Effect of Coregistration Error on Patchy Target Detection using High-Resolution Imagery

Keith T. Weber, GISP, GIS Director, Idaho State University, GIS Training and Research Center, 921 S. 8th Ave., Stop 8104, Pocatello Idaho 83209-8104 (webekeit@isu.edu)

Jérôme Théau, Idaho State University, GIS Training and Research Center, 921 S. 8th Ave., Stop 8104, Pocatello Idaho 83209-8104 (theajero@isu.edu)

Kindra Serr, Idaho State University, GIS Training and Research Center, 921 S. 8th Ave., Stop 8104, Pocatello Idaho 83209-8104 (serrkind@isu.edu)

ABSTRACT
Many factors influence classification accuracy and a typical error budget includes uncertainty arising from the 1) selection of processing algorithms, 2) selection of training sites, 3) quality of orthorectification, and 4) atmospheric effects. With the development of high spatial resolution imagery, the impact of errors in geographic coregistration between imagery and field sites has become apparent -- and potentially limiting-- for classification applications, especially those involving patchy target detection. The goal of this study was to document and quantify the effect of coregistration error between imagery and field sites on classification accuracy. Artificial patchy targets were randomly placed over a study area covered by a QuickBird image. Classification accuracy of these targets was assessed at two levels of coregistration. Results showed that producer’s accuracy of target classification increased from 37.5% to 100% between low and high levels of coregistration respectively. In addition, “Error due to Location”, a measure of how well pixels were located within respective classes, decreased to zero at high coregistration levels. This study highlights the importance of considering coregistration between imagery and field sites in the error budget, especially with studies involving high spatial resolution imagery and patchy target detection.

KEYWORDS: Quickbird, coregistration, field sites, positional accuracy, classification accuracy, patchy targets
INTRODUCTION

Much has been written about the effect of various input parameters and processing decisions on resulting image classification accuracy. Toward that end, researchers have investigated and published details describing the 1) selection of appropriate classification algorithms (Foody & Arora, 1997), 2) effect of the purity of training sites relative to a minimum ground cover threshold (Mundt et al., 2006), 3) influence of the orthorectification process (Cheng et al., 2003; Robertson, 2003; Toutin & Chenier, 2004; Wijnant & Steenberghen, 2004; Parcharidis et al., 2005), 4) impact of misregistration between image layers (Townshend et al., 1992; Dai & Khorram, 1998; Stow, 1999; Roy, 2000; Verbyla & Boles, 2000; Wang & Ellis, 2005), 5) influence of spectral resolution (Mehner et al., 2004), as well as the 6) influence of atmospheric anomalies and correction processes (Lillesand & Kiefer, 2000). The result of these and other efforts has allowed geospatial scientists to construct a fairly complete error budget and, thereby, better understand and interpret image classification results.

With the development and proliferation of high spatial resolution imagery (i.e., QuickBird and IKONOS) and positioning technologies that readily enable highly accurate training site location (i.e., sub-meter resource-grade GPS), another segment of the error budget has become apparent; geographic coregistration between imagery and field sites (i.e., training and validation samples). Prior to the development of high spatial resolution technologies, it was fairly easy to correctly locate a field site within the correct pixel of existing imagery (such as Landsat with 28.5 m x 28.5 m pixels) using even uncorrected GPS locations. Today, however, it has become a challenge to reliably locate training sites within the correct and representative pixel (Weber, 2006). Further, as we explore the ability of remote sensing technologies to detect patchy and rare land features (e.g., those that may occupy only one QuickBird or IKONOS pixel), it is not only important but critical that field sites be placed within the correct pixel if one expects results with reliable accuracy (≥ 75% overall accuracy; Goodchild et al., 1994). However, accurate field site positioning may become less critical when target features grow larger and occupy numerous, contiguous pixels. The purpose of our research was to explore the effect of geographic coregistration between imagery and field sites (henceforth referred to as coregistration) as it relates to the detection and accurate classification of patchy and rare land features.

To our knowledge, no study has been performed to quantify the effect of this type of coregistration error on classification accuracy. However, Sanchez & Kooyman (2004) described limitations for classification of penguin habitat in Antarctica due to the positional accuracy of QuickBird imagery. This was not quantified, and the error examined was not coregistration between imagery and field sites, but rather the effect of image coregistration alone.

The potential significance of a coregistration effect was first noticed in the authors’ work while using QuickBird and SPOT 5 multi-spectral imagery to produce predictive presence and distribution models of a patchy invasive weed, leafy spurge (Euphorbia esula), at study sites in southeastern Idaho (Weber et al., 2006). The authors acquired QuickBird and SPOT 5 imagery during a time period when leafy spurge was believed to be most spectrally distinct from the matrix of other species and cover types (i.e., the pre-flowering and flowering stage). Using the location of known leafy spurge infestations (+/- 0.9m @ 95% CI), the authors applied a maximum-likelihood classifier to produce predictive presence/absence models of leafy spurge. The results indicated that QuickBird multispectral imagery could not produce reliable models (≥ 75% overall accuracy; Goodchild et al., 1994) in contrast with the models derived from SPOT
5 imagery which did produce reliable results (76%). The authors sought answers to explain why QuickBird imagery did not perform as well as SPOT 5 imagery under those conditions. What was most puzzling was the fact that the QuickBird sensor appeared to be far more technologically advanced compared to all other multispectral sensors available at that time (2003). The QuickBird sensor had far better 1) radiometric resolution (11-bit compared to 8-bit), 2) spatial resolution (2.4 m compared to 10 m), 3) comparable spectral resolution (blue, green, red, and near-infrared compared to SPOT’s green, red, near-infrared, short-wave infrared), and 4) very good signal-to-noise ratios. Yet, the models derived from the QuickBird imagery failed to achieve the same level of accuracy as those derived from the “simpler” SPOT 5 imagery.

The differences between the platforms were categorically addressed and new predictive models created for comparison: 1) imagery was converted from 11-bit to 8-bit by performing a linear histogram stretch, 2) imagery was resampled to produce a QuickBird product with 10m spatial resolution (using cubic convolution resampling) and thereby absorb georegistration errors, and 3) classifications were performed using only those bands in common between QuickBird and SPOT 5 platforms (green [560 nm versus 545 nm band centers, respectively], red [660 nm versus 645 nm band centers, respectively], and near-infrared [830 nm versus 840 nm band centers, respectively]) (note: these “common bands” allowed the authors to produce and use vegetation indices such as NDVI). After each adjustment was made, another classification was performed, producing a new presence/absence model for leafy spurge. In each case, the SPOT 5 imagery out-performed the QuickBird imagery. The only discrepancy that helped explain the performance difference was the fact that SPOT 5 imagery appeared to be better georeferenced (however, this was not quantifiable due to the remote nature of the study area and lack of ground control features). It was at this point that the authors designed an experiment to test and quantify the effect of coregistration error between imagery and field sites on classification accuracy. The paper focuses on a description of the experiment and its results.

METHODS

Study Area
The experiment was performed in sagebrush-steppe rangelands of southeastern Idaho approximately 30 km south of Pocatello, Idaho, at the O’Neal Ecological Reserve. The O’Neal Ecological Reserve (http://www.isu.edu/departments/CERE/o’neil.htm) was donated to the Department of Biological Sciences by Robin O’Neal. This 50 ha site, located along the Portneuf River, contains riparian areas along the river and typical sagebrush steppe upland areas located on lava benches. The O’Neal Ecological Reserve receives <0.38 m of precipitation (primarily in the winter) annually and is relatively flat with an elevation of approximately 1400 m (1401-1430 m). The dominant plant species is big sagebrush (Artemisia tridentata) with various native and non-native grasses, including indian rice grass (Oryzopsis hymenoides) and needle-and-thread (Stipa comata).

Field Data
Throughout the study area we placed 22 bright blue tarps (2.4m x 3.0 m) approximately equal in size to a single QuickBird pixel (2.4 x 2.4 m) (Fig. 1). The positioning of the tarps was random but with the following set of criteria established for final placement in the field 1) no part of the tarp was placed beneath vegetation, 2) tall vegetation (>1m) was not located near the tarps (+/−2m) that could cast a shadow on a portion of the tarp during image acquisition, and 3) tarps were installed flat and horizontal to
avoid deformation and changes in their apparent size within the imagery. The location of the tarps was taken using a Trimble ProXR GPS receiver and post-processed using base station files from Pocatello, Idaho (+/- 0.9m @ 95% CI) (Serr et al. 2006). QuickBird imagery was ordered and acquired while the tarps were in the field. When the imagery was delivered, the tarps were removed from the field.

Figure 1. The study site, location of blue tarps used for classification, and ground control platforms (with silver tarps) used for georectification.

In order to compare spectral properties of blue tarps with common rangeland elements, spectral signatures of various targets were acquired using an Analytical Spectral Device (ASD) hand-held FieldSpecPro field spectroradiometer. Measurements were made during a sunny day (without clouds) at +/- 1 hour of solar noon. For each target, between 15 and 25 spectral recordings were taken. Spectral comparison included blue tarps, bare ground, basalt, low sagebrush (Artemisia arbuscula), and big sagebrush (Artemisia tridentata).
To improve georegistration of the imagery within the relatively flat study area we constructed five permanent ground control platforms (Fig. 1). Each platform was 2.4 m x 2.4 m in size and stood 1.2 m above the ground. During satellite image acquisition periods, highly reflective silver tarps were tightly secured to the platforms. The location of the platform’s corners were acquired and processed with Trimble ProXR receivers in the same fashion as noted above (+/- 0.9 m @ 95% CI). All five ground control platforms were used to georectify the Quickbird imagery used in this study.

**Imagery**

Standard QuickBird imagery (28-June-2006) was delivered by DigitalGlobe Corporation projected into Idaho Transverse Mercator. The authors georectified the imagery using the ArcGIS 9.1 georectify tool. All five ground control platforms were clearly visible within the imagery making the positioning of control points quite easy. In this case, the “from” location was on-screen digitized and the “to” location was entered from the keyboard using the known GPS-based locations. Georectification was performed using a first order affine transformation with cubic convolution resampling.

Frequently, imagery is atmospherically corrected before georectification is performed to best preserve the original radiometric data. However, since the imagery was already projected upon delivery, the authors chose to perform atmospheric correction twice, once using the standard imagery as delivered and again using the georectified imagery. Atmospheric correction was performed with Idrisi Kilimanjaro (v14) using the ATMOSC module. All imagery was corrected using the Cos(t) model (Chavez, 1996) with input parameters reported in the metadata supplied by DigitalGlobe Corporation. Both the georectified and standard images (bands 1-4) were atmospherically corrected yielding four distinct datasets for use in this experiment: 1) standard imagery as delivered (standard), 2) standard imagery that was atmospherically corrected (atmos), 3) georectified imagery (geo), and 4) georectified imagery that was atmospherically corrected (geo-atmos).

A geodatabase feature class containing 50 points representing the location of the “target” blue tarps ($n=22$) and non-target points ($n=28$) was randomly resampled without replacement using Hawth’s tool in ArcGIS 9.1. This produced two datasets for use in the classification process. The first ($n=28$) was used as training sites (14 blue tarp and 14 non-target points) and the second ($n=22$) was used as validation sites (8 blue tarp and 14 non-target points)(note: ideally 14 blue tarps would have been available for validation, however based upon the author’s tarp positioning criteria, a total of only 22 tarps could be positioned in the field and remain in place throughout the satellite acquisition time window of approximately one month). Spectral signatures were extracted from each of the four imagery datasets (standard, atmos, geo, and geo-atmos) using the training site points within the MAKESIG module in Idrisi Kilimanjaro.

A series of maximum-likelihood classifications (Richards 1986) were performed using Idrisi Kilimanjaro (MAXLIKE) and validated using the ERRMAT module, which calculates both a standard contingency table (Congalton & Green, 1999) and Kappa statistic (Kappa Index of Agreement [KIA]) (Cohen, 1960; Titus et al., 1984; Foody, 1992; Monserud & Leemans, 1992). To better identify the source of classification error, the VALIDATE module of Idrisi was also used, which calculates a variety of statistics quantifying agreement between a classified image and reference image relative to the 1) quantity of cells in each class and 2) location of cells in each class (Pontius, 2000). The reference image was a
raster layer of the validation sites. The “Error due to Location” statistic is reported here as it indicates how well pixels are located within each class, and hence, best communicates the results of this study.

RESULTS AND DISCUSSION
The root mean square error (RMSE) reported during the georectification process of the QuickBird imagery was 0.20 m. True horizontal positional error was determined by measuring the distance from the known location (determined using GPS) to the center of each blue tarp’s location within the imagery. This calculation was performed twice, once using the standard imagery and again using the georectified imagery. The mean distance between the known blue tarp locations and 1) its location within the standard imagery was 1.55 m (median = 1.61 m) and 2) its location within the georectified imagery was 0.80 m (median = 0.55 m) (Fig. 2). The latter error was <50 % of the size of each QuickBird pixel while the former was >50 % of the size of each pixel. The measured difference in positional accuracy was tested using a paired t-test and Wilcoxon Signed Ranks and the improvement was found to be significantly different (P < 0.0001). Further, this difference is notable as other authors have hypothesized that field site locational error must be <50 % of the pixel size to yield reliable classification results (Peleg & Anderson, 2002; Weber, 2006), particularly in the case of patchy target detection. For instance, if a patch of leafy spurge covers only the area of a single QuickBird pixel, then a shift in the correct location of the field site—relative to the satellite imagery—of as little as half a pixel can not only lower classification accuracy but introduce a misclassification error into the spectral signatures that will propagate throughout the classification process when field sites are positioned over an entirely different class.

![Figure 2. Euclidean distance from known locations and the location of points (n=22) determined from georectified (mean and median = 0.80 and 0.55 m respectively) and standard QuickBird imagery (mean and median = 1.55 and 1.61 m respectively). The dotted diagonal line represents a hypothetical 1:1 relationship where no difference between standard and georectified locations would be measurable.](image-url)

The result of the maximum-likelihood classifications are given in Table 1(a-d). The classification results presented under sections a and c reveal a precision effect; that is the effect of coregistering imagery relative to the location of “known” field site locations used in the classification-validation process. By reducing the horizontal positional error between the target’s true location relative to its location within the
imagery, we were able to improve producer’s accuracy from an unacceptable 37.5% to a very reliable 100%. User accuracy was reduced during the georectification process (100% to 89%). However, upon closer inspection one can see that when using standard imagery (Table 1a), only 3 of the 8 blue tarp locations were detected, albeit each was correctly detected. A user would be able to find the three targets but would be blind to over 50% of the target population. Using the georectified imagery (Table 1c), users would find 100% of the target population and only one false-positive site. The latter is actually a much better scenario of operation for weed managers or other users of predictive maps.

Table 1. Error matrix describing maximum-likelihood classification results for detection of randomly placed blue tarps (a rare, patchy target) using QuickBird multispectral imagery.

<table>
<thead>
<tr>
<th>Known ground truth</th>
<th>Blue tarp target</th>
<th>Non-target</th>
<th>Total</th>
<th>Commission Error</th>
</tr>
</thead>
<tbody>
<tr>
<td>a) Standard imagery</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Blue tarp target</td>
<td>3</td>
<td>0</td>
<td>3</td>
<td>0.000</td>
</tr>
<tr>
<td>Non-target</td>
<td>5</td>
<td>14</td>
<td>19</td>
<td>0.263</td>
</tr>
<tr>
<td>Total</td>
<td>8</td>
<td>14</td>
<td>22</td>
<td></td>
</tr>
<tr>
<td>Omission Error</td>
<td>0.625</td>
<td>0.000</td>
<td></td>
<td>Overall error 0.227</td>
</tr>
<tr>
<td>KIA = 0.43; Error due to Location = 0.00*</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>b) Atmospherically corrected standard imagery</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Blue tarp target</td>
<td>0</td>
<td>4</td>
<td>4</td>
<td>1.000</td>
</tr>
<tr>
<td>Non-target</td>
<td>8</td>
<td>10</td>
<td>18</td>
<td>0.444</td>
</tr>
<tr>
<td>Total</td>
<td>8</td>
<td>14</td>
<td>22</td>
<td></td>
</tr>
<tr>
<td>Omission Error</td>
<td>1.000</td>
<td>0.286</td>
<td></td>
<td>Overall error 0.546</td>
</tr>
<tr>
<td>KIA = -0.32; Error due to Location = 0.36</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>c) Georectified imagery</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Blue tarp target</td>
<td>8</td>
<td>1</td>
<td>9</td>
<td>0.111</td>
</tr>
<tr>
<td>Non-target</td>
<td>0</td>
<td>13</td>
<td>13</td>
<td>0.000</td>
</tr>
<tr>
<td>Total</td>
<td>8</td>
<td>14</td>
<td>22</td>
<td></td>
</tr>
<tr>
<td>Omission Error</td>
<td>0.000</td>
<td>0.071</td>
<td></td>
<td>Overall error 0.046</td>
</tr>
<tr>
<td>KIA = 0.90; Error due to Location = 0.00*</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>d) Atmospherically corrected georectified imagery</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Blue tarp target</td>
<td>8</td>
<td>1</td>
<td>9</td>
<td>0.111</td>
</tr>
<tr>
<td>Non-target</td>
<td>0</td>
<td>13</td>
<td>13</td>
<td>0.000</td>
</tr>
<tr>
<td>Total</td>
<td>8</td>
<td>14</td>
<td>22</td>
<td></td>
</tr>
<tr>
<td>Omission Error</td>
<td>0.000</td>
<td>0.071</td>
<td></td>
<td>Overall error 0.046</td>
</tr>
<tr>
<td>KIA = 0.90; Error due to Location = 0.00*</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
</tbody>
</table>

Where KIA is the Kappa Index of Agreement

* indicates the spatial allocation of the pixels is as accurate as possible relative to the validation sites (Pontius, 2000).

The “Error due to Location” statistic further corroborated the inferred results. Where georectification was < 50% the size of a pixel, “Error due to Location” was zero, indicating that none of the disagreement between the predictive model and the reference image was due to locational error. In comparison, “Error due to Location” was as high as 36% in the case of the atmos imagery dataset (Table 1b) (Pontius, 2000).
No additional improvement in overall classification accuracy was seen using the atmospherically corrected data. This is principally because there was little potential for improvement as only one of the 22 target sites was incorrectly classified. These results should not be interpreted to indicate that atmospheric correction is unimportant, but rather that the effect of coregistration plays a significant role in a classification’s error budget.

Results from this study underline the impact of coregistration error. In this study, blue tarps were used to simulate homogeneous patchy targets, and while these targets were artificial, we investigated the spectra of these targets relative to adjacent, natural targets to better understand the classification results. Fig. 3 shows the spectra of the blue tarps has some similarities with adjacent, natural targets like sagebrush and bare ground, especially in the red and near-infrared regions. As a result, reducing coregistration error should benefit classification results of other natural patchy targets. However, the authors acknowledge another important factor influencing patchy target detection, target cover thresholds (Mundt et al., 2006). In this study, the targets (blue tarps) had 100% ground cover whereas many natural, patchy targets will exhibit a much lower ground cover making accurate classification more challenging. Other studies have developed methods to reduce the impact of positional error using spatial aggregation (Carmel et al., 2006) and epsilon band in a change detection context (Mas, 2006). However, these techniques may be of limited application in the context of patchy target detection (i.e. where target size is approximately the same as pixel size) due to the critical loss spatial resolution during aggregation and the subsequent loss of information.

![Figure 3. A comparison of spectral signatures from common rangeland targets and the artificial blue tarps used in this study. Signatures were acquired with a spectroradiometer and the mean signature of n (15-25) spectra are shown. QuickBird image bands are shown in grey for reference.](image-url)

This study was performed at the QuickBird spatial resolution as this sensor highlighted the challenge to accurately locate field sites within the correct and representative pixel. In semi-arid environments, this accuracy is critical because landscape features such as sagebrush, shrubs, patches of invasive weeds, and patches of bare ground are frequently found at the same spatial order (i.e., 1-4m). The authors hypothesize
that the coregistration effect described in this paper will diminish as the size of the pixel increases and the likelihood of field sites “automatically” being incorporated into the correct/representative pixel increases. As a result, coregistration error will be nil where field site positional error is small in proportion to a pixel’s size.

**CONCLUSIONS**

Coregistration error between imagery and field sites is an important consideration when evaluating classification results. With the development and proliferation of high spatial resolution imagery, a need arises to use high accuracy positioning technologies to ensure that field sites are correctly located within the representative pixel(s). Without appropriate allowances, classification accuracy may be seriously hindered especially when attempting to detect patchy and rare targets on the landscape.

**ACKNOWLEDGEMENTS**

This study was made possible by a grant from the National Aeronautics and Space Administration Goddard Space Flight Center (NNG06GD82G). Idaho State University (ISU) would also like to acknowledge the Idaho Delegation for their assistance in obtaining this grant and the assistance of Brett Thomassie and Jeff Adams of DigitalGlobe for their technical advice, and Jacob Tibbitts and Jamey Underwood (ISU) for their efforts in the field. This paper has also benefited from comments and suggestions by Drs. T. Sankey, R. G. Pontius, and two anonymous reviewers.

**LITERATURE CITED**


Recommended citation style:
[THIS PAGE LEFT BLANK INTENTIONALLY]