficult to obtain. Blumenthal et al. (2007) were successful in finding true VLSA image locations, but did not report how much effort was required in finding these locations.

With the degree of accuracy and precision achieved using the described georeferencing method, the 80th percentile search area for finding matching image and ground points is less than 2% of the search area that would be required from aircraft-based locations alone. For 80% of the images, the search area was less than 21 m<sup>2</sup>, 1.8% of the corresponding 80th percentile search area from bias-corrected aircraft locations; 50% of the images had a search area of less than 3.7 m<sup>2</sup>, 1% of the corresponding 50th percentile search area from bias-corrected aircraft locations.

For 80% of the 16-MP images studied, the georeferenced FOV overlapped greater than 50% of the true ground area. This indicates that searching for matched points might not be necessary to obtain reasonably similar paired methods data. In other words, with 50% overlap between the georeferenced and true FOV, ground sampling results for either area would likely give similar results in this environment.

### Aircraft and Ground Based GSD

The aircraft-based GSD was biased, tending to underestimate the true GSD. Several sources of error, such as pitch and roll of the aircraft, might explain this bias (Booth et al. 2006b). Another source of error explained in Booth et al. (2006b) would be misalignment of the laser altimeter and the camera line of sight. Booth et al. (2006b) reported agreement between mean aerial measurement and ground measurements of object width on the basis of image interpretation and aircraft-estimated GSD. Their study compared ground and VLSA width measurements of objects in the 0.13 m to 1.44 m range (Booth et al. 2006b). The imagery was collected with an 11-MP camera fitted with a 100-mm focal-length lens at altitudes ranging between 50 m and 150 m AGL (Booth et al. 2006b). The 100-mm focal-length lens would have greater variation in GSD within an image than the 840-mm focal-length lens due to differences in the angle of view. If relative sizes are being compared, then the limit of agreement is sufficient even when uncalibrated. For absolute measurements, the aircraft-based GSD may need to be calibrated. For critical measurements, or where laser altimeter and camera alignment are uncertain, bias is easily corrected with calibration.

### Shrub Canopy Cover and Shrub Number Agreement

In the present study, regardless of the shrub species, there was good agreement between VLSA- and ground-measured shrub cover. Half the LoA width between two methods cannot be smaller than the repeatability coefficient for either method, and for each shrub type the half LoA width was within 1.4% of the repeatability coefficient (Tables 1 and 2).

In general, the intercept for the regression relationships were not different from 0, and the slopes were not different from 1. The exceptions were point-intercept-measured horsebrush cover and delineated-region-measured bitterbrush cover. The relationship between ground- and VLSA-measured horsebrush cover had an intercept greater than 0, but the slope was not different from 1. The VLSA-measured cover underestimated ground-measured horsebrush cover by 1.2% regardless of the amount of horsebrush cover. Mature horsebrush plants are generally shorter than mature sagebrush or bitterbrush and are more likely to be obstructed from view. The slope for the

relationship between ground-based and delineated region-based measurements of bitterbrush cover was less than 1, but the intercept was not different from 0 and the mean difference was not different from 0. Bitterbrush cover was overestimated with this method when there was > 5% bitterbrush cover. It is unlikely that this is due to mistaking another shrub for bitterbrush because the error was not observed for the point-sample method. A more likely explanation is that the software plug-in used did not facilitate exclusion of regions within a canopy (i.e., islands) that were not bitterbrush. These islands were more likely at high bitterbrush cover levels where shrubs were typically larger. This could be addressed by delineating these islands and subtracting their area from the total.

Other authors have compared vegetation cover classification from ground-based high-resolution imagery with ground measurements of cover (Seefeldt and Booth 2006, Laliberte et al. 2007). Laliberte et al (2007) made paired comparisons of cover at the plot level and used appropriate methods to validate the cover measurements. The ground-based imagery used in that study differed from VLSA imagery in terms of altitude (< 2.8 m), camera angle of view, and camera motion; furthermore, the image measurement method used was automated rather than based on human interpretation.

Numbers of bitterbrush plants were measured in the VLSA images and on the ground, but agreement was less than for cover. Disagreement may have arisen from 1) not being able to determine whether a region was a cluster of many plants or a single plant and 2) not being able to count small plants that contributed little to the cover. It is doubtful that VLSA imagery could be used to make scientifically sound estimates of bitterbrush density because of difficulty in obtaining good bitterbrush plant counts and GSD errors. The distribution of bitterbrush individuals is often clumped, with several plants growing in close proximity to one another, making them difficult to differentiate. Another error in counting bitterbrush individuals occurs when stems contact the ground and become rooted so that a single genet appears as several distinct individuals in the imagery. Determining density in the field often requires close examination of stems. The utility of VLSA imagery for estimating shrub density for other species was not determined; however, it is conceivable that VLSA may be useful for other plant species that do not layer or aggregate if the GSD were calibrated so the area estimated for each image was accurate.

### Minimum Sample Size

The importance of making measurements on the same plot when designing a validation study is clear. Fewer pairs were required to determine that image-measured cover was unbiased when paired measurements were made at the plot level, compared with making paired measurements at the management unit level. The number of pairs needed to detect a 4% difference in cover between methods was always less when paired measurements were made on the same experimental unit rather than on independent samples within the same management unit (Table 3). In this study, within-management unit sampling variability was reduced to be stratified into seven levels of year since last fire, which accounted for much of the variability in shrub cover. The difference in the number of pairs

required depends on the amount of within-management unit variability for the methods. As within management unit variability increases (e.g., bitterbrush has an aggregated distribution with large variability among < 100 m<sup>2</sup> plot measurements), the number of paired management unit samples required increases. Conversely, as within-management unit variability decreases (e.g., horsebrush had little spatial variability in cover within a management unit) the number of paired management unit samples required decreases and the need for true plot locations and ground measurements on the same plot is reduced.

## MANAGEMENT IMPLICATIONS

Two important questions relative to the suitability of VLSA imagery were addressed in this study. First, "Can shrub cover be accurately measured using VLSA imagery?" Second, "How well is the location of a VLSA image known?" The results of this study indicate that, in general, image-based shrub cover measurements are unbiased and have agreement between methods similar to the repeatability of conventional ground-based methods in mountain big sagebrush communities. Furthermore, several shrub species can be easily distinguished. VLSA imagery is an efficient and accurate substitute for ground-based shrub cover measurement in the sagebrush steppe. With respect to image location accuracy and precision, the question was framed from two perspectives. First, "How likely is it that a preselected location will be imaged?" The true image location can be expected to be placed within about a 4 000-m<sup>2</sup> area surrounding the planned location, which is typically sufficient for placing images within allotment or management unit boundaries. However, VLSA imagery is not a good candidate for repeated sampling of small plots (< 100 m<sup>2</sup>) nor is it a good candidate for one-time sampling of established small plots. A 100-m<sup>2</sup> VLSA image is unlikely to overlap a plot or previously imaged area at this scale. A second question is, "How well is the true location of a photographed area known?" This question is important if a ground visit is necessary to collect ground measurements. From bias-corrected aircraft information alone, the best that could be expected was to be within a 2 675-m<sup>2</sup> area of the true location. When the imagery was georeferenced using QuickBird and an intermediate scale VLSA image, the locations were known to within an 82-m<sup>2</sup> area of the true location. Georeferenced 100-m<sup>2</sup> FOV images overlap much of the true imaged area. These two locations are close enough so that the true locations were easily identified in the field from comparing patterns common to both the ground area and image. In spatially heterogeneous landscapes, the cost of georeferencing the images is justified for studies that require paired ground measurements. The reduction in sample size and the ability to accurately estimate the degree of method agreement justifies the cost of georeferencing.

## ACKNOWLEDGMENTS

The author thanks Lana Barg, Keith Klement, Tonya Thelen, and Cole Wagoner for assistance with data collection. The author would like to

acknowledge that the VLSA imagery used in this study was collected as part of a larger collaborative project with Dr Terry Booth at the US Department of Agriculture–Agricultural Research Service High Plains Grassland Research Station in Cheyenne, Wyoming.

## LITERATURE CITED

- BLAND, J. M., AND D. G. ALTMAN. 1986. Statistical methods for assessing agreement between two methods of clinical measurement. *Lancet* i:307–310.
- BLAND, J. M., AND D. G. ALTMAN. 1999. Measuring agreement in method comparison studies. *Statistical Methods in Medical Research* 8:135–160.
- BLUMENTHAL, D., D. T. BOOTH, S. E. COX, AND C. E. FERRIER. 2007. Large-scale aerial images capture details of invasive plant populations. *Rangeland Ecology and Management* 60:523–528.
- BOOTH, T. D., AND S. E. COX. 2006. Very large scale aerial photography for rangeland monitoring. *Geocarto International* 21:27–34.
- BOOTH, T. D., AND S. E. COX. 2008. Image-based monitoring to measure ecological change in rangeland. *Frontiers in Ecology and the Environment* 6:185–190.
- BOOTH, T. D., S. E. COX, AND R. D. BERRYMAN. 2006a. Point sampling digital imagery with 'samplepoint.' *Environmental Monitoring and Assessment* 123:97–108.
- BOOTH, T. D., S. E. COX, AND R. D. BERRYMAN. 2006b. Precision measurements from very-large scale aerial digital imagery. *Environmental Monitoring and Assessment* 112:293–307.
- BOOTH, T. D., S. E. COX, T. MEIKLE, AND H. R. ZUURING. 2008. Ground-cover measurements: assessing correlation among aerial and ground-based methods. *Environmental Management* 42:1091–1100.
- LALIBERTE, A. S., A. RANGO, J. E. HERRICK, E. L. FREDRICKSON, AND L. BURKETT. 2007. An object-based image analysis approach for determining fractional cover of senescent and green vegetation with digital plot photography. *Journal of Arid Environments* 69:1–14.
- NATIONAL OCEANIC AND ATMOSPHERIC ADMINISTRATION. 2003. Climatography of the United States no. 20, 1971–2000. Available at: [http://cdo.ncdc.noaa.gov/climate\\_normals/clim20/id/102707.pdf](http://cdo.ncdc.noaa.gov/climate_normals/clim20/id/102707.pdf). Accessed 5 December 2008.
- RASBAND, W. S. 2006. ImageJ. Bethesda, MD, USA: US National Institutes of Health. Available at: <http://rsb.info.nih.gov/ij/>. Accessed 5 December 2008.
- R DEVELOPMENT CORE TEAM. 2008. R: a language and environment for statistical computing. Vienna, Austria: R Foundation for Statistical Computing. Available at: <http://www.R-project.org>. Accessed 5 December 2008.
- SEEFELDT, S. S., AND D. T. BOOTH. 2006. Measuring plant cover in sagebrush steppe rangelands: a comparison of methods. *Environmental Management* 37(5):703–711.
- SERR, K., T. WINDHOLZ, AND K. WEBER. 2006. Comparing GPS receivers: a field study. *URISA Journal* 18:19–23.
- WEST, N. E. . Accounting for rangeland resources over entire landscapes. In: D. Eldridge and D. Freudenberger [EDS.]. People and rangelands: building the future. VI International Rangeland Congress; 19–23 July 1999. Volume 2. Aitkenvale, Queensland, Australia: VI International Rangeland Congress, Inc.