

Article

Mitigating the Impact of Field and Image Registration Errors through Spatial Aggregation

John Hogland ^{1,*} and David L.R. Affleck ²¹ Rocky Mountain Research Station, U.S. Forest Service, Missoula, MT 59801, USA² W.A. Franke College of Forestry and Conservation, University of Montana, Missoula, MT 59812, USA; david.affleck@umontana.edu

* Correspondence: jshogland@fs.fed.us; Tel.: +1-406-329-2138

Received: 18 December 2018; Accepted: 20 January 2019; Published: 22 January 2019



Abstract: Remotely sensed data are commonly used as predictor variables in spatially explicit models depicting landscape characteristics of interest (response) across broad extents, at relatively fine resolution. To create these models, variables are spatially registered to a known coordinate system and used to link responses with predictor variable values. Inherently, this linking process introduces measurement error into the response and predictors, which in the latter case causes attenuation bias. Through simulations, our findings indicate that the spatial correlation of response and predictor variables and their corresponding spatial registration (co-registration) errors can have a substantial impact on the bias and accuracy of linear models. Additionally, in this study we evaluate spatial aggregation as a mechanism to minimize the impact of co-registration errors, assess the impact of subsampling within the extent of sample units, and provide a technique that can be used to both determine the extent of an observational unit needed to minimize the impact of co-registration and quantify the amount of error potentially introduced into predictive models.

Keywords: attenuation; registration; aggregation; spatial correlation; co-registration

1. Introduction

Remotely sensed data play an ever-increasing role in characterizing and quantifying landscapes. These types of data have been used to study our surroundings [1], stratify the terrestrial environment [2,3], and build a wide range of data products depicting terrestrial characteristics, such as topography [4], land use and cover [5], vegetative indices [6], vegetation communities [7,8], fire severity [9], land cover change [10], and temperature [11]. Due to the success and relatively low cost of using remotely sensed data to depict landscape patterns and changes in those patterns, fields like landscape ecology [12] and concepts like spatial connectivity and the relationships between patterns and processes are now at the forefront of many land management and planning endeavors [13–16].

Ideas such as spatial contiguity, patch size, and patch juxtapositioning, and their relationships to processes and concepts such as forest management, land use planning, and sustainable forestry have in part fueled the desire to precisely and accurately define existing patterns at fine spatial detail, across broad extents [17–19]. Coupled with the availability of fine-grained remotely sensed data (≤ 5 m) and advancements in computer-based hardware and software [20], a fine-scaled depiction of the landscape can now be produced across broad extents relatively quickly, at a low cost [21–23]. At the same time, the fine-grain nature of these types of data provide unique opportunities to relate measured characteristics of the landscapes for small spatial extents (response variables) to remotely sensed data (predictor variables).

Many have capitalized on this point to develop mathematical, statistical, and spatial models that can be used to create surfaces depicting landscape variables of interest using geo-rectified field and remotely sensed data [18,22–24]. Generally, this process can be described as: (1) registering both field and remotely sensed data to a known coordinate system, (2) using the spatial coordinates of the field and remotely sensed data to link measured values in the field to remotely sensed data, (3) building a model for the linked variable as a function of variables derived from the remotely sensed data, and (4) applying the model to remotely sensed surfaces to create a continuous surface of estimated characteristics. While straightforward, the linking process is subject to error (co-registration error) owing to the imperfectly identified spatial coordinates of the response and predictors, and this can have a negative impact on the accuracy of the model estimates (i.e., increased bias and imprecision). With regression models, predictor variables (X_i) are assumed to be measured without error. Response variables (Y_i) can be measured with error, and this is accounted for within the modeling process, often by specifying an additive random discrepancy, typically denoted as ε_i [25]. Take for example a simple linear model equation:

$$Y_i = \beta_0 + \beta_1 X_i + \varepsilon_i, \quad (1)$$

where β_0 and β_1 correspond to the intercept and slope, and ε_i corresponds to model error which includes any potential error associated with measuring the response variable. When co-registration errors occur, this amounts to the introduction of error into the ability to measure X_i (e.g., spectral values) coincident with Y_i (e.g., basal area per hectare). Measurement error in X_i is not typically accounted for in regression models and will cause attenuation bias [25,26], which manifests in estimates trending towards the global mean of the response variable.

To circumvent the impacts of co-registration errors, analysts have employed a wide variety of solutions, ranging from rectifying images in a relative manner [27] to ignoring these errors and assuming them to be of little importance in predictions [28]. Regardless of the precision of measuring the true or relative surface location, spatial error will always be part of the rectification process and will have an impact on the underlying predictive model.

Within remote sensing literature, the impact of co-registration error has been recognized, especially for Light Detection And Ranging (LiDAR) data [29–32], but typically is not directly quantified. Often studies cite co-registration as an additional source of error that should be minimized, but fall short in describing the effects of those errors or providing suggestions to minimize the influence of those errors on predicted values. In this study, we address this knowledge gap by developing techniques to quantify this source of error and mitigate co-registration errors in applied work. Through simulation using Landsat 8 and National Agriculture Imaging Program (NAIP) imagery and images created with specific spatial correlation, based on Landsat 8 and NAIP images, we investigate co-registration errors and their impacts on the modeling process, and test the hypothesis that co-registration errors can be mitigated through spatial aggregation. Additionally, given estimates of global spatial continuity and co-registration errors, we provide recommendations on the size and layout of field observations with respect to the grain size of remotely-sensed data that will help to minimize the impact of co-registration errors.

2. Materials and Methods

2.1. Theoretical Background

The impact of co-registration errors on any predictive models should be related to four primary factors: (1) the horizontal misalignment between response and predictor variables, (2) the spatial extent of the sample unit, (3) the spatial correlation of predictor and response variables, and (4) the strength and form of the relationship between response and predictor variables. Prior to performing a study, researchers typically do not know the spatial correlation of response variables, nor the strength or form of the relationship between response and predictor variables. To address this lack of information in our study, remove issues of measurement error, and focus our study solely on the impacts of co-registration

error, we constrain our predictor surfaces to have a one-to-one relationship with our response variables. In this scenario, the Y and X surfaces, in the absence of co-registration errors, should exhibit a perfect linear relationship (i.e., an intercept of 0, a slope of 1, and a coefficient of determination of 1). Also, where the relationship between X and Y is linear, aggregated values (i.e., averages over multiple adjacent pixels) will also exhibit the same one-to-one relationship as non-aggregated values. Given this design, deviations from a one-to-one relationship can be solely attributed to co-registration errors.

Additionally, assuming that co-registration errors manifest as random noise within the regression models, we anticipate that the proportion of variation in Y explained by X (R^2) should follow the squared geometric relationship between the sample unit size (A_s) and the area of overlap (A_o) between the X and Y units, when the X values are distributed independently at random over space (Figure 1, Appendix A). This can be expressed as follows:

$$R^2 = \left(\frac{A_o}{A_s} \right)^2 \quad (2)$$

In concept, each sample unit's Y values are related to a combination of the corresponding X values now attached to an area only partially overlapping with the sample unit (cross-hatched area in Figure 1), as well as to X values attached to distinct spatial areas that have been falsely aligned with the sample unit. The latter occurs only because co-registration errors incorrectly identify a spatial match. If the X values are distributed independently at random over space, then on average the proportion of information on Y that can be explained by X should correspond to the average amount of area shared between response and predictor sample units, given the registration errors. Given this assumption, deviation from this condition in our simulations can be attributed to the spatial correlation within a landscape, and provide a rationale for using measures like global Moran's index (GMI) [33] as predictors, to estimate the proportion of modeling error contributed by co-registration errors.

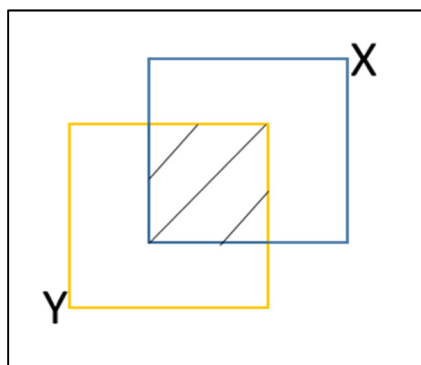


Figure 1. Graphical depiction of co-registration error. The global positioning system (GPS) (Y) and Image (X) sample units represent the same extents located on surface of the earth, but due to co-registration errors they only share a portion of the same area in projected space (diagonal black lines).

2.2. Overview

All analyses within this study were performed using R [34]. Images created with specified amounts of spatial correlation (virtual images) were built using the raster [35] and gstat [36,37] packages. Our simulations use one Landsat 8 [38] and five NAIP [39] images as baseline datasets taken from varying landscapes (Figure 2, Table 1), to produce nine virtual Landsat images and ten virtual NAIP images, respectively. To determine the amount of spatial correlation associated with the Landsat and NAIP baseline images, a uniform random selection of 20 locations were used to extract raster cell values within a 200 by 200 cell window, and to calculate empirical omnidirectional covariogram statistics [40].

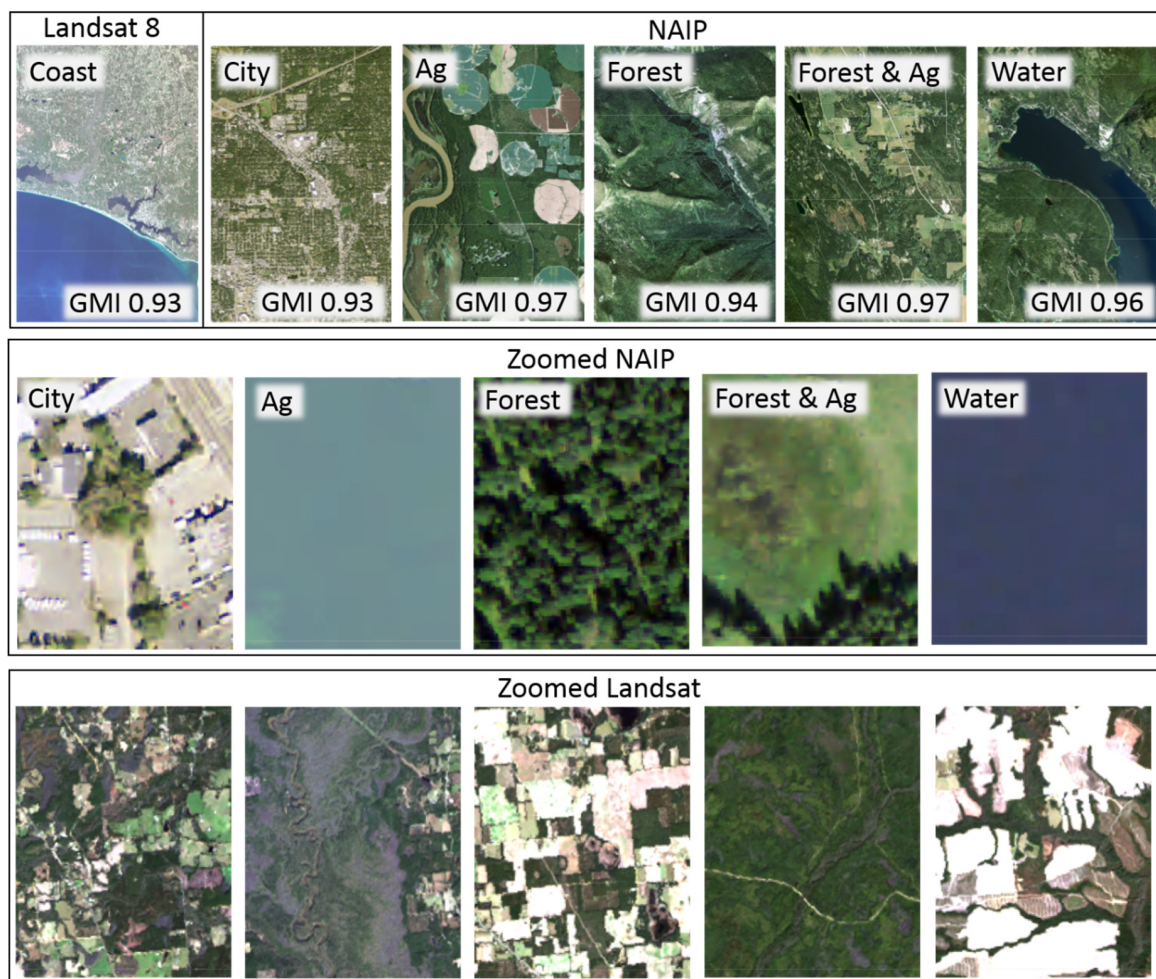


Figure 2. Base images used in simulations. Zoomed-in areas illustrate the extent for which images were subset and summarized to estimate mean digital number, sill, nugget, and range values.

Table 1. Average digital number (MDN), sill, nugget, and global Moran's index (GMI) and maximum range (in number of cells) values for Landsat and National Agriculture Imaging Program (NAIP) imagery. Averages and maximum values were based on all bands within an image.

Name	Label	MDN	Sill	Range	Nugget	GMI
Landsat 8 Coast	Coast	7478.1	552,790.7	40.5	180,321.3	0.93
NAIP City	City	140.2	1630.3	30.1	348.7	0.93
NAIP Agriculture	Ag	115.7	449.3	46.0	161.0	0.97
NAIP Forest	Forest	86.3	525.0	31.2	187.4	0.94
NAIP Forest & Agriculture	Forest & Ag	119.0	593.9	41.9	78.7	0.97
NAIP Forest & Water	Water	88.9	548.3	33.8	113.9	0.96

Cell values within each 200 by 200 cell window were summarized to estimate a mean digital number (DN) value, as well as sill, nugget, and range values for empirical omnidirectional covariograms. Mean DN, sill, and nugget statistics from each band were then averaged across image sources and used as inputs for creating Landsat- and NAIP-based virtual surfaces. To mimic different degrees of spatial correlation, range values were allowed to vary from 0.5 cells (completely random image) to the maximum range found among bands within each image source. Together, mean DN, sill, nugget, and ranges with a spherical spatial model were used to create virtual NAIP and Landsat surfaces (36,37). A complete listing of the code used to estimate spectral and spatial statistics and create virtual Landsat and NAIP images can be found in [Appendix B](#).

After creating each single band virtual images, two simulated sampling experiments were conducted, using actual and virtual Landsat 8 and NAIP images to evaluate the impacts of co-registration errors, spatial aggregation, sampling intensity, and spatial correlation on model prediction. The first set of simulations (stage I) were used to quantify the impacts of spatial aggregation of individual cells into multi-cell sampling units with regards to model prediction given co-registration error and defined spatial correlations (Figure 3). To account for potential logistical constraints of sampling large areas in the field, a second simulation was performed (stage II) that explored the impacts of alternative subsampling configurations corresponding to varying levels of measurement intensity and sample unit extent.

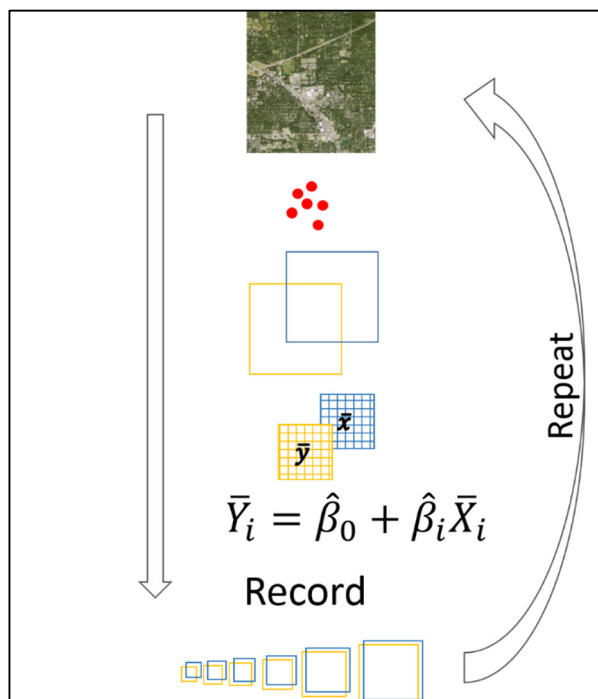


Figure 3. Visualization of Stage I simulations. A total of 200 sample locations (red points) were used to extract and calculate mean values from an image for different spatial extents around a point before and after a spatial shift was introduced (yellow and blue squares). Values were then regressed against one another to determine the impact of co-registration errors. This process was performed for each image used in the study.

Due to computational limitations associated with calculating range values for the extent of Landsat and NAIP imagery, we explored using GMI as a surrogate for range. GMI, while different than range, quantifies spatial correlation as an index value bounded between -1 (negative correlation) and 1 (positive correlation), with a value of zero corresponding to no spatial correlation (completely random image). For each band within each image of our simulations, GMI was calculated as follows:

$$\text{GMI} = \frac{N \sum_i \sum_j w_{ij} (x_i - \bar{x})(x_j - \bar{x})}{W \sum_i (x_i - \bar{x})^2}, \quad (3)$$

With x equal to values within a raster surface indexed by I and j rows and columns, w_{ij} representing a weighted spatial matrix (rook's case), N being the number of cells, and W being the sum of all weights. The remainder of this section describes in detail the design and implementation of each simulation stage within our study and model fitting used to estimate the impact of co-registration errors.

2.3. Stage I Simulations

Co-registration errors were mimicked based on published NAIP (6 m) [39], Landsat 8 (37 m) [41], and global positioning system (GPS; 7 m) [42] horizontal errors. For each image, 200 sample locations (L_1) were selected spatially at random and used to extract DN values for sequentially increasing spatial extents, with side lengths ranging from 1–100 cells (response). Mean cell DN values were calculated and recorded for the extent of the sample unit size at each L_1 location. L_1 locations were then randomly shifted based on co-registration errors between GPS locations and imagery to produce L_2 locations. Random shifts were implemented to the nearest cell by adding random distances and azimuths to easting and northing coordinates, based on a normal distribution with mean 0 and standard deviation expressed as the root mean squared error (RMSE) for each source of spatial error. Because Landsat 8 absolute geodetic accuracy is reported at the 90% confidence level, while NAIP imagery is reported at the 95% confidence level, we adjusted Landsat 8 horizontal error to the 95% confidence level. For Landsat 8, this transition amounts to an absolute error of 48 m (1.6 cells). The source code used to perform spatial shifts (function `shiftXY`), image value extractions (function `extractRC`), and mean calculations (function `getMeanBlockValue`) can be found in Appendix B (jLib.r).

L_2 locations follow the same DN extraction and summarization process as L_1 locations (predictor). Using response and predictor variables for each image, band, and sample unit size, we performed a simple linear regression using ordinary least squares and recorded RMSE (measured in units of mean DN value), as well as intercept, slope, and coefficient of variation (R^2) fit statistics. To minimize the effects of sampling variation, this procedure was performed 10 times, and regression results were averaged across all iterations. Additionally, for each image and band GMI was calculated. Regression fit statistics and coefficients were then compared across sample unit sizes and spatial correlation to identify and quantify the impact of co-registration errors and determine an array of suitable field sampling extents, to evaluate measurement intensity for Stage II of the simulations.

2.4. Stage II Simulations

Preferably, when relating remotely sensed data to field samples, the entire area within a sample unit would be measured on the ground. However, due to practical limitations related to collecting field data for sample units with large spatial extents, this is often not economically feasible. This situation can lead to instances when the only practical way to estimate a mean for a spatial extent is to use subsampling. To quantify the impact of six common subsampling (subplots) layouts and various subsampling intensities (area measured) within a given sample unit size (plot), we investigated multiple plot/subplot layouts. A depiction of plot extents, subplot layouts, and subsampling intensities are illustrated in Figure 4.

Identical to Stage I simulations, Stage II simulations mimic registration errors for 200 observations and extract cell values surrounding each plot location for each sample unit size. For the response variable, the mean values for each sample unit size are calculated based on the spatial extent of one of six subplot layouts, and subsampling intensities ranging from 0.05 to 0.95 of the plot extent, by increments of 0.05. Subplot layouts include one subplot located in the center of the plot (One), four subplots located systematically in the corners of the plot (Sys 4), four randomly located subplots within the plot (Rnd4), four subplots oriented in a similar fashion as the Forest Inventory Analysis (FIA) program plot protocol (FIA 4) [43], five subplots systematically placed within the plot extent (Sys 5), and nine plots systematically placed within the subplot (Sys 9). For predictor variables, mean values were calculated using all cell values within the extent of the plot (P_{all}), and for only the areas within the subplots (P_{sub}). Regression fit statistics and coefficients were then compared with results from 100% of the sample unit size measured in stage I.

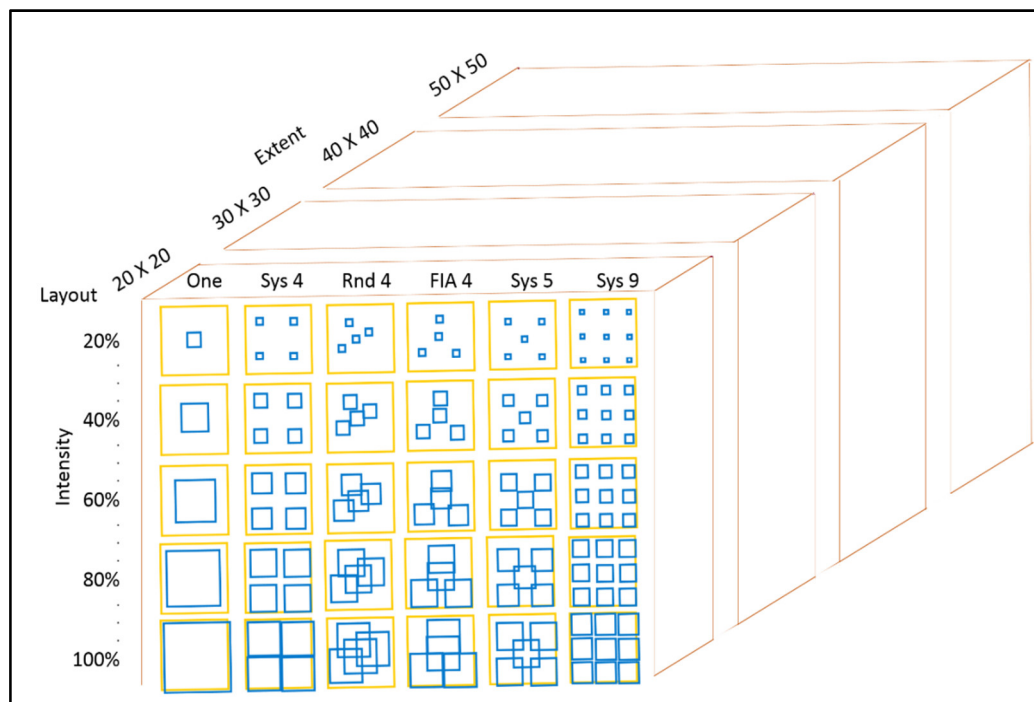


Figure 4. Subsampling layout and subsampling intensity for varying sample unit sizes. Yellow square boxes define the spatial extent sampled within an image, while the shifted blue boxes illustrate the impact of co-registration errors and define a subsampling layout and proportion of area measured within each yellow extent. Large brown cubes denote iterations for potential sample unit sizes.

2.5. Modelling the Impacts of Co-Registration Error

After performing each simulation and recording error and fit statistics for each image, we developed a suite of models to relate those statistics to predictors measuring spatial correlation in the images (GMI) and the magnitude of spatial co-registration errors (expected proportion of area overlapped between field plots and corresponding image locations). While the overlap between two rectangles can be calculated if both the distance and direction of co-registration errors are known, the direction of co-registration errors is seldom calculated or reported. Therefore, within our iterations we estimated the expected proportion of overlap (PO) for each sample unit size, given the offsets used to simulate co-registration errors.

For virtual images with no spatial correlation, we hypothesized a one-to-one relationship between PO^2 and the proportion of variation explained (R^2). However, as spatial correlation increases within images, we anticipate that the ratio between R^2 and PO^2 will be greater than one and will interact with spatial correlation metrics. Additionally, we recognize that PO^2 would be difficult to calculate in practice given commonly reported horizontal rectifications. Therefore, when modeling the impact of co-registration errors on R^2 in the presence of spatial correlation, we used only sample unit size and GMI as predictors and beta regression with a logit link [44]. Similar in concept to logistic regression, beta regression was developed to work with observations between zero and one, and is typically used to characterize natural rates or proportions on a continuous scale. Using a logit link, our proposed model takes the following form:

$$\ln\left(\frac{R^2}{1-R^2}\right) = \beta_0 + \beta_1 f(|sample\ unit|) + \beta_2 g(GMI) + \beta_3 f(|sample\ unit|) * g(GMI) \quad (4)$$

where $f()$ and $g()$ are known transformations of the sample unit size and image GMI, respectively, and the β_k are parameters estimated from the data. Transformations of predictor variables were determined based on graphical analyses. While we anticipated needing sample unit size, GMI,

and their interaction to estimate R^2 , we also evaluated nested models using only sample unit size and GMI. All beta regression models were compared using Akaike's information criterion (AIC) [45,46].

3. Results

3.1. Datasets

Estimated mean DN, sill, and nugget and maximum range values varied by image (Table 1). While most of these characteristics varied substantially by data source due to pixel depth (Landsat 16-bit pixel depth versus NAIP 8-bit pixel depth), range values, measured in cells, were quite similar. Using the average DN, sill, and nugget and maximum range values of each data source, we created nine virtual Landsat images and ten virtual NAIP images of varying spatial correlation (Figure 5). It should be noted that virtual image GMI values were less than actual image GMI values, suggesting that there was less positive spatial correlation in the virtual images than in the actual images. However, the range of simulated autocorrelations produced virtual images with a variety of spatial structures and aggregated patterns that closely resembled patterns found within homogenous patches of actual images (Figure 2 zoomed-in examples and Figure 5). Generally, the boundaries between patches representing different DN values within the virtual images were not as sharp when compared to the base images. However, the patterns created in the virtual images provide an objective way to evaluate varying levels of spatial correlation.

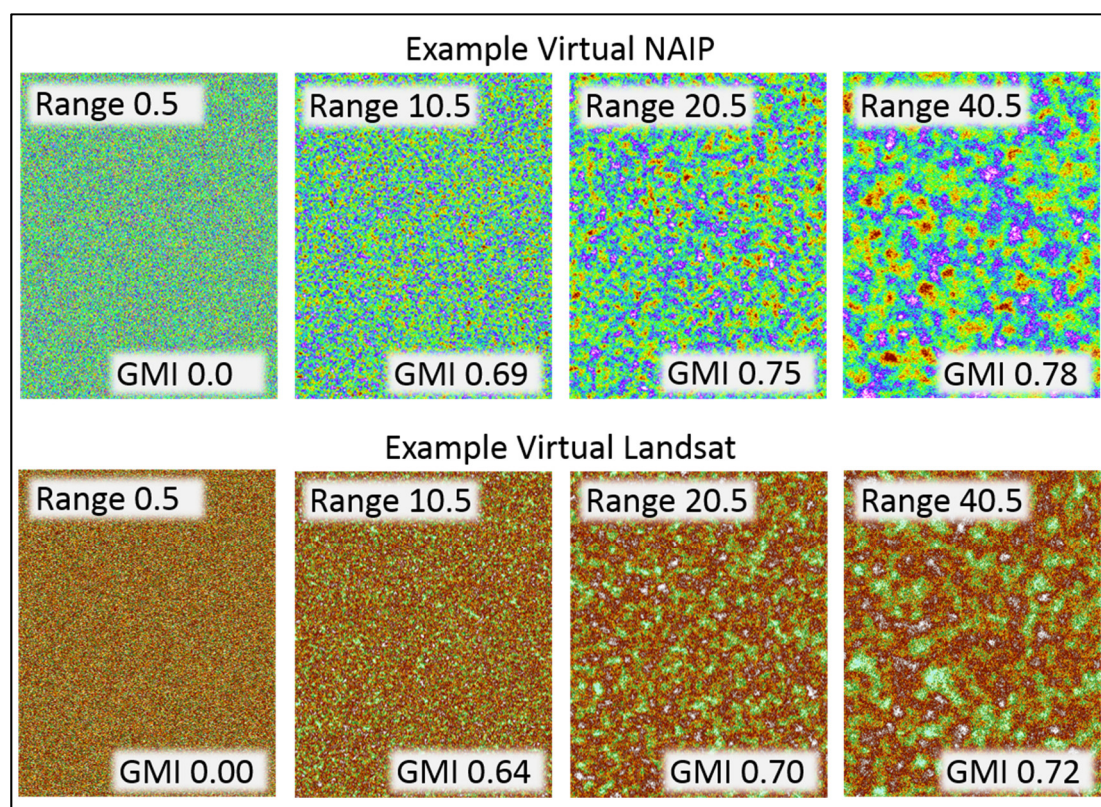


Figure 5. Subset of images created from average digital number (DN), sill, and nugget and maximum range values derived from NAIP and Landsat 8 imagery.

3.2. Stage I

Comparisons in stage I indicate that increasing the spatial footprint of a sample unit can mitigate the effects of co-registration errors on predictive models. On average, horizontal shifts between L_1 and L_2 locations were 1.6 and 7.8 cells for Landsat 8- and NAIP-based images, respectively. For all images and bands analyzed, the extent of the sample unit was strongly related to the magnitude of deviation from the anticipated one-to-one regression relationship (intercept, slope, and R^2 equal to 0, 1, and 1, respectively). Linear models derived from raster datasets with large spatial correlation, in terms of range or GMI, produced slope and intercept estimates closer to 1 and 0, respectively (less attenuation), than raster datasets, with less spatial correlation for both Landsat- and NAIP-based datasets (Figures 6 and 7). While larger sample unit sizes reduced attenuation bias, for spatial correlation ranges above 30 cells, sample unit sizes greater than approximately 9 and 40 cells for Landsat- and NAIP-based images, respectively, appear to produce only marginal reductions in parameter bias or improvements in R^2 . For NAIP-based imagery, this suggests that a field plot with an extent as large as 40 m by 40 m might be required to mitigate the effects of co-registration errors between NAIP imagery and GPS locations. Similarly, for Landsat-based images a field plot with an extent as large as 270 m by 270 m may be required to mitigate model error introduced by co-registration error.

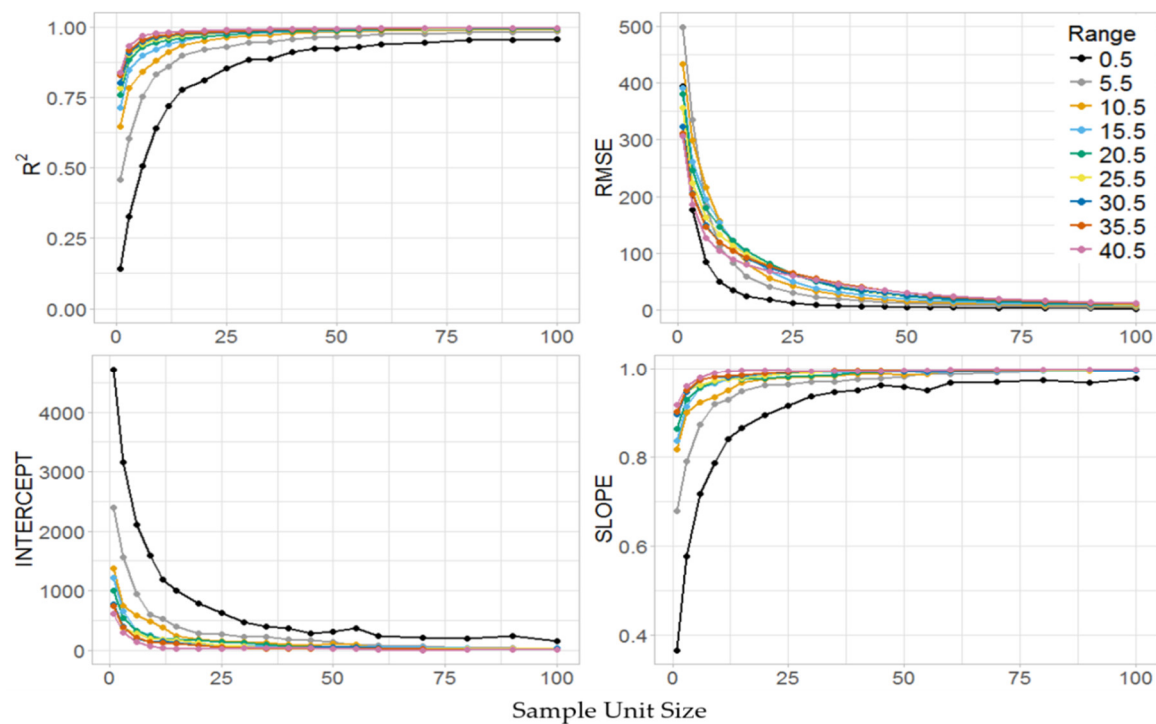


Figure 6. Stage I Virtual Landsat image regression statistics for varying sample unit sizes and spatial correlation (Range), given an average image registration error of 1.6 cells and an average GPS navigational unit error of 0.23 cells. Actual Landsat image regression statistics are shown in Appendix C (Figure A1).

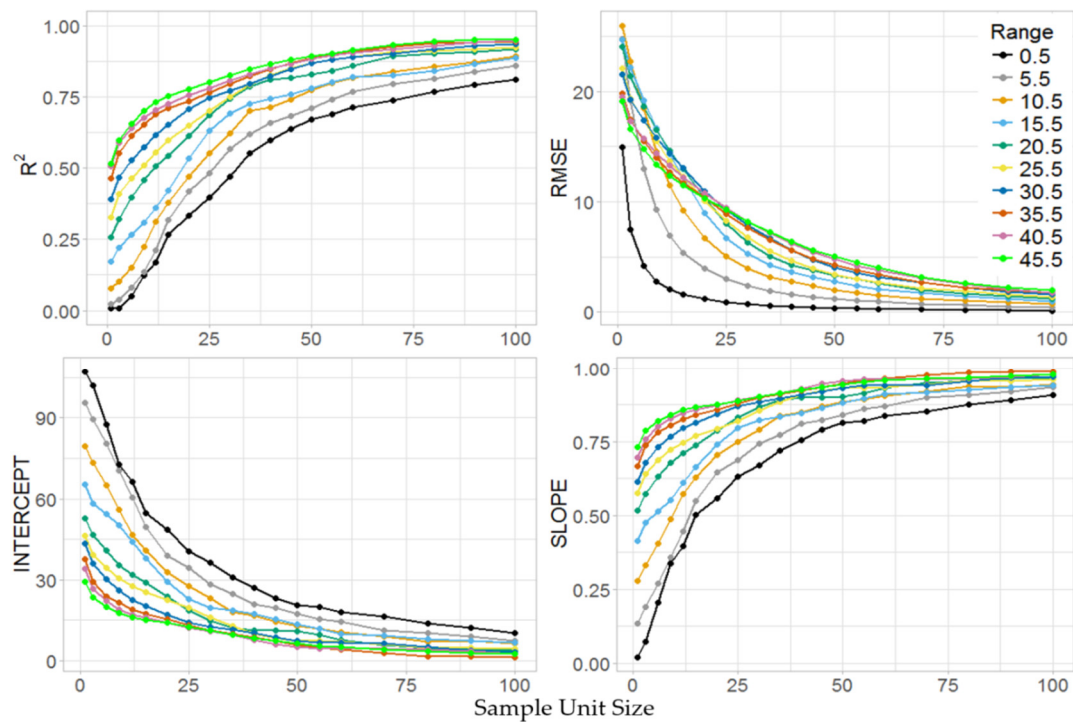


Figure 7. Stage I virtual NAIP image regression statistics for varying sample unit sizes and spatial correlation (Range), given average raster and GPS registration errors of 6 and 7 cells respectively. Actual NAIP image regression statistics are shown in Appendix C (Figure A2).

3.3. Stage II

Comparisons in stage II had similar trends as found in stage I, and verify that subsampling intensity and layout also impacted the amount of variation explained by models in the P_{all} subsampling scenario. For sample unit sizes between 20 and 50 cells and 5 and 20 cells for NAIP- and Landsat-based imagery, respectively, larger proportions of the area subsampled within a sample unit consistently explained more variation within the data, and produced smaller RMSE across all levels of spatial correlation and data sources. After the proportion of area subsampled reached approximately 80% of the plot extent (P_{sub}), R^2 appeared to differ only marginally relative to the R^2 associated with P_{all} (Figure 8). This was also the case for RMSE. Across all subsampling intensities and sample unit sizes, the worst-performing subplot layouts were Rnd 4, FIA 4, and Sys 5. Subplot layouts One, Sys 4, and Sys 9 produced similar results, especially when the proportion of area measured within the plot extent was greater than 75%. As expected, P_{sub} generally produced better results than P_{all} , given that the response and predictor variables shared the same spatial configurations. However, there was little difference between P_{sub} and P_{all} subsampling techniques when greater than 80% of the plot extent was measured. As one might expect, smaller subsampling intensities (<20% of the sample unit extent) substantially reduce R^2 in our linear models. In some cases, when subsampling intensity and spatial correlation was small, the reduction in R^2 , compared to measuring all the area within a plot extent, was greater than 60%. However, for actual Landsat 8 and NAIP images, which have relatively high levels of spatial correlation, the reduction in variation ranged from approximately 0.4% to 30%, depending on the data source, subsampling intensity, spatial correlation, and co-registration errors (Appendix C, Figures A3 and A4). Similar to stage I simulations, increased amounts of spatial correlation generally dampened the negative effects of co-registration errors in stage II simulations. Additionally, this same dampening effect carried over to subsampling intensities when estimating means for all cells within a sample unit extent of a predictor variable.

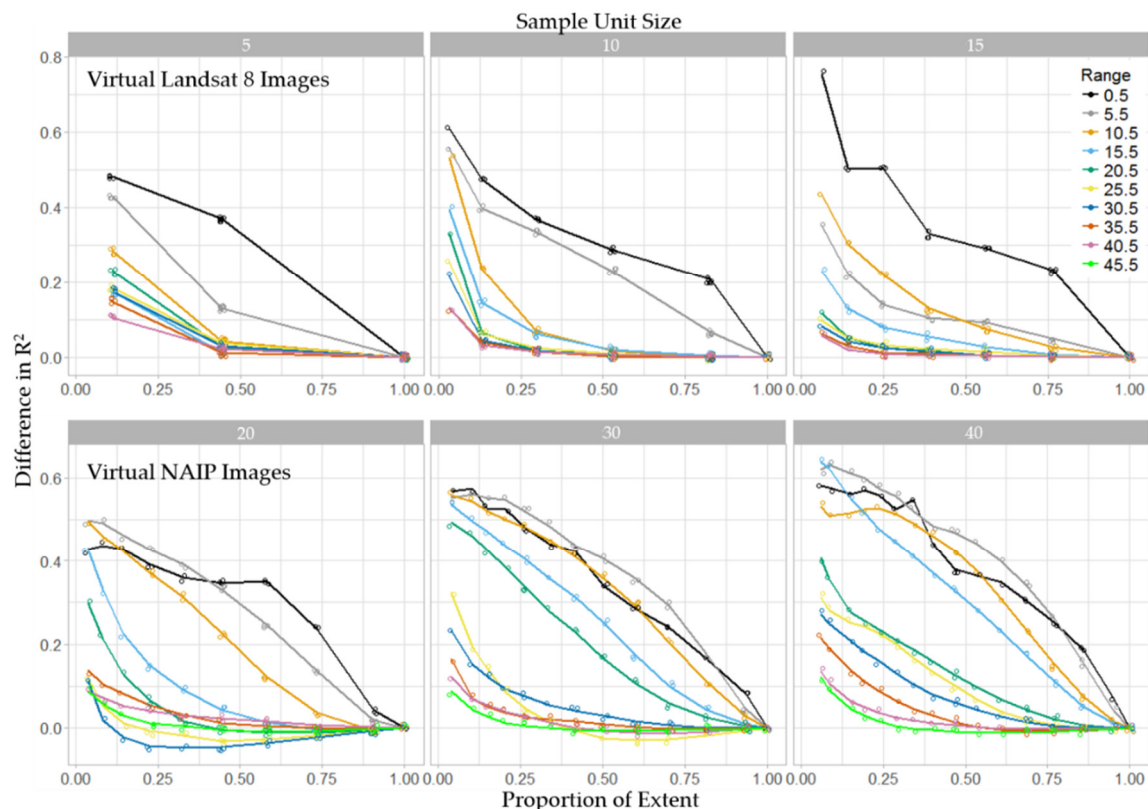


Figure 8. Reduction in the proportion of variation explained (R^2) for Landsat 8 and NAIP virtual images by subsample intensity (Proportion of Extent), sample unit size (in cells), and spatial correlation (Range) using SYS 4 and P_{all} subsampling layout. Figures A3 and A4 in Appendix C show actual Landsat 8 and NAIP image reductions.

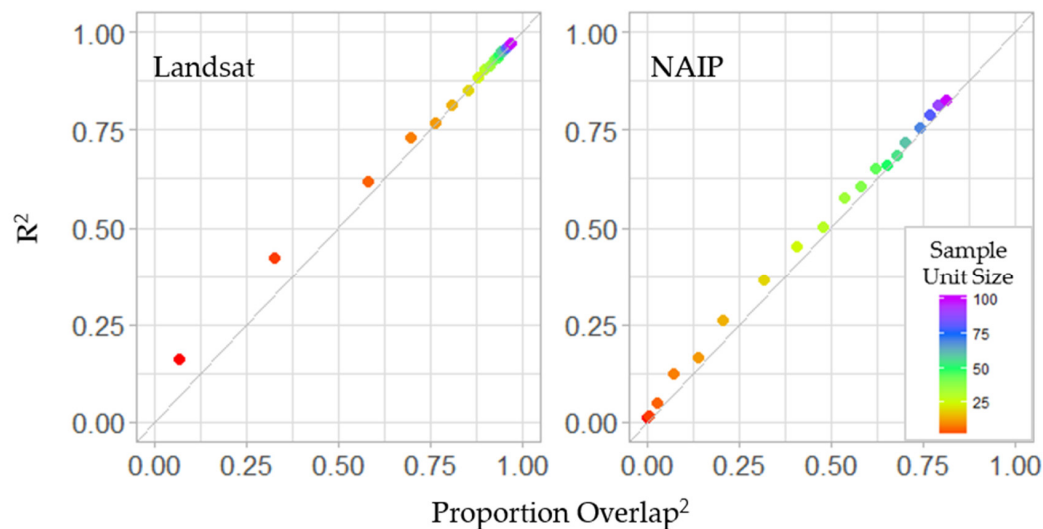
3.4. Model Fitting

Regressed mean DN values for L_1 and L_2 locations in both simulations indicate that co-registration errors can have substantial impacts on model fit, and can bias DN estimates. Globally, across the extent of each image, estimates of mean DN were necessarily unbiased. However, local estimates tended to over- or under-estimate DN values that were respectively smaller or larger than the mean (attenuation). The degree of attenuation in our models, identified by deviations from theoretical intercept and slope, was strongly related to both the spatial extent of an observation (sample unit size) and the spatial correlation of predictor variables (Figures 6 and 7).

For completely independent virtual images, the amounts of variation explained in our linear models were closely related to PO^2 (Table 2, Figure 9). For both Landsat 8- and NAIP-based imagery with average co-registration errors of 1.6 and 7.8 cells, respectively, R^2 and PO^2 closely followed a one-to-one ratio. For images with spatial correlation, exploratory analysis revealed that sample unit size and GMI did not appear to be linearly related to the logit of R^2 . However, the natural log of sample unit size (LSS) and the exponentiation of GMI (EGMI) did appear to be linearly related to R^2 . Therefore, we included LSS and EGMI in our suite of models for comparison (Table 3). Our top fitting models were statistically significant (p -value < 0.001), and included LSS, GMI, and the interaction between LSS and GMI for both Landsat 8- and NAIP-based images (Table 4). RMSE values for top-fitting Landsat 8 and NAIP models were 0.019 and 0.089, respectively (expressed on the scale of R^2). Regression diagnostics of our top-fitting models are shown in Appendix A, Figure A5. Untransformed, observed versus predicted R^2 followed a one-to-one relationship for both Landsat- and NAIP-based imagery (Figure 10), and the latter was constrained to fall between 0 and 1, with more variation occurring within the middle portion of the observed domain, as expected.

Table 2. Linear regression statistics for 19 independently random images, given the average proportion of overlap (PO) determined by sample unit size and simulated co-registration errors.

Model	Equation	Slope	R^2	RSE	F-Stat	P-Value
Landsat 8	$R^2 = 0 + PO^2$	1.008	0.9984	0.03456	11440	<0.001
NAIP	$R^2 = 0 + PO^2$	1.035 *	0.9983	0.02364	10330	<0.001

* Statistically different than one at $\alpha = 0.01$.**Figure 9.** Scatter plot of proportion of variation explained (R^2) versus the squared proportion of overlap between L_1 and L_2 locations, given various sample unit sizes, independent random surfaces, and Landsat 8 and NAIP horizontal registration errors. The gray diagonal line is a one-to-one line, for the purpose of comparison.**Table 3.** Suite of potential models and their associated AIC and Δ AIC values. Interaction term denoted by * specifies a full interaction model.

Model	Rank	Source	Predictors	AIC	Δ AIC
1	6	Landsat 8	<i>sample unit size</i>	−1723.823	−651.018
2	4	Landsat 8	<i>sample unit size + GMI</i>	−1920.781	−454.06
3	3	Landsat 8	<i>sample unit size * GMI</i>	−1938.580	−436.261
4	5	Landsat 8	$\ln(\text{sample unit size})$	−1870.267	−504.574
5	2	Landsat 8	$\ln(\text{sample unit size}) + e^{GMI}$	−2368.824	−6.017
6	1	Landsat 8	$\ln(\text{sample unit size}) * e^{GMI}$	−2374.841	0
1	6	NAIP	<i>sample unit size</i>	−1086.606	−681.595
2	5	NAIP	<i>sample unit size + GMI</i>	−1469.469	−298.732
3	3	NAIP	<i>sample unit size * GMI</i>	−1494.508	−273.693
4	4	NAIP	$\ln(\text{sample unit size})$	−1193.989	−574.212
5	2	NAIP	$\ln(\text{sample unit size}) + e^{GMI}$	−1728.674	−39.527
6	1	NAIP	$\ln(\text{sample unit size}) * e^{GMI}$	−1768.201	0

Table 4. Beta regression coefficients and statistics for top fitting Landsat 8 and NAIP based images given natural log sample unit size (LSS), exponent of global Moran's index (EGMI), interaction between LSS and EGMI, and simulated average co-registration errors.

Model	N	Intercept ⁺	LSS ⁺	EGIM ⁺	EGMI * LSS ⁺	Pseudo R^2 [45]	P-Value
Landsat 8	304	−3.743	1.089	2.423	−0.085	0.918	<0.001
NAIP	570	−9.364	1.883	3.481	−0.419	0.8255	<0.001

⁺ Statistically different from zero at $\alpha = 0.01$.

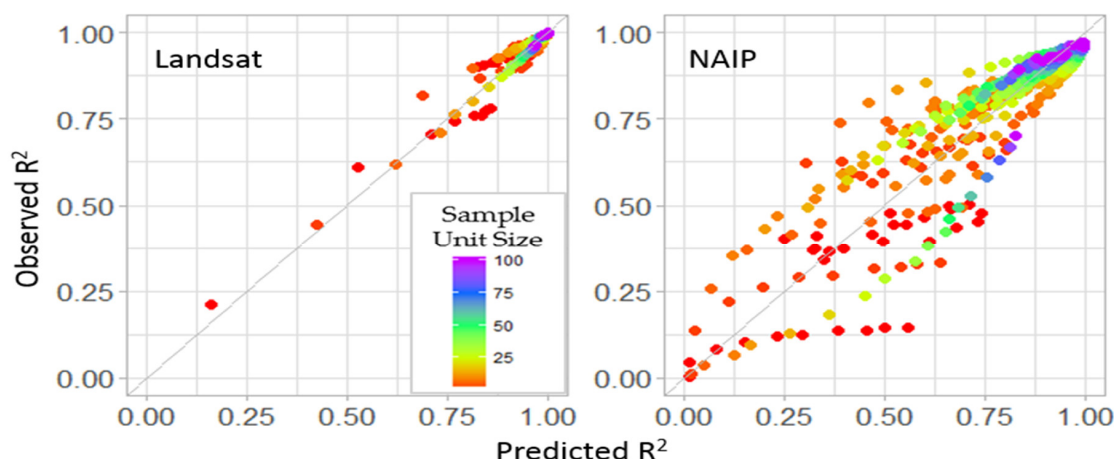


Figure 10. Observed versus predicted proportion of variance explained (R^2) for co-registration errors associated with Landsat 8 and NAIP imagery and virtual imagery, given various sample unit sizes, measures of spatial correlation, and top-fitting beta regression models. The gray diagonal line is a one-to-one reference line.

4. Discussion

Through our simulations, we have documented that spatial co-registration errors produce attenuation bias in linear models (Figures 6 and 7). For studies that relate field data located using GPS to geo- or ortho-rectified remotely sensed data, this bias will manifest in regression coefficients biased toward 0 and regression estimates trending towards the sampled mean value of the variable of interest (Appendix C, Table A1). While every attempt to minimize the amount of co-registration error should be taken, technical and financial limitations often make it impractical to completely remove this source of error. Due to these limitations, we explored the impacts of spatial aggregation of observational units on model performance when predictor variables have spatial co-registration errors.

Our findings demonstrate that increasing the spatial extent of sample units can help to reduce the impacts of imperfect co-registration. This result further verifies that larger field plots can mitigate the effects of co-registration error found by others [29,30,47,48]. However, when choosing the extent of a field sample unit, one must take into consideration practical issues associated with the costs of implementation and measurement, as well as the fact that large field sampling units can have a smoothing effect on spatial variability [29]. Moreover, subsampling within the extent of a field plot, regardless of the subplot layout, introduces additional variability into the predictive models, and should be used sparingly when spatially relating field measurements to remotely sensed information.

Given the sample unit sizes, co-registration errors, and spatial correlation we investigated, we recommend selecting a field plot extent large enough to substantially reduce bias in linear regression, while also keeping the extent of the field plot as small as possible to retain spatial detail. In the case of NAIP imagery, this recommendation would correspond to a field plot with an area between 400 m² and 1600 m². For Landsat 8 imagery, this recommendation corresponds to a field plot with an area between 8100 m² and 72,900 m². Fortunately, most NAIP and Landsat 8 images have a large degree of spatial correlation, suggesting that the lower end of these recommendations may suffice in mitigating the impacts of co-registration errors. For other sources of remotely sensed information that have different co-registration errors, simulations similar to those presented in this study should be completed to help determine suitable field plot extents and sampling intensities.

If subplots are used to estimate mean values within the extent of a sample unit, it is important that the subplot layout covers as much of the area within the extent of the sample unit as possible. For NAIP imagery, we recommend measuring 75% or more of the sample unit area to minimize the negative effects of subsampling. When it is too costly to measure 75% of the area within a sample unit, a tradeoff between cost and precision must be made. In this situation, collecting more sample units with

less than 75% of the subsample area measured can help to offset the losses in precision associated with subsampling (Figure 9). Additionally, when subsampling is used, layouts should be chosen such that there is no overlap among subplots, such as layout Sys 4 from our study (e.g., Figure 11).

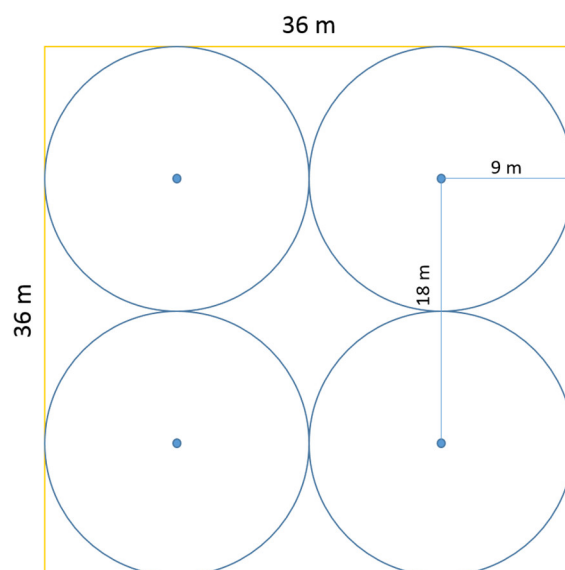


Figure 11. Example of a recommended field plot size and layout for NAIP imagery.

The actual extent of a sampling unit should depend on the amount of co-registration error, the spatial correlation within the imagery, and the amount of model error one is willing to accept. For readily available Landsat and NAIP imagery, their reported horizontal accuracies, and their estimated spatial correlations, we can estimate the co-registration error-induced reduction in variation explained by linear regression for various sample unit sizes (Equation (4) and Table 4). From these estimates, one can select a sample unit extent that both reduces estimation bias and quantifies error in predictor variables due to co-registration. For example, if a project was to use NAIP imagery with a sample unit size of 20 cells (field plot extent of 400 m²) and an estimated GMI of 0.92, then one would expect the logit of R^2 to be approximately 1.744, and the loss in predictive ability associated with co-registration errors to be $1 - R^2 = 0.149$.

While Equation (4) and the coefficients from Table 4 can be used to help guide the size of a field plot needed to mitigate the negative impacts of co-registration (Appendix C, Table A2), they should be interpreted as a best-case scenario. Specifically, our simulations were developed under the premise that there was a perfect one-to-one relationship between response and predictor variables. In many applications this will not be the case, and co-registration errors will be coupled with model error. To decouple co-registration errors from model errors, model coefficients can be dis-attenuated [26,49]. Within that context, simulations similar to the ones performed in our study, which use a random sample of the predictor variables and regress those values against shifted locations, can be used to estimate a ratio adjustment factor for model coefficients, as described by Forest and Thompson [49]. Appendix B provides examples of R coding that can be used to simulate co-registration errors and determine ratio adjustment factors.

Fortunately, most remotely-sensed images have relatively high levels of spatial correlation, which in turn dampens the impacts of co-registration errors. In our study, we evaluated the effects of co-registration on model error for levels of spatial correlation that spanned independent random landscapes, to those commonly found in terrestrial environments. For all actual landscapes used in our study, the minimum spatial correlation found had a GMI value of 0.93. Interestingly, virtual images with ranges comparable to actual image ranges had corresponding GMI values that were substantially less than those found in the actual images. This is likely due to the dramatic transitions found between

land use and cover types that can occur within actual landscapes (e.g., a forest bounded by grass lands). This further suggests that natural landscapes have more localized spatial correlation than our virtual landscapes, and that edges between land use and cover types constitute a substantial amount of the overall area within an image. Because these edge areas can make up a substantial component of the landscape, it is important that they are included in future investigations that use simulated landscapes, and more importantly, model training. Mapping endeavors that omit these transition areas from training sets do so at the cost of extrapolating model results to potentially large portions of an image.

5. Conclusions

In this study, we looked at the impacts of co-registration errors on model prediction. We found that increasing field plot size helps to mitigate the negative impacts of co-registration errors by reducing attenuation bias. Additionally, we identified that increased positive spatial correlation within imagery reduces the negative impacts of co-registration for a given sample unit size. Finally, we presented a simulation methodology that can be easily applied to remotely sensed data that both quantifies the impact of co-registration on model prediction and can be used to estimate measurement error in predictor variables. Using our plot size recommendation and components of the simulation techniques described, estimation bias can be mitigated, which in turn should help managers to precisely define the complex spatial relationships needed to promote spatially informed decision making.

Author Contributions: Conceptualization, J.H. and D.L.R.A.; data curation, J.H.; formal analysis, J.H.; investigation, J.H. and D.L.R.A.; methodology, J.H. and D.L.R.A.; software, J.H.; supervision, D.L.R.A.; validation, J.H.; visualization, J.H.; writing—original draft, J.H.; writing—review and editing, J.H. and D.L.R.A.

Funding: This research was funded by the Gulf Coast Ecosystem Restoration Council (RESTORE Council) through an interagency agreement with the U.S. Department of Agriculture (USDA) Forest Service (17-IA-11083150-001) for the Apalachicola Tate's Hell Strategy, and by the U.S. Forest Service, Southern Region, and Rocky Mountain Research Station.

Acknowledgments: We would like to thank the independent reviewers for their comments and suggestions. Their advice has helped to improve the quality of this manuscript.

Conflicts of Interest: The authors declare no conflict of interest

Appendix A

The relationship between the proportion of overlapping area between two square sampling units offset by a specified direction and distance (PO) and Pearson's correlation.

Let $X(p)$ be the DN value of pixel p in a random raster and $X(p) = \mu + e(p)$ where $e(p)$ are offsets from the mean DN μ with expected value 0 and variance σ_e^2 . Then let $X(b)$ be the mean DN value of a block b of pixels. Denote the size (in pixels) of b by $|b|$. Then

$$X(b) = |b|^{-1} \sum_{p \in b} X(p) = \mu + |b|^{-1} \sum_{p \in b} e(p) \quad (5)$$

For blocks selected uniformly at random, the expected value of $X(b)$ is μ and its variance is $|b|^{-1} \sigma_e^2$ provided that $e(p)$ are uncorrelated (but $\text{var}[X(b)] > |b|^{-1} \sigma_e^2$ if there is positive spatial autocorrelation among the $e(p)$). The covariance between any $X(b)$ and $X(b')$ is

$$\text{cov}[X(b), X(b')] = E[e(b)e(b')] \quad (6)$$

$$= \frac{1}{|b|^2} E \left[\left(\sum_{p \in b} e(p) \right) \left(\sum_{p \in b'} e(p) \right) \right] \quad (7)$$

$$\geq \frac{1}{|b|^2} E \left[\left(\sum_{p \in (b \cap b')} e(p)^2 \right) \right] = \frac{|b \cap b'|}{|b|^2} \sigma_e^2 \quad (8)$$

with equality holding only if the $e(p)$ are spatially uncorrelated.

Co-registration error can be simulated by shifting the original raster X by some random amount, resulting in a shifted raster Y where $Y(p) = X(p')$ and thus $Y(b) = X(b')$. If the $e(p)$ are spatially uncorrelated, then

$$\text{cov}[X(b), X(b')] = \text{cov}[X(b'), X(b)] = \frac{|b \cap b'|}{|b|^2} \sigma_e^2. \quad (9)$$

Furthermore,

$$\text{corr}[Y(b), X(b)] = \frac{\text{cov}[X(b'), X(b)]}{\sqrt{\text{var}[X(b')] \text{var}[X(b)]}} = \frac{|b \cap b'| |b|^{-2} \sigma_e^2}{|b|^{-1} \sigma_e^2} = \frac{|b \cap b'|}{|b|} \quad (10)$$

where the last quantity is the proportion of the original block b that is overlapped by the shifted block b' . As a result, the coefficient of determination (R^2) obtained by regressing $Y(b)$ on $X(b)$ will be directly related to the proportion of block overlap:

$$R^2 = (\text{corr}[Y(b), X(b)])^2 = \left(\frac{|b \cap b'|}{|b|} \right)^2 \quad (11)$$

Appendix B

R code developed to perform all analyses and simulations within the study.



Appendix C

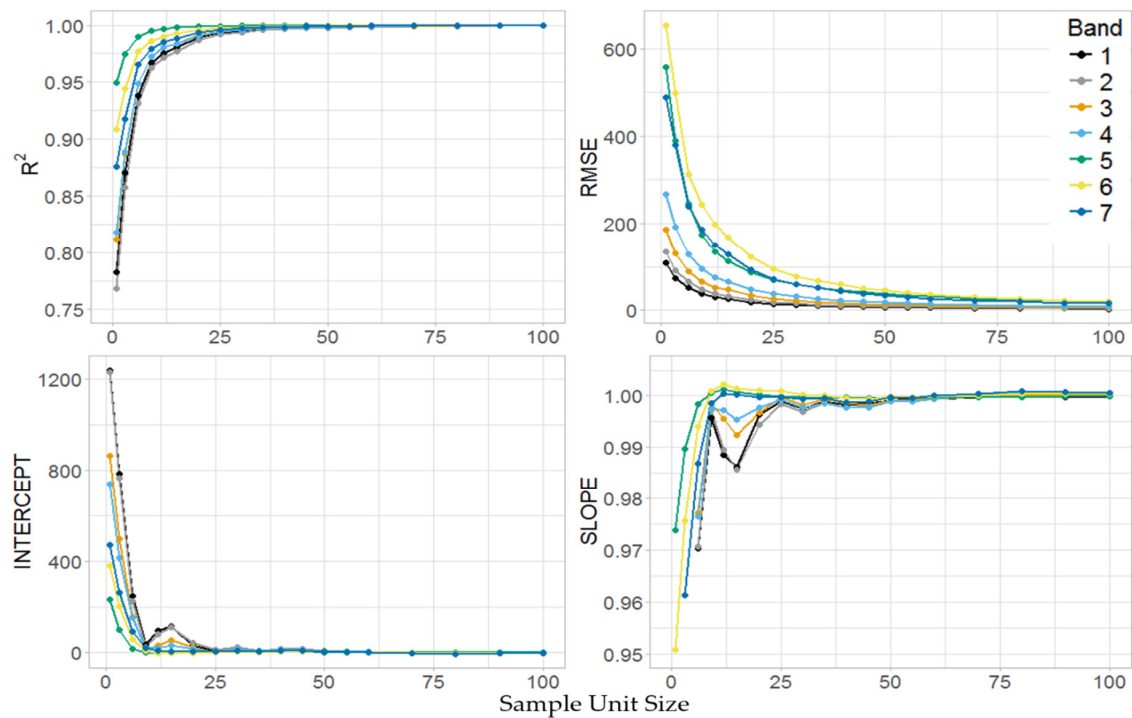


Figure A1. Stage I Landsat image regression statistics for varying bands and sample unit size lengths given an average image registration error of 1.6 cells and an average GPS navigational unit error of 0.23 cells.

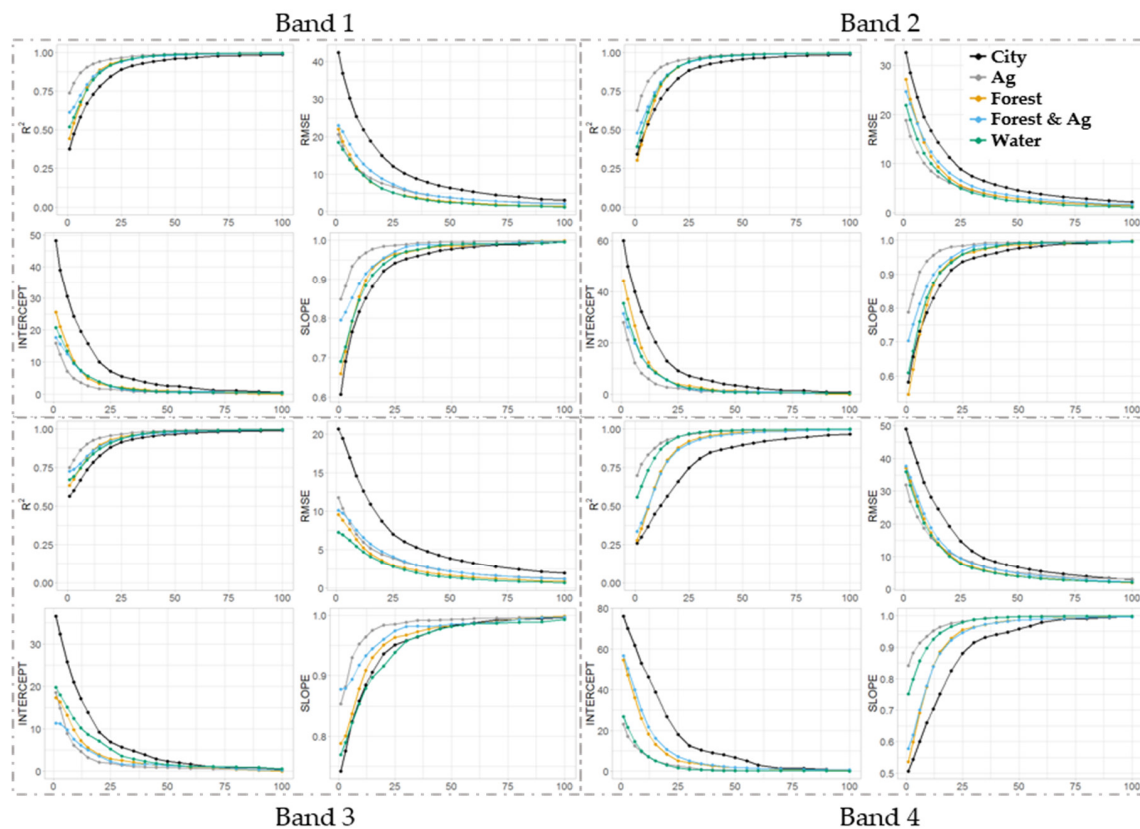


Figure A2. Stage I NAIP image regression statistics for varying sample unit sizes and spectral bands given average raster and GPS registration errors of 6 and 7 cells respectively.

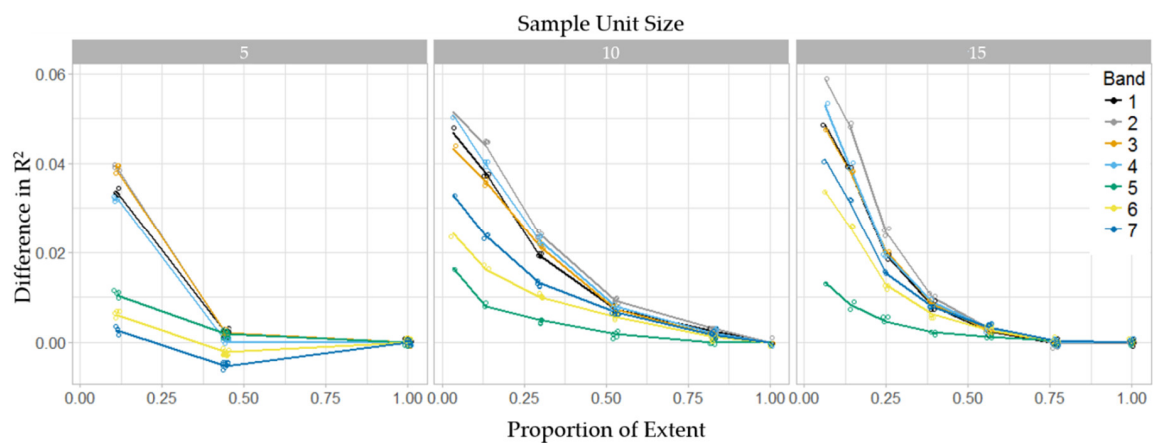


Figure A3. Reduction in the proportion of variation explained (R^2) for Landsat 8 image by subsample intensity (Proportion of Extent), sample unit size, and spectral band using SYS 4 and P_{all} subsampling layout.

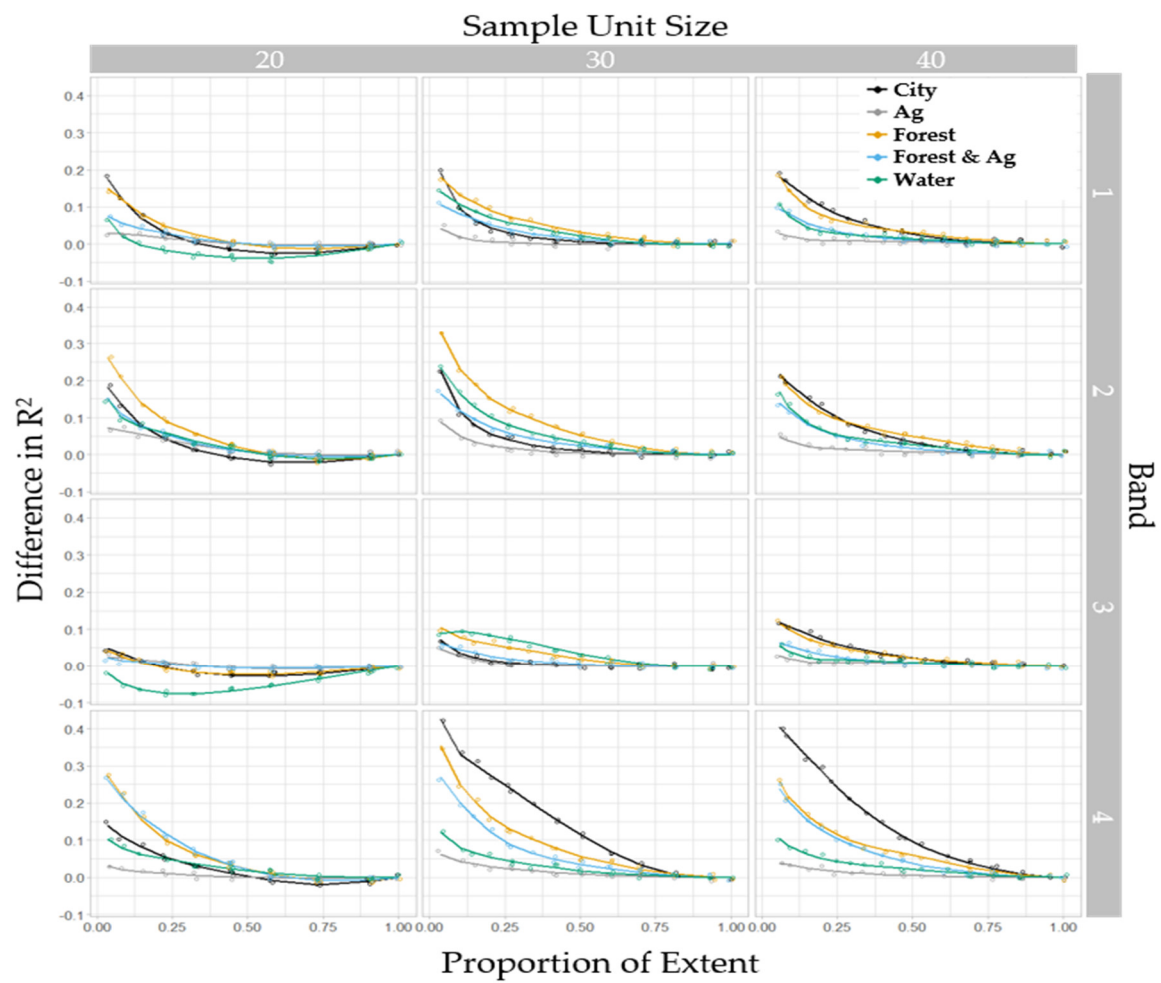


Figure A4. Reduction in the proportion of variation explained (R^2) for NAIP images by subsample intensity (Proportion of Extent), sample unit size and spectral band using SYS 4 and P_{all} subsampling layout.

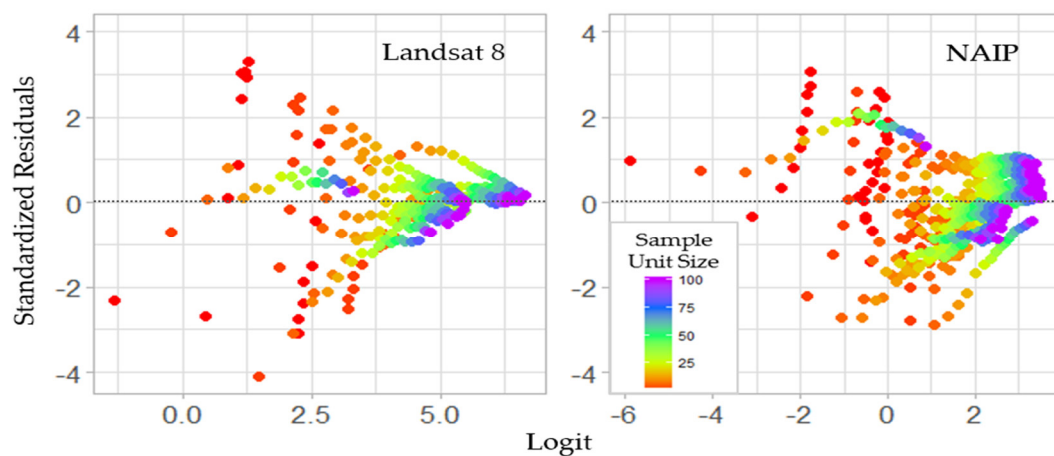


Figure A5. Scatter plot of standardized Residuals vs linear predictors (Logit) given block size for beta regression Landsat 8 and NAIP based models.

Table A1. Examples of remote sensing applications impacted by co-registration errors and the impact on model fit and estimates.

Application	Data Source	Impact on Model Fit and Estimates
Mapping forest characteristics using models derived from field data and remotely sensed data.	Landsat imagery	Attenuated estimates and reduction in model fit. The amount depends on the spatial correlation within the imagery, the co-registration error between imagery and field data, and the spatial extent of the field data observations (Table A2).
	NAIP imagery	Attenuated estimates and reduction in model fit. The amount depends on the spatial correlation within the imagery, the co-registration error between imagery and field data, and the spatial extent of the field data observations (Table A2).
	Other remotely sensed data.	Attenuated estimates and reduction in model fit. The amount depends on the spatial correlation within the imagery, the co-registration error between imagery and field data, and the spatial extent of the field data observations.
Change detection derived from multiple images of a given area.	Satellite and aerial based imagery	Attenuated estimates and reduction in model fit.
Image radiometric normalization	Satellite and aerial based imagery	Attenuated estimates and reduction in model fit.
Image segmentation	Attenuated outputs	Less variation in estimated values potentially reducing the accuracy of the segmentation process.
Practitioner use of attenuated spatial data products derived from field plots and remotely sensed imagery.	Attenuated outputs	Mean estimates derived from the entire surface will not be bias. Subsets of the derived surface will be biased and will either over estimate (values < mean) or under estimate (values > mean) the true values.

Table A2. Estimated reduction in R^2 (ΔR^2) for Landsat 8 and NAIP imagery given Equation (4), sample unit size, GMI value, and published horizontal image and GPS errors.

Source	Sample Unit Size (Cells Wide)	GMI	ΔR^2
Landsat 8	3	0.8	0.067
Landsat 8	5	0.8	0.043
Landsat 8	9	0.8	0.026
Landsat 8	3	0.9	0.040
Landsat 8	5	0.9	0.026
Landsat 8	9	0.9	0.016
Landsat 8	3	0.95	0.030
Landsat 8	5	0.95	0.019
Landsat 8	9	0.95	0.012
NAIP	20	0.8	0.226
NAIP	30	0.8	0.166
NAIP	40	0.8	0.131
NAIP	20	0.9	0.148
NAIP	30	0.9	0.109
NAIP	40	0.9	0.088
NAIP	20	0.95	0.116
NAIP	30	0.95	0.087
NAIP	40	0.95	0.070

References

1. Jensen, J. *Remote Sensing of the Environment: An Earth Resource Perspective*; Prentice Hall: Upper Saddle River, NJ, USA, 2000; p. 544.
2. Bechtold, W.A.; Patterson, P.L. (Eds.) *The Enhanced Forest Inventory and Analysis Program—National Sampling Design and Estimation Procedures*; General Technical Report, SRS-80; Department of Agriculture, Forest Service, Southern Research Station: Asheville, NC, USA, 2005; p. 85.

3. Omernik, J.; Griffith, G. Ecoregions of the conterminous United States: Evolution of a hierarchical spatial framework. *Environ. Manag.* **2014**, *54*, 1249–1266. [[CrossRef](#)] [[PubMed](#)]
4. Gesch, D.; Oimoen, M.; Greenlee, S.; Nelson, C.; Steuck, M.; Tyler, D. The National Elevation Dataset. *Photogr. Eng. Remote Sens.* **2002**, *68*, 5–11.
5. Homer, C.; Dewitz, J.; Yang, L.; Jin, S.; Danielson, P.; Xian, G.; Coulston, J.; Herold, N.; Wickham, J.; Megown, K. Completion of the 2011 National Land Cover Database for the conterminous United States—Representing a decade of land cover change information. *Photogr. Eng. Remote Sens.* **2015**, *81*, 345–354.
6. Gandhi, G.; Parthiban, S.; Thummalu, N.; Christy, A. Ndvi: Vegetation Change Detection Using Remote Sensing and Gis—A Case Study of Vellori District. *Procedia Comput. Sci.* **2015**, *57*, 1199–1210. [[CrossRef](#)]
7. LANDFIRE. Existing Vegetation Type Layer, LANDFIRE 1.1.0, U.S. Department of the Interior, Geological Survey. 2008. Available online: <http://landfire.cr.usgs.gov/viewer/> (accessed on 28 April 2018).
8. Lowry, J.; Ramsey, R.; Boykin, K.; Bradford, D.; Comer, P.; Falzarano, S.; Kepner, W.; Kirby, J.; Langs, L.; Prior-Magee, J. *Southwest Regional Gap Analysis Project: Final Report on Land Cover Mapping Methods*; RS/GIS Laboratory, Utah State University: Logan, UT, USA, 2005.
9. Escuin, S.; Navarro, R.; Fernandez, P. Fire severity assessment by using NBR (Normalized Burn Ratio) and NDVI (Normalized Difference Vegetation Index) derived from LANDSAT TM/ETM images. *Int. J. Remote Sens.* **2008**, *29*, 1053–1073. [[CrossRef](#)]
10. Davids, C.; Doulgeris, A. Unsupervised change detection of multitemporal Landsat imagery to identify changes in land cover following the Chernobyl accident. In *Proceedings of the Geoscience and Remote Sensing Symposium, Barcelona, Spain, 8–11 July 2008*; pp. 3486–3489.
11. Weng, Q.; Fu, P.; Gao, F. Generating daily land surface temperature at Landsat resolution by fusing Landsat and MODIS data. *Remote Sens. Environ.* **2014**, *145*, 55–67. [[CrossRef](#)]
12. Turner, M.; Gardner, G.; O'Neil, R. *Landscape Ecology in Theory and Practice: Pattern and Process*; Springer: New York, NY, USA, 2001; p. 406.
13. Pietsch, M. Contribution of connectivity metrics to the assessment of biodiversity—Some methodological considerations to improve landscape planning. *Ecol. Indic.* **2018**, *94*, 116–127. [[CrossRef](#)]
14. Stancioiu, P.; Nita, M.; Lazar, G. Forestland connectivity in Romania—Implications for policy and management. *Land Use Policy* **2018**, *76*, 487–499. [[CrossRef](#)]
15. Lechner, M.; Harris, R.; Doerr, V.; Doerr, E.; Drielsma, M.; Lefroy, E. From static connectivity modelling to scenario-based planning at local and regional scales. *J. Nat. Conserv.* **2015**, *28*, 78–88. [[CrossRef](#)]
16. Shafer, G. Land use planning: A potential force for retaining habitat connectivity in the Greater Yellowstone Ecosystem and beyond. *Glob. Ecol. Conserv.* **2015**, *3*, 256–278. [[CrossRef](#)]
17. Ersoy, E.; Jorgensen, A.; Warren, P. Identifying multispecies connectivity corridors and the spatial pattern of the landscape. *Urban For. Urban Green.* **2018**. [[CrossRef](#)]
18. Hogland, J.; Anderson, N. Function Modeling Improves the Efficiency of Spatial Modeling Using Big Data from Remote Sensing. *Big Data Cogn. Comput.* **2017**, *1*, 3. [[CrossRef](#)]
19. Hogland, J.; Anderson, N.; St. Peters, J.; Drake, J.; Medley, P. Mapping Forest Characteristics at Fine Resolution across Large Landscapes of the Southeastern United States Using NAIP Imagery and FIA Field Plot Data. *Int. J. Geo-Inf.* **2018**, *7*, 140. [[CrossRef](#)]
20. Masek, J.; Hayes, D.; Hughes, M.; Healey, S.; Turner, D. The role of remote sensing in process-scaling studies of managed forest ecosystems. *For. Ecol. Manag.* **2015**, *355*, 109–123. [[CrossRef](#)]
21. Hogland, J.; Anderson, N.; Chung, W. New Geospatial Approaches for Efficiently Mapping Forest Biomass Logistics at High Resolution over Large Areas. *Int. J. Geo-Inf.* **2018**, *7*, 156. [[CrossRef](#)]
22. St. Peters, J.; Hogland, J.; Anderson, N.; Drake, J.; Medley, P. Fine resolution probabilistic land cover classification of landscapes in the southeastern United States. *Int. J. Geo-Inf.* **2018**, *7*, 107. [[CrossRef](#)]
23. Graves, S.; Caughlin, T.; Asner, G.; Bohlman, S. A tree-based approach to biomass estimation from remote sensing data in a tropical agricultural landscape. *Remote Sens. Environ.* **2018**, *218*, 32–43. [[CrossRef](#)]
24. Faga, M.; Morton, D.; Cook, B.; Masek, J.; Zhao, F.; Nelson, R.; Huang, C. Mapping pine plantation in the southeastern U.S. using structural, spectral, and temporal remote sensing data. *Remote Sens. Environ.* **2018**, *216*, 415–426. [[CrossRef](#)]
25. Cheshner, A. The effect of measurement error. *Biometrika* **1991**, *78*, 451–462. [[CrossRef](#)]
26. Fuller, W. *Measurement Error Models*; Wiley: New York, NY, USA, 1987; p. 500.

27. Greenberg, J.; Dobrowski, S.; Ustin, S. Shadow allometry: Estimating tree structural parameters using hyperspatial image analysis. *Remote Sens. Environ.* **2005**, *97*, 15–25. [CrossRef]
28. Sheridan, R.; Popescu, S.; Gatzliolis, D.; Morgan, C.; Ku, W. Modeling Forest Above ground Biomass and Volume Using Airborne LiDAR metrics and Forest Inventory and Analysis Data in the Pacific Northwest. *Remote Sens.* **2015**, *7*, 229–255. [CrossRef]
29. Frazer, G.; Magnussen, S.; Wulder, M.; Niemann, K. Simulated impact of sample plot size and co-registration error on the accuracy of and uncertainty of LiDAR-derived estimates of forest stand biomass. *Remote Sens. Environ.* **2011**, *115*, 636–649. [CrossRef]
30. Bobakken, T.; Naesset, E. Assessing effects of positioning errors and sample plot size on biophysical stand properties derived from airborne laser scanner data. *Can. J. For. Res.* **2009**, *39*, 1036–1052. [CrossRef]
31. Saarela, S.; Schnell, S.; Tuominen, S.; Balazs, A.; Hyypä, J.; Grafstrom, A.; Stahl, G. Effects of positional errors in model-assisted and model-based estimation of growing stock volume. *Remote Sens. Environ.* **2016**, *172*, 101–108. [CrossRef]
32. Van Niel, T.; McVicar, T.; Li, L.; Gallant, J.; Yang, Q. The impact of misregistration on SRTM and DEM image differences. *Remote Sens. Environ.* **2008**, *112*, 2430–2442. [CrossRef]
33. Moran, P. Notes on Continuous Stochastic Phenomena. *Biometrika* **1950**, *37*, 17–23. [CrossRef] [PubMed]
34. Core Team. *R: A Language and Environment for Statistical Computing*; R Foundation for Statistical Computing: Vienna, Austria, 2014. Available online: <http://www.R-project.org/> (accessed on 28 April 2018).
35. Hijmans, R.; Etten, J. Raster: Geographic Analysis and Modeling with Raster Data. R Package Version 2.0-12. 2012. Available online: <http://CRAN.R-project.org/package=raster> (accessed on 24 September 2018).
36. Pebesma, E. Multivariable geostatistics in S: The gstat package. *Comput. Geosci.* **2004**, *30*, 683–691. [CrossRef]
37. Gräler, B.; Pebesma, E.; Heuvelink, G. Spatio-Temporal Interpolation using gstat. *R J.* **2016**, *8*, 204–218.
38. Landsat. Landsat Project Description. 2018. Available online: <https://landsat.usgs.gov/landsat-project-description> (accessed on 27 April 2018).
39. National Agriculture Imagery Program [NAIP]. National Agriculture Imagery Program (NAIP) Information Sheet. Available online: http://www.fsa.usda.gov/Internet/FSA_File/naip_info_sheet_2013.pdf (accessed on 14 May 2014).
40. Cressie, N. *Statistics for Spatial Data Revised Edition*; Wiley Classics Library, John Wiley: Hoboken, NJ, USA, 2015; p. 928.
41. Loveland, T.; Irons, J. Landsat 8: The plans, the reality, and the legacy. *Remote Sens. Environ.* **2016**, *185*, 1–6. [CrossRef]
42. Beyerhelm, C. Head-to-Head Comparison of Four SiRF-Based GPS Receivers. 2009 Report. Available online: <https://www.fs.fed.us/database/gps/documents/SiRFComp.pdf> (accessed on 24 September 2018).
43. Forest Inventory and Analysis Program [FIA]. *Forest Inventory and Analysis National Core Field Guide: Field Data Collection Procedures for Phase 2 Plots. Version 6.0. Vol. 1*; Internal Report; U.S. Department of Agriculture Forest: Washington, DC, USA, 2012. Available online: http://www.fia.fs.fed.us/library/field-guides-methods-proc/docs/2013/Core%20FIA%20P2%20field%20guide_6-0_6_27_2013.pdf (accessed on 5 June 2014).
44. Cribari-Neto, F.; Zeileis, A. Beta Regression in R. *J. Stat. Softw.* **2010**, *34*. [CrossRef]
45. Akaike, H. Information theory and an extension of the maximum likelihood principle. In *Selected Papers of Hirotugu Akaike*; Petrov, B.N., Csake, F., Eds.; Springer: New York, NY, USA, 1973; pp. 267–281.
46. Akaike, H. A new look at the statistical model identification. *IEEE Trans. Autom. Control* **1974**, *19*, 716–723. [CrossRef]
47. Patterson, P.L.; Williams, M.S. Effects of registration error between remotely sensed and ground data on estimators of forest area. *For. Sci.* **2003**, *49*, 110–118.
48. Zhang, M.; Lin, H.; Zeng, S.; Li, J.; Shi, J.; Wang, G. Impacts of plot location errors on accuracy of mapping and scaling up aboveground forest carbon using sample plot and landsat tm data. *IEEE Geosci. Remote Sens. Lett.* **2013**, *10*, 1483–1487. [CrossRef]
49. Frost, C.; Thompson, S. Correcting for regression dilution bias: Comparison of methods for a single predictor. *J. R. Statist. Soc. A* **2000**, *163*, 173–189. [CrossRef]

