Available online at www.sciencedirect.com
ScienceDirect
Agriculture, Ecosystems and Environment 118 (2007) 201–210
Agriculture Ecosystems & Environment
www.elsevier.com/locate/agee

Monitoring and verifying agricultural practices related to soil carbon sequestration with satellite imagery

Ross S. Bricklemyer [a,*], Rick L. Lawrence [a], Perry R. Miller [a], Norov Battogtokh [b]

[a] *Department of Land Resources and Environmental Sciences, Montana State University-Bozeman, P.O. Box 173120, Bozeman, MT 59717-3120, United States*
[b] *Mongolian Society for Sustainable Development, Baga Toiruu 38/A, Fokus tov, Ulaanbaatar 13, Mongolia*

Received 20 October 2005; received in revised form 5 April 2006; accepted 12 May 2006
Available online 27 June 2006

Abstract

The Kyoto Protocol entering into force on 16 February 2005 continues to spur interest in development of carbon trading mechanisms internationally and domestically. Critical to the development of a carbon trading effort is verification that carbon has been sequestered, and field level measurement of C change is likely cost prohibitive. Estimating C change based on agricultural management practices related to carbon sequestration seems more realistic, and analysis of satellite imagery could be used to monitor and verify these practices over large areas. We examined using Landsat imagery to verify crop rotations and quantify crop residue biomass in north central Montana. Field data were collected using a survey of farms. Standard classification tree analysis (CTA) and boosted classification and regression tree analysis (BCTA) were used to classify crop types. Linear regression (LM), regression tree analysis (RTA), and stochastic gradient boosting (SGB) were used to estimate crop residue. Six crop types were classified with 97% accuracy (BCTA) with class accuracies of 88–99%. Paired *t*-tests were used to compare the difference between known and predicted mean crop residue biomass. The difference between known and predicted mean residues using SGB was not different than 0 (p-value = 0.99); however root mean square error (RMSE) was large (1981 kg ha^{-1}), implying that SGB accurately predicted regional crop residue biomass but not local predictions (i.e., field or farm level). The results of this study, and previous research classifying tillage practices and estimating soil disturbance, supports using satellite imagery as an effective tool for monitoring and verifying agricultural management practices related to carbon sequestration over large areas.
Published by Elsevier B.V.

Keywords: Carbon sequestration; Crop types; Crop residue biomass; Classification and regression tree; Boosting; Landsat imagery

1. Introduction

The Kyoto Protocol (KP) entered into force on 16 February 2005 (http://unfccc.int, 22 April 2005). This event has continued interest in the development of carbon trading mechanisms. The United States chose not to participate in KP; however, the current administration has vowed to address domestic carbon dioxide emissions, and carbon trading will be a part of that effort (Pianin, 2002). Agricultural soils have the potential to sequester C from the atmosphere and help mitigate global climate change (Lal et al., 1998). Critical to the development of a carbon trading effort is verification that carbon has been sequestered. Field level measurement of C change is cost prohibitive with currently available technologies. Thus, estimating C change based on agricultural soil management practices is more feasible at this time.

Previous efforts modeling agricultural C dynamics reveal that several key agricultural practices are primarily responsible for changes in agricultural soil C (Parton et al., 1988, 1987). These practices include tillage systems, levels of soil disturbance, crops grown, including crop rotation practices, and amount of residual crop left after harvest. Satellite remote sensing has the potential to monitor and verify all of these practices over regional scales.

* Corresponding author. Tel.: +1 406 994 5119; fax: +1 406 994 3933.
E-mail address: rsb@montana.edu (R.S. Bricklemyer).

0167-8809/$ – see front matter. Published by Elsevier B.V.
doi:10.1016/j.agee.2006.05.017

Tillage disturbance has been shown to greatly influence soil C dynamics due to increased erosion and microbial decomposition (Paustian et al., 1997). The adoption of no-till (NT) can reduce losses of soil and can increase soil organic C (Lal et al., 1998). Previous studies have used remote sensing to predict tillage systems using various classification methods. Logistic regression (LR) of Landsat Enhanced Thematic Mapper Plus (ETM+) imagery had >95% accuracy in verifying NT fallow fields in a study in north central Montana (Bricklemyer et al., 2002). LR had 93% map accuracy using Landsat Thematic Mapper (TM) data in a corn/soybean rotation in Ohio (vanDeventer et al., 1997). Landsat TM and logistic regression have also been used to map tillage practices in the lower Minnesota River watershed using logistic equations developed by vanDeventer et al. (1997) and TM band 5 or the difference between TM bands 3 and 5 with 70–77% accuracy (Gowda et al., 2001). Logistic models applied to IKONOS imagery principal component (PC) 2 and PC 4 had 80 and 77% overall accuracy for discriminating corn/soybean residues and conventional/conservation tillage in Nebraska, respectively (Vina et al., 2003). Finally, the Crop Residue Index Multiband (CRIM) model using ETM+ imagery of the Minnesota River Basin, although not specifically addressing the NT/tillage question, had 79–80% accuracy classifying two categories, 0–30 and 31–100% residue cover, which were equivalent to conventional and NT management, respectively (Thoma et al., 2004). Additionally, soil disturbance has been estimated using regression tree analysis of Landsat ETM+ images (Bricklemyer et al., 2006a).

Monitoring cropping systems requires identification of various crop types. Identifying crop types and estimating yields using Landsat satellite imagery has been a focus of remote sensing experiments beginning with the Large Area Crop Inventory Experiment or LACIE in the middle 1970s (MacDonald et al., 1975). Studies subsequently have investigated improving classification methods for increasing crop type discrimination accuracy by overcoming a primary issue of separating crops. That primary issue was identified as the variability in crop maturity that can occur within a Landsat scene (Wheeler and Misra, 1980). Methods used to improve classification accuracy include the use of maximum likelihood classification (MLC) of single and multidate Landsat imagery, principal component analysis, discriminant analysis, and using active microwave response. Using an iterative MLC approach for classifying rice, maize, sorghum, and soybean, accuracy increased from 89% using single date imagery to >95% using 2 and 3 image dates (Van Niel and McVicar, 2004). Using MLC of a single date and including middle infrared bands and principal component analysis, 97% accuracy was achieved for discriminating oilseed crops from orchards, scrubs, and forest (Sharma et al., 1995). Classification of maize, durum wheat, and bread wheat using MLC and a single image date had an overall accuracy of 72% for Landsat ETM+ imagery and 81% for Earth Observing-1 Advanced Land Imager imagery (Lobell and Asner, 2003). Discriminant analysis of combined visible, near infrared, and active microwave response data had 92% accuracy for classifying corn, bare soil, bare soil + weeds, pasture, millet, weeds, and wheat stubble (Macelloni et al., 2002; Rosenthal and Blanchard, 1984). Dryland agriculture in the northern Great Plains, particularly Montana, has been a void for agricultural remote sensing research related to carbon sequestration and crop types. The diversity in crop types and seeding dates in semi-arid farming can be substantial.

Related to crop type is the amount of crop residue remaining after harvest. The Century model for agricultural C dynamics uses a crop production submodel to estimate crop biomass, yield, and residue biomass using inputs of crop type, fertilizer application, annual climatic data, and harvest practices (Parton et al., 1987). Site-specific data could be used instead of existing databases and would likely enhance the predictive capabilities of the model (Bricklemyer, 2003). Studies specifically quantifying crop residue biomass have not been documented. Previous studies, however, have successfully estimated the proportion of crop residue covering the soil using remote sensing techniques such as radar satellite data (McNairn et al., 1998), laser induced flouresence (Daughtry et al., 1996; McMurtrey et al., 1993), and Landsat TM and ETM+ (McNarin and Protz, 1993; Thoma et al., 2004; vanDeventer et al., 1997).

The overall objective of this research was to monitor and verify agricultural practices that influence soil carbon sequestration in farm fields, namely tillage systems, soil disturbance, crop types, and crop residue biomass. Tillage systems and soil disturbance have been previously addressed (Bricklemyer et al., 2006a). The present study focused on predicting crop types and estimating residue biomass. To meet these objectives we compared using: (1) classification tree analysis (CTA) and boosted classification tree analysis (BCTA) for predicting crop types and (2) linear regression (LM), regression tree analysis (RTA), and stochastic gradient boosting (SGB) for estimating crop residue biomass.

2. Methods

2.1. Study area

The study area was located in the north central Montana region known as the ''Golden Triangle''. The Golden Triangle is predominantly a dryland wheat (*Triticum aestivum* L.) production region roughly bounded by Great Falls to the south, Cut Bank to the northwest, and Havre to the northeast (Fig. 1). Spring wheat (SW), winter wheat (WW), and barley (Bly) (*Hordeum vulgare* L.) are the primary crops grown in the region with smaller acreages of lentils (Len) (*Lens culinaris* Medik). The area is semi-arid with 250–375 mm of average annual precipitation, the majority of which occurring May to mid-July.

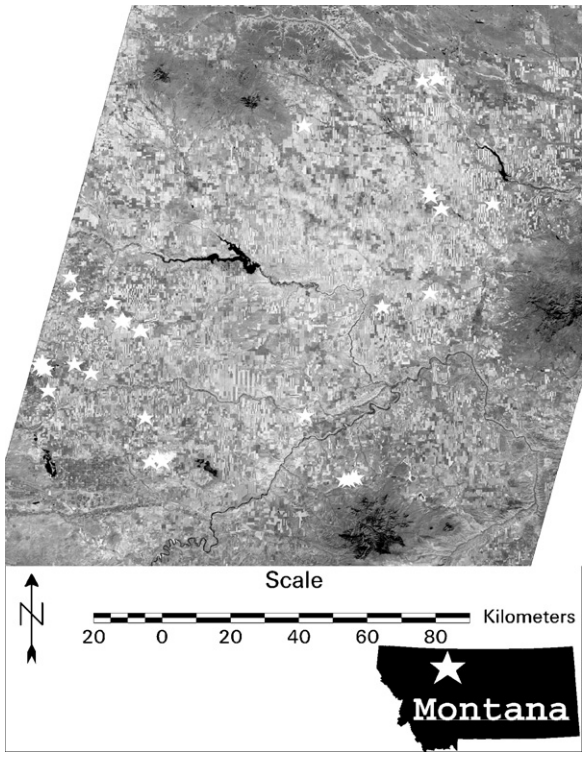

Fig. 1. 12 July 2002 Landsat ETM+ image of north central Montana. Small white stars are locations of fields used in the analyses. Inset image is the state of Montana with the large white star indicating the region of the study area.

2.2. Data collection

A farm survey of the study area was used to obtain field level farm management data. Farmers in the region were contacted by phone to determine if they would be interested in participating in the study. Those who agreed supplied legal descriptions (township, range, and section) of farm fields that were currently: (1) in crop, (2) in fallow, and (3) in the Conservation Reserve Program (CRP). True-color Landsat ETM+ subset images from a 26 June 2002 image were created for each participant using supplied legal descriptions. The 26 June image was chosen because it was the earliest cloud-free image for the 2002 growing season. The farm survey was built such that farmers identified the field(s) of interest on the subset image of their farm. Field level information about fallow management and equipment, seeding operations and equipment, pest control, fertilizer management, crop types, seeding/harvest dates and methods, and crop yields was collected corresponding to the fields identified by each farmer. Data collected in the survey effort provided the necessary information for identifying crop types and estimating crop residue. Fields surveyed were not visited to verify survey responses; all responses were assumed to be correct in 'good faith'.

2.3. Imagery

Three Landsat ETM+ images were used in this study. 26 June 2002 and 12 July 2002 images were used in predicting crop types and a 14 September 2002 image was used to estimate crop residue biomass. The June and July images were chosen because they were the highest quality (i.e., cloud-free) growing season images available. 14 September was chosen because it was necessary to use an image acquired post-harvest to estimate crop residue biomass.

2.4. Data preparation

2.4.1. Crop type

Currently cropped fields were chosen from survey responses for crop type analysis. Spectral digital numbers (DNs) of 42 cropped fields were extracted from center portions of selected fields and grouped by field. Pixel data from 21 of the fields, stratified by crop type, were randomly chosen for training, and the remaining fields were used for independent validation datasets. The crop types included in this study were SW, WW, Bly, Alfalfa, and Len. The majority of fields were evenly distributed between cereal crops; however two lentil fields were present.

Crop types mature at different times depending on seeding date and different grow habits. Those phenological differences should allow better discrimination of crop types. Two image dates, 26 June and 12 July, therefore were included in the analysis in order to take advantage of crop growth differences. All seven bands from each image date along with band difference values (July–June) and 14 principal components from a combined June + July image were used in the analysis.

2.4.2. Crop residue biomass

Spectral data for 39 cropped fields were used to estimate crop residue biomass. Three of the 42 fields included in the crop type analysis were excluded from the residue analysis because they were either not harvested by 14 September or were obstructed by partial cloud cover in the 14 September image. Spectral data from 20 of the 39 fields, randomly selected within crop types, were used as training data, and the 19 remaining fields were used for independent validation.

The crops included in this portion of the study were SW, Bly, Len, and WW. Grain yields reported in the survey results were used to calculate crop residue biomass. Grain yield data and harvest indices (HI) were representative of the conditions in the study area. Yield values were well distributed and were consistent for dryland and irrigated fields (i.e., dryland yields lower than irrigated yields). HI varies temporally (year to year at the same location) and spatially (between locations in the same year) for each crop. The HI values used to convert SW, Bly, and WW grain yield to residue biomass were averages calculated from a study with data from five locations across Montana (Bozeman, Amsterdam, Denton, Havre, and Dutton) (Miller and Holmes, 2005). That study experienced high, average, and low growing season precipitation, creating robust average harvest indices for each crop included in the present

study. The HI value for Len was calculated from a multiyear study in southwestern Saskatchewan (Miller et al., 2003). Southwestern Saskatchewan is in the same agroecoregion as the present study area (Padbury et al., 2002), thus the use of this data is appropriate. The HI value used for SW, Bly, and Len was 0.365 (S.E. = 0.04 for SW and Bly and S.E. = 0.02 for Len), and for WW was 0.33 (S.E. = 0.03). The equations used to calculate crop residue biomass were:

$$\begin{aligned} \mathrm{HI} &= \frac{\mathrm{grain}}{\mathrm{agbio}} \\ \mathrm{agbio} &= \frac{\mathrm{grain}}{\mathrm{HI}} \\ \mathrm{crop\,residue\,biomass} &= \mathrm{agbio} - \mathrm{grain} \end{aligned} \tag{1}$$

where HI is the harvest index, grain the dry mass of grain estimated from survey yield (kg ha^{-1}) and agbio is the aboveground biomass (kg ha^{-1}). Eq. (1) assumes that only grain was removed from the fields during harvest or post-harvest. Farm survey responses supported this assumption. The farm survey asked questions specifically regarding harvest and post-harvest residue management (i.e., harrowing and baling of residues). None of the fields in this study were harrowed or baled.

2.5. Data analysis

2.5.1. Crop type

Classification tree analysis is becoming more common for classification of remotely sensed data due, among other things, to higher accuracies than previously used methods (Freund and Schapire, 1996; Lawrence et al., 2004). Boosted classification tree analysis, a variant of CTA, has shown promising results for increasing accuracy over standard CTA (Lawrence et al., 2004) because of the multiple recursive partitioning trees developed using training data. BCTA works by developing multiple CTA trees. Each new tree developed is based on misclassifications of the previous tree. The "boosting" algorithm essentially focuses on the more difficult classifications in the previous trees (Freund and Schapire, 1996; Lawrence et al., 2004). Final classification is ultimately the result of a plurality "vote" of the multiple classification trees. CTA and BCTA models were built using the See-5 data mining statistical software (Quinlan, 2005).

2.5.2. Crop residue biomass

Crop residue biomass, a continuous variable, was estimated using linear regression, regression tree analysis (a regression version of CTA), and stochastic gradient boosting (SGB). SGB is a variant of RTA and can substantially increase classification accuracy over standard RTA (Lawrence et al., 2004). LM, RTA, and SGB analyses were performed using the R 2.0.1 statistical software (R, 2004). Regsubset, an all subsets regression model building procedure in the R leaps package (Lumley, 2004), was used to determine the best linear models using combinations of all seven ETM+ bands and their squares. The best model was chosen to have the highest adjusted coefficient of determination (adjusted r^2) value with all of the predictors in the model being significant ($\alpha = 0.05$). Regression trees were built using the rpart package (Therneau and Atkinson, 2004) and SGB was performed using the gbm package (Ridgeway, 2005). Mean values predicted by each method were compared to known mean values from an independent validation dataset using a two-sample t-test comparison of means. The t-test determined if the difference of the two means was significantly different than zero. A significant difference implies that the model is not accurately predicting the known mean value. Statistical differences are a function the number of observations (n), however, and n was large in this analysis, thus statistical significance might be more related to n than a meaningful difference in means. Additionally, root mean square error (RMSE) and 95% prediction intervals were calculated for each method.

3. Results

3.1. Crop type

Both CTA and BCTA performed well for classifying crop types. CTA had overall accuracy of 92% with class accuracies ranging from 77 to 95% (Table 1). Kappa (K_{hat}) was 0.87 suggesting that CTA was significantly better for classifying crop types than random chance (p-value < 0.0001). The majority of WW, Bly, Len, and Alf confusion occurred with SW. SW was most confused with Bly, then with WW and Len. CRP was best discriminated

Table 1
Non-boosted classification tree analysis confusion matrix and accuracy assessment for crop types using independent validation dataset pixels

	Reference						
	SW	WW	Bly	CRP	Len	Alf	Total
Class							
SW	20165	348	554	45	180	122	21414
WW	254	6094	77	37	19	12	6493
Bly	496	99	3239	7	6	14	3861
CRP	19	24	2	1218	3	0	1266
Len	158	14	10	4	951	55	1192
Alf	97	5	8	1	70	1537	1718
Total	21189	6584	3890	1312	1229	1740	35944
Producer's accuracy (%)	95	93	83	93	77	88	
User's accuracy (%)	94	94	84	96	80	89	
Overall accuracy (%)	92						
K_{hat}	0.87						
Z	374						
p-Value	<0.0001						

SW, spring wheat; WW, winter wheat; Bly, barley; Len, lentils; Alf, alfalfa; K_{hat}, Kappa; Z, z-score for significance testing.

Table 2
Boosted classification tree analysis confusion matrix and accuracy assessment for crop types (99 boosts) using independent validation dataset pixels

	Reference						
	SW	WW	Bly	CRP	Len	Alf	Total
Class							
SW	20906	117	291	21	106	57	21498
WW	57	6452	14	22	5	0	6550
Bly	136	13	3581	4	1	0	3735
CRP	2	0	2	1265	0	0	1269
Len	71	2	1	0	1077	32	1183
Alf	17	0	1	0	40	1651	1709
Total	21189	6584	3890	1312	1229	1740	35944
Producer's accuracy (%)	99	98	92	96	88	95	
User's accuracy (%)	97	99	96	99	91	97	
Overall accuracy (%)	97						
K_{hat}	0.95						
Z	653						
p-Value	<0.0001						

SW, spring wheat; WW, winter wheat; Bly, barley; Len, lentils; Alf, alfalfa; K_{hat}, Kappa; Z, z-score for significance testing.

(Table 1). BCTA out performed CTA and had overall accuracy of 97% with class accuracies ranging from 88 to 99% (Table 2). K_{hat} was 0.95, again suggesting that BCTA was significantly better at predicting crop type than random chance (p-value < 0.0001). WW, Bly, Len, and Alf were most confused with SW, and SW was most confused with Bly. BCTA improved accuracy in all classes by 3–12%; the largest increases occurred in the Bly and Len classes. BCTA was the more accurate classifier in this study with all accuracies exceeding the recommended 85% minimum level of interpretation accuracy using remotely sensed data (Lillesand and Kiefer, 2000).

Fig. 2 shows the base CTA decision tree. The complete tree was unable to be presented due to the number of secondary trees See-5 developed to predict crop types. Multiple image dates played a major role in distinguishing crop types. Difference in the blue portion of the spectrum (band 1) from June to July was the only predictor needed to classify CRP and was the primary decision point for WW. Secondary decision points for WW required the multidate PC 1 and 7 to distinguish it from SW. Alfalfa was most similar to SW and required the June predictors band 6 and PC 12 for final classification. Lentil was also similar to SW and required data from both image dates, predominantly in the form of band differences and principal components rather than bands from individual image dates. Bly and SW were most similar for they were not distinguishable until >10 levels into the decision tree. The majority of Bly was distinguished from SW based on differences in band 6. BCTA decision trees unfortunately are not easily interpretable or presentable because BCTA starts with the same base decision tree as CTA and then develops subsequent trees from the misclassifications in the first tree.

3.2. Crop residue biomass

Linear regression, regression tree analysis, and stochastic gradient boosting were used to estimate crop residue

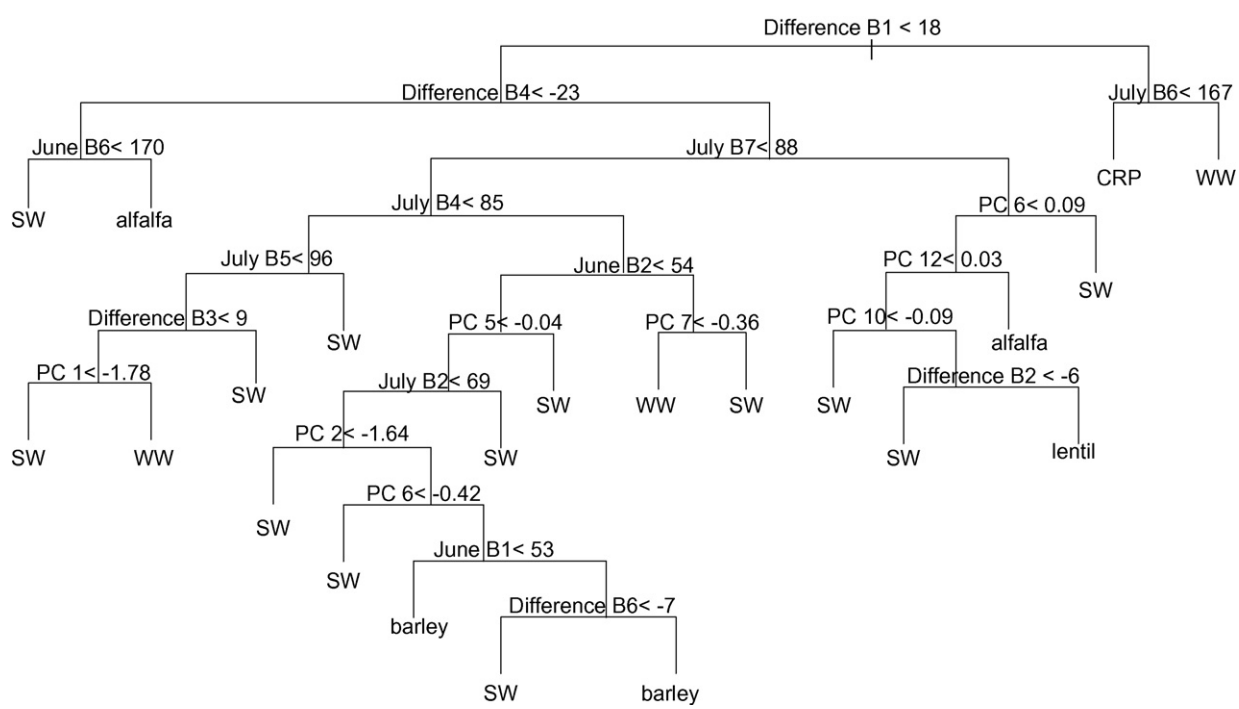

Fig. 2. Base classification tree analysis decision tree for classifying crop types in north central Montana. SW, spring wheat; WW, winter wheat; B, band from Landsat ETM+ image; PC, principal component from combined 26 June and 12 July 2002 ETM+ images.

Table 3
Two-sample *t*-test comparisons of known validation data and linear regression (LM), regression tree analysis (RTA), and stochastic gradient boosting (SGB) for crop residue biomass

Method	d.f.	Mean (kg ha^{-1})	Predicted mean (kg ha^{-1})	*p*-Value	RMSE (kg ha^{-1})	Difference from mean (%)
LM	36320	5893	4225	<0.001	2658	−28
RTA	36320	5893	4418	<0.001	3116	−25
SGB	36320	5893	5893	0.99	1981	<1

biomass (Table 3). The best LM (adjusted $r^2 = 0.29$) was:

$$\begin{aligned}\text{Crop residue biomass} &= 1081 + 51(\text{blue}) - 77(\text{NIR}) + 716(\text{MIR1}) \\ &\quad - 869(\text{MIR2}) - 41(\text{MIR1} \times \text{MIR2}) + 13(\text{MIR1}^2) \\ &\quad + 32(\text{MIR2}^2) \qquad (2)\end{aligned}$$

where blue was the Landsat ETM+ band 1, NIR the near infrared (Landsat ETM+ band 4), MIR1 the middle infrared (Landsat ETM+ band 5) and MIR2 was the middle infrared (Landsat ETM+ band 7). Blue and MIR1 both had positive relationships with crop residue biomass. NIR and MIR2 both had negative relationships with crop residue biomass.

Fig. 3 shows the RTA decision tree for predicting crop residue biomass. NIR, MIR1, thermal, and MIR2 (bands 4–7, respectively) were used to partition the data in 19 of the 21 RTA splits. Higher crop residue predictions tended to have low values in bands NIR, MIR1, and MIR2 and higher values in the thermal band; the relationship of bands NIR and MIR2 was consistent with LM. Green (band 2) was used in two decision nodes with inconsistent effects on predicted residue quantity. SGB results are not easily interpretable other than reporting the relative influence of each predictor. MIR1 (36%), NIR (30%), and MIR2 (25%) contributed 91% of the relative influence (i.e., relative importance) in predicting crop residue biomass, with bands 6 (6%) and 2 (3%) contributing the remaining 9%.

All three methods used for predicting crop residue biomass included bands 4, 5, and 7. This was consistent with results from other studies using Landsat imagery to determine the proportion of crop residue covering the soil also found bands 4, 5, and 7 important. Bands 4 and 5 were used in the normalized difference index (McNarin and Protz, 1993), bands 5 and 7 were used in the normalized difference tillage index (vanDeventer et al., 1997, and regression models using bands 3, 5, and 7 were found to be good predictors of residue cover (Thoma et al., 2004).

Paired *t*-test of means found that the LM predicted mean value of 4225 kg ha^{-1} underestimated the known crop residue mean of 5893 kg ha^{-1} (p-value < 0.0001) by 28% and had a RMSE of 2658 kg ha^{-1} (Table 3). RTA had a predicted mean value of 4418 kg ha^{-1} and a RMSE of 3116 kg ha^{-1}. The RTA predicted mean was 25% lower than the known mean (p-value < 0.0001). SGB better predicted the actual mean residue biomass than both LM and RTA. Comparison of means found that the SGB predicted mean value of 5893 kg ha^{-1} was not significantly different than

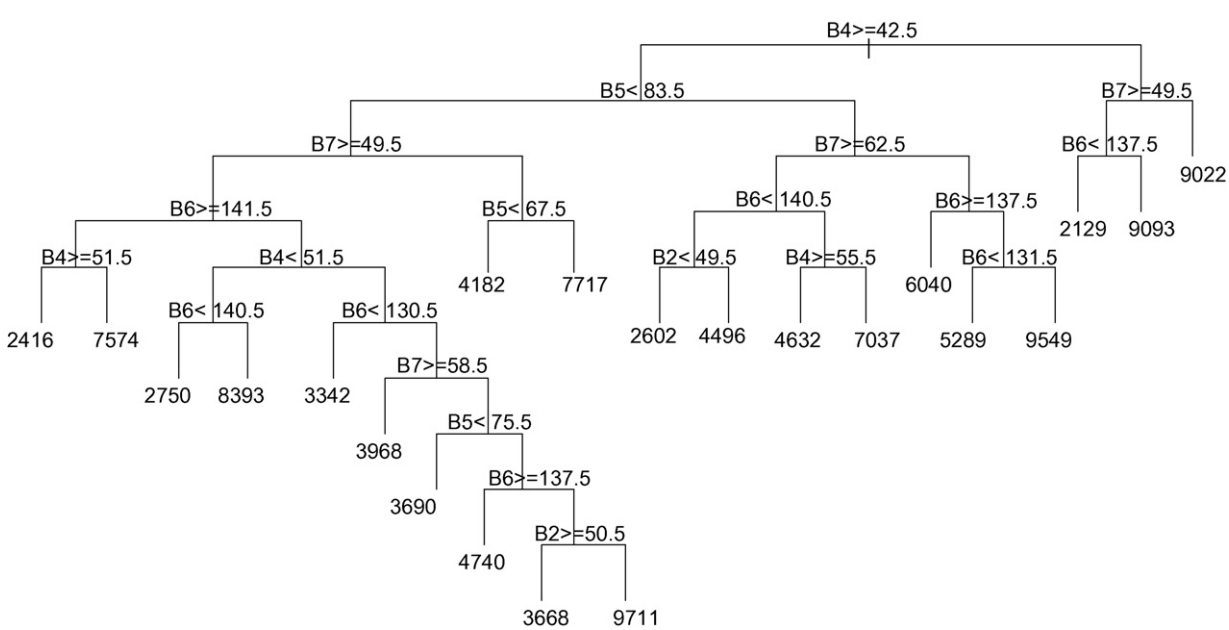

Fig. 3. Regression tree analysis decision tree for estimating the quantity of crop residue biomass in north central Montana. B is the band from 14 September Landsat ETM+ image.

the known crop residue mean (p-value = 0.99). SGB had a RMSE of 1981 kg ha^{-1}, smaller than both LM and RTA. Plots of predicted vs. actual crop residue values showed that the majority of predictions fell near the 1:1 line; however there is significant spread in the data for all methods (Fig. 4). Prediction intervals (95%) were widest for RTA, whereas LM and SGB were similar in width (Fig. 4). Although the prediction intervals were similar in width, fewer SGB predictions fell outside of the 95% prediction interval than LM predictions. This result is quantified by SGB having a smaller RMSE than LM. Residue prediction data clustered in columns because known residue values were calculated using Eq. (1) based on average crop yield per acre from survey results. Substantial variability in crop yield and thus calculated crop residues existed among pixels, as would be expected.

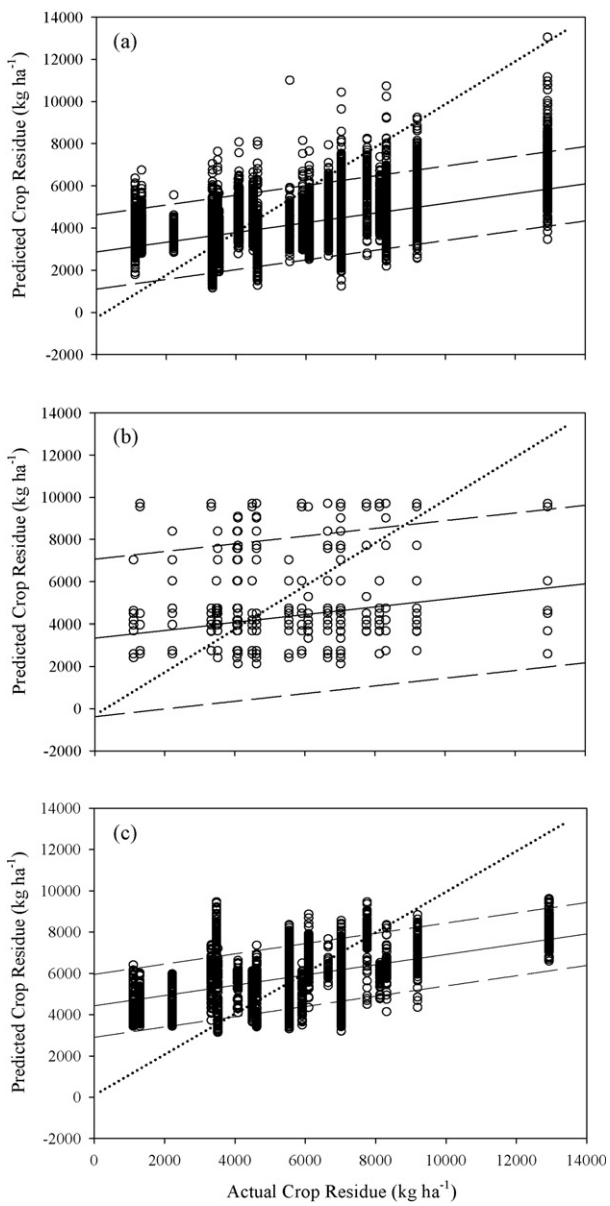

Fig. 4. Plot showing predicted crop residue biomass using linear regression (a), regression tree analysis (b), and stochastic gradient boosting (c) vs. computed crop residue biomass using farmer supplied crop yields from farm surveys and crop-specific harvest indices (kg ha^{-1}). Solid line is regression line, dashed lines are 95% prediction interval, and dotted line indicates 1:1 agreement.

4. Discussion

Many methods have been used for classifying various agricultural management practices using satellite imagery. In this study, classification of crop types was highly accurate (97%) and estimation of mean crop residue biomass was not different than known values (p-value = 0.99). No-till and tilled fields were classified in previous studies with >93% accuracy (Bricklemyer et al., 2002; vanDeventer et al., 1997). Finally, regional estimates of soil disturbance were not different than known values (Bricklemyer et al., 2006b). Using satellite imagery can be an effective tool for large area monitoring and verification of agricultural practices related to carbon sequestration.

4.1. Crop type

Discriminating crop types is a function of variability in crop growth habits. The use of multiple image dates was important for distinguishing crop types in this study. This was indicated by difference variables and principal components comprising 13 of the 21 decision nodes in CTA. Important difference variables were in bands 1, 5, and 7 and are likely related to crop growth habits and canopy closure. The proportion of soil and vegetation spectral mixing becomes dominated by vegetation as crops mature and the canopy closes. The amount that bands 1 (blue), 5 (MIR), and 7 (MIR) reflected changes as crops matured was attributable to effects of the crop canopy closure related to chlorophyll absorption of band 1, the sensitivity of band 5 to water in growing plants, and cellulose absorption (band 7) (Jensen, 1996; Thoma et al., 2004). The importance of individual bands 1, 5, and 7 for discriminating crop types was also likely a function of canopy cover for similar reasons. Bands 2 (green) and 4 (NIR) are known to be associated with green and NIR wavelengths reflected by healthy, photosynthetically active vegetation. Bands 2 and 4 could have been used in the models to separate fall seeded (WW) and perennial crops (alfalfa and CRP) from more recently spring seeded crops (SW, barley, and lentil) due again to differences in crop biomass and canopy development.

The diversity in crop types and whether a crop was fall or spring seeded likely contributed to confusion between crop types expressing spectrally similar crop phenology. Confusion occurred between SW, WW, and Bly (Tables 1 and 2). SW and Bly are both spring seeded cereal crops, have similar growth habits, and visually appear quite similar in the field until the formation of the seed head (i.e., heading). Heading

might have begun between the June and July images thus allowing for accurate discrimination between SW and Bly. The confusion between WW and SW could be due to similar spectral signatures, considering both are wheat crops. Accurate separation likely occurred due to differences in crop canopy cover and more advanced maturation of WW in the July image. WW in June and July would have had the most visible aboveground biomass compared to the spring seeded crop types in this study. Crop canopy cover and similar above ground biomass production in late June and early July might have contributed to confusion between late seeded SW and Len. CRP had the least confusion with other crops probably because the mixing of previous year's senescent biomass and present year's green growth would spectrally look quite different than the mono-crop conditions of the other crops. BCTA was a particularly effective method for classifying crop types given the variability in crop types and seeding dates.

BCTA was the better crop type predictor compared to CTA. BCTA increased accuracy across all crop type classes by 3–12%. The boosting algorithm appears to have been able make finer adjustments in the decision trees to better differentiate crops that look similar, SW and Bly, for example. The methods of producing subsequent decision trees that focus on the more difficult classifications, and the use of a plurality vote, more accurately distinguished crop types. It also appeared that both CTA and BCTA were not affected by the unbalanced nature of the dataset, although this is known to be an issue with these algorithms (Lawrence et al., 2004). There were substantially more SW fields/pixels than other classes (see column and row totals in Tables 1 and 2).

4.2. Crop residue biomass

The results of this study support that combinations of bands 4, 5, and 7 are important for predicting the quantity of crop residue biomass and residue cover. Band 4, as discussed above, is known to be associated with actively growing vegetation; hence a negative relationship might imply that more actively growing vegetation would correspond to lower crop residue biomass. This might be true for a post-harvest image; however this might also be causing inaccurate estimates. Band 5 is sensitive to the amount of water in plants (Jensen, 1996). The positive relationship between band 5 and crop residue biomass was likely a function of senescent crop residues having very little water content. Absorption of band 7 could be related to cellulose absorption by crop residues (Thoma et al., 2004). The negative relationship of band 7 would imply that with greater crop residue biomass, which is senescent and high in cellulose, more absorption of band 7 would occur.

Predictions using LM, RTA, and SBG have two spatial scales of accuracy to consider when tested against known values using paired *t*-tests. Accurately predicting the mean value signifies that the method would be a strong predictor of regional crop residue biomass, areas of over-prediction would be balanced by areas of under-prediction (Fig. 4). A small RMSE and narrow 95% prediction interval would suggest that the model also would accurately predict individual pixel and/or field values. Both the LM and RTA models significantly underestimated the mean crop residue biomass (p-value < 0.0001 for both methods) by 28 and 25%, respectively, and had large RMSE values. This suggested that neither LM nor RTA were good methods for predicting crop residue biomass in this study. SGB accurately predicted the mean crop residue biomass value (p-value = 0.99) and had a RMSE smaller than the other prediction methods; however RMSE was 34% of the known mean average. The results suggested that SGB was a strong predictor of regional crop residue biomass, however large RMSE values suggested that none of the methods were good local predictors of crop residue biomass when given known crop yields and representative harvest indices.

5. Conclusion

Analysis of Landsat imagery can be an effective tool for monitoring and verifying agricultural practices related to carbon sequestration. This study showed that six crop types in dryland Montana agriculture could accurately be classified (97% accuracy) using BCTA. The ability to classify various crops across a region has the potential to enhance regional soil C sequestration estimates by verifying and documenting crop rotations and cropping intensity, both of which are important factors in modeling soil C dynamics.

Multiple image dates within a year would be necessary for monitoring and verifying management practices that influence soil carbon sequestration. The earliest post-seeding cloud-free spring image date (i.e., early May) would be most useful for determining tillage practices and soil disturbance (Bricklemyer et al., 2006a). Image dates in the early to middle (i.e., June) and late (i.e., late July–August) portions of the growing season would be necessary for accurate crop type determination. In addition to multiple within-year images, multiple year images are required for documenting crop rotations. A post-harvest image date (i.e., mid-late September) is necessary for estimating crop residues. Timely image acquisition would be high priority for monitoring and verifying agricultural management practices related to carbon sequestration. An interesting side effect of image timing is that producers would not know the imagery dates ahead of time, so attempts to wait until after image acquisition to till, for example, would not be practical.

Crop residue biomass could be accurately estimated at regional scales using SGB. None of the methods used, however, were able to predict accurately at the local (i.e., field or farm) level. Estimates of crop residue biomass could be used in conjunction with biophysical models, such as Century, for more accurate estimation of soil carbon changes over time. Additionally, with further research in C cycling in semi-arid Montana agriculture, residue estimates could be

used to develop regional indicators of annual C sequestration by estimating the amount of above- and below-ground biomass C returned to the soil (knowing crop types and crop yield) and estimating the proportion of that C that is converted to soil organic carbon.

The specific models developed in this study are scene dependant do to the nature of empirical models, however the methods used to develop the models are not scene dependant and are easily employed given ground-truth information. Annual field data on at least a subset of fields will be required because of the scene dependant nature of empirical models. The farm survey provided excellent field level data; however the process was time and resource intensive. Local governmental agencies could play a major role in supporting this type of monitoring and verification of C sequestering practices, the Farm Service Agency (FSA) and conservation districts, for example. Analysis of Landsat ETM+ imagery can be an effective, low cost tool for monitoring and verifying agricultural management practices related to carbon sequestration across large areas with high accuracy at local and region scales.

Acknowledgements

The authors wish to thank the Upper Midwest Aerospace Consortium (UMAC) and the Consortium for Agricultural Soils Mitigation of Greenhouse Gases (CASMGS) for funding this research; all farm survey participants in north central Montana for their time and valuable field information, and Doug Weist for organizing and administering the farm survey.

References

Bricklemyer, R.S., 2003. Sensitivity of The Century Model for Estimating Sequestered Soil Carbon Using Coarse- and Fine-Scale Map Data Sources in North Central Montana. Montana State University, Bozeman.

Bricklemyer, R.S., Lawrence, R.L., Miller, P.R., 2002. Documenting no-till and conventional till practices using Landsat ETM plus imagery and logistic regression. J. Soil Water Conserv. 57 (5), 267–271.

Bricklemyer, R.S., Lawrence, R.L., Miller, P.R., Battogtokh, N., 2006a. Predicting tillage practices and agricultural soil disturbance in north central Montana with Landsat imagery. Agr. Ecosyst. Environ. 114 (2–4), 210–216.

Bricklemyer, R.S., Lawrence, R.L. Battogtokh, N., Miller, P.R., 2006b. Monitoring and verifying agricultural practices related to soil carbon sequestration. Agr. Ecosyst. Environ., in press.

Daughtry, C.S.T., McMurtrey, J.E., Chappelle, E.W., Hunter, W.J., Steiner, J.L., 1996. Measuring crop residue cover using remote sensing techniques. Theor. Appl. Climatol. 54 (1–2), 17–26.

Freund, Y., Schapire, R.E., 1996. Experiments With A New Boosting Algorithm. Morgan Kaufman, San Francisco.

Gowda, P.H., Dalzell, B.J., Mulla, D.J., Kollman, F., 2001. Mapping tillage practices with landstat Thematic Mapper based logistic regression models. J. Soil Water Conserv. 56 (2), 91–96.

Jensen, J.R., 1996. Introductory Digital Image Processing: A Remote Sensing Perspective, second ed. Prentice Hall Inc., Upper Saddle River, New Jersey, pp. 152–153.

Lal, R., Kimble, J.M., Follet, R.F., Cole, C.V., 1998. The Potential of U.S. Cropland to Sequester Soil Carbon and Mitigate The Greenhouse Effect. Ann Arbor Press, Chelsea, MI.

Lawrence, R., Bunn, A., Powell, S., Zambon, M., 2004. Classification of remotely sensed imagery using stochastic gradient boosting as a refinement of classification tree analysis. Remote Sens. Environ. 90 (3), 331–336.

Lillesand, T.M., Kiefer, R.W., 2000. Remote Sensing and Image Interpretation, fourth ed. Wiley & Sons, New York.

Lobell, D.B., Asner, G.P., 2003. Comparison of Earth Observing-1 ALI and Landsat ETM+ for crop identification and yield prediction in Mexico. IEEE Trans. Geosci. Remote Sens. 41 (6), 1277–1282.

Lumley, T., R Development Core Team, 2004. R: A Language and Environment for Statistical Computing. R Foundation for Statistical Computing, Vienna, Austria. , ISBN: 3-900051-07-0, http://www.R-project.org.

MacDonald, R.B., Hall, F.G., Erb, R.B., 1975. The use of LANDSAT data in Large Area Crop Inventory Experiment (LACIE). In: Proceedings of the Symposium on Machine Processing of Remotely Sensed Data, LARS, Purdues University, West Lafayette, IN.

Macelloni, G., Palosci, S., Panipaloni, P., Ruisi, R., Dechambre, M., Valentin, R., Chanzy, A., Wigneron, J.P., 2002. Active and passive microwave measurements for the characterization of soils and crops. Agronomie 22 (6), 581–586.

McMurtrey, J.E., Chappelle, E.W., Daughtry, C.S.T., Kim, M.S., 1993. Fluorescence and reflectance of crop residue and soil. J. Soil Water Conserv. 48 (3), 207–213.

McNairn, H., Wood, D., Gwyn, Q.H.J., Brown, R.J., Charbonneau, F., 1998. Mapping tillage and crop residue management practices with RADARSAT. Can. J. Remote Sens. 24 (1), 28–35.

McNarin, H., Protz, R., 1993. Mapping corn residue cover on agricultural fields in Oxford County, Ontario, using Thematic Mapper. Can. J. Remote Sens. 19 (2), 152–159.

Miller, P.R., Holmes, J.A., 2005. Cropping sequence effects of four broadleaf crops on four cereal crops in the northern Great Plains. Agron. J. 97, 189–200.

Miller, P.R., Gan, Y., McConkey, B.G., McDonald, C.L., 2003. Pulse crops for the northern Great Plains. 1. Grain productivity and residual effects on soil water and N. Agron. J. 95, 972–979.

Padbury, G., Waltman, S., Caprio, J., Coen, G., McGinn, S., Mortensen, D., Nielsen, G., Sinclair, R., 2002. Agroecosystems and land resources of the northern Great Plains. Agron. J. 94 (2), 251–261.

Parton, W.J., Stewart, J.W.B., Cole, C.V., 1988. Dynamics of C, N, P and S in grassland soils: a model. Biogeochemistry 5 (1), 109–131.

Parton, W.J., Schimel, D.S., Cole, C.V., Ojima, D.S., 1987. Analysis of factors controlling soil organic-matter levels in Great-Plains grasslands. Soil Sci. Soc. Am. J. 51 (5), 1173–1179.

Paustian, K., Andrén, O., Janzen, H.H., Lal, R., Smith, P., Tian, G., Tiessen, H., Van Noordwijk, M., Woomer, P.L., 1997. Agricultural soils as a sink to mitigate CO2 emissions. Soil Use Manage. 13 (4), 230–244.

Pianin, E., 2002. Bush Touts Greenhouse Gas Plan. , p. A01http://www.washingtonpost.com.

Quinlan, J.R., 2005. Data Mining Tools See5 and C5.0. RULEQUEST RESEARCH, St. Ives, NSW, Australia. , www.rulequest.com/see5-info.html.

R Development Core Team, 2004. R: A Language and Environment for Statistical Computing. R Foundation for Statistical Computing, Vienna, Austria. , ISBN: 3-900051-07-0, http://www.R-project.org.

Ridgeway, G., 2005. gbm: Generalized Boosted Regression Models. R Package Version 1.4-2. http://www.i-pensieri.com/gregr/gbm.shtml.

Rosenthal, W.D., Blanchard, B.J., 1984. Active microwave responses—an aid in improved crop classification. Photogramm. Eng. Remote Sens. 50 (4), 461–468.

Sharma, S.A., Bhatt, H.P., Ajai, 1995. Oilseed crop discrimination—selection of optimum bands and role of middle infrared. ISPRS J. Photogramm. Remote Sens. 50 (5), 25–30.

Therneau, T.M., Atkinson, B., 2004. rpart: Recursive Partitioning. R Package Version 3.1-20. S-PLUS 6.x original at: http://www.mayo.edu/hsr/Sfunc.html.

Thoma, D.P., Gupta, S.C., Bauer, M.E., 2004. Evaluation of optical remote sensing models for crop residue cover assessment. J. Soil Water Conserv. 59 (5), 224–233.

Van Niel, T.G., McVicar, T.R., 2004. Determining temporal windows for crop discrimination with remote sensing: a case study in south-eastern Australia. Comput. Electron. Agr. 45 (1–3), 91–108.

vanDeventer, A.P., Ward, A.D., Gowda, P.H., Lyon, J.G., 1997. Using Thematic Mapper data to identify contrasting soil plains and tillage practices. Photogramm. Eng. Remote Sens. 63 (1), 87–93.

Vina, A., Peters, A.J., Ji, L., 2003. Use of multispectral ikonos imagery for discriminating between conventional and conservation agricultural tillage practices. Photogramm. Eng. Remote Sens. 69 (5), 537–544.

Wheeler, S.G., Misra, P.N., 1980. Crop classification with Landsat Multispectral Scanner Data 2. Pattern Recogn. 12 (4), 219–228.