Available online at www.sciencedirect.com
SCIENCE DIRECT®

Agriculture, Ecosystems and Environment 114 (2006) 210–216

Agriculture Ecosystems & Environment

www.elsevier.com/locate/agee

Predicting tillage practices and agricultural soil disturbance in north central Montana with Landsat 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, P.O. Box 173120, Bozeman, MT 59717-3120, USA
[b] Mongolian Society for Sustainable Development, Baga Toiruu 38/A, Fokus tov, Ulaanbaatar 13, Mongolia

Received 3 May 2005; received in revised form 19 October 2005; accepted 28 October 2005
Available online 20 December 2005

Abstract

Management of agricultural soils, most notably tillage, influences wind, and water erosion, which in turn has implications for non-point source pollution of pesticides, fertilizer, and sediment in agro-ecosystems. No-till (NT) practices improve soil, water, and aquatic ecosystem quality by reducing soil erosion and chemical runoff. The ability of cropland soils to sequester C from the atmosphere might help mitigate global warming. Classification of Landsat ETM+ satellite images has the potential to identify tillage practices and soil disturbance over large areas, enabling efficient monitoring of these agricultural practices. Previous studies predicting tillage management had relatively small study areas (located in a single county), relatively low numbers of fields (6–51), and were temporally focused on non-planted fields to reduce the potential effects of crop canopy interference and/or field patterning. Our objectives were to predict in the presence of crop canopy and over a spatially large, management diverse study area (1) tillage systems (NT versus tilled) and (2) soil disturbance. A farm survey of the study area, north central Montana, was used to as a means to obtain extensive field-level farm management data. We compared logistic regression (LR), traditional classification tree analysis (CTA), and boosted classification tree analysis (BCTA) for identifying NT fields. Logistic regression had an overall accuracy of 94%, BCTA 89%, and CTA 87%, but tillage was not well distinguished. Soil disturbance was estimated using linear regression (LM), regression tree analysis (RTA), and stochastic gradient boosting (SGB), an RTA variant. Classification of soil disturbance was best achieved using RTA (predicted mean soil disturbance not significantly different than known soil disturbance, p-value = 0.08). Classification of Landsat ETM+ imagery showed promise for predicting tillage and agricultural soil disturbance over large, heterogeneous areas.
© 2005 Elsevier B.V. All rights reserved.

Keywords: Tillage; Logistic regression; Classification and regression tree analysis; Boosted classification tree analysis; Stochastic gradient boosting; Carbon sequestration

1. Introduction

Management of agricultural soils affects many facets of both natural and agro-ecosystems. Agricultural practices, most notably tillage, influence wind and water erosion, which in turn has implications for non-point source pollution of pesticides, fertilizer, and sediment. Tillage and soil disturbance also affect soil organic matter (SOM) dynamics. Reducing or eliminating tillage, managing crop residue, increasing cropping intensity, diversifying crop rotations, and efficiently managing fertilizer are management practices that diminish the potential for environmental impacts from agriculture and tend to increase SOM content in cropland soils (Campbell et al., 2000; Lal, 1998; Liang et al., 1999; Peterson et al., 1998; Potter et al., 1997; Mickelson et al., 2001). National inventories of potential non-point pollution source areas and regions with high carbon sequestration potential related to tillage and soil disturbance are needed in order to efficiently target pollution mitigation strategies and carbon sequestration opportunities.

Mapping areas with substantial acreage managed with tillage could help identify areas prone to soil erosion and chemical runoff. These same areas might also have the

* Corresponding author. Tel.: +1 406 994 5409; fax: +1 406 994 5122.
E-mail address: rickl@montana.edu (R.L. Lawrence).

0167-8809/$ – see front matter © 2005 Elsevier B.V. All rights reserved.
doi:10.1016/j.agee.2005.10.005

potential to store C if management changed. Survey data from the study area show a greater proportion of no-till management compared to tillage management; however, it has been estimated that 2 million acres or approximately 24% of cropland in Montana are in no-till (CTIC, 2004). Reducing soil erosion and chemical runoff would improve soil, water, and aquatic ecosystem quality (Mickelson et al., 2001). The ability of cropland to sequester C from the atmosphere, helping mitigate global warming, has the potential to add value to farmland and agricultural farm management (Lal et al., 1998).

Using remote sensing to determine tillage practices has been limited, especially in dryland wheat (*Triticum aestivum* L.) regions. Landsat Enhanced Thematic Mapper Plus (ETM+) imagery and logistic regression (LR) had >95% accuracy in verifying NT fallow fields in a limited study in north central Montana (Bricklemyer et al., 2002). Classification of conservation tillage practices, including NT, has been successful in climates moister than north central Montana. Landsat Thematic Mapper (TM) data were used to determine tillage practices in a corn (*Zea mays* ssp.)/soybean (*Glycine max* L.) rotation in Ohio using six logistic regression models with resulting map accuracy of 93% (vanDeventer et al., 1997). Landsat TM and logistic regression have also been used to map tillage practices in the Lower Minnesota River watershed using the logistic equations developed by vanDeventer et al. (1997) (Gowda et al., 2001). Models using TM band 5 or the difference between TM bands 3 and 5 had 70–77% accuracy. Ikonos imagery has been used more recently to discriminate conventional and conservation tillage practices in Nebraska (Vina et al., 2003). Five Ikonos bands and four principal components (PCs) were evaluated to determine which band(s) best discriminated between corn and soybean residues and conventional and conservation tillage. Logistic models applied to PC 2 and PC 4 had 80 and 77% overall accuracy for discriminating corn/soybean residues and conventional/conservation tillage, respectively (Vina et al., 2003). Finally, the Crop Residue Index Multiband (CRIM) model, although not specifically addressing the NT/tillage question, was used to classify residue cover into 2, 3, and 5 categories using ETM+ imagery in the Minnesota River Basin (Thoma et al., 2004). The highest accuracy (79–80%) occurred when classifying two categories, 0–30 and 31–100% residue cover, which were equivalent to conventional and NT management, respectively (Thoma et al., 2004).

Application of the methods used in the previous studies is limited due to temporal, spatial, and cultural reasons. All of the previous tillage prediction studies temporally focused on non-planted fields in the analysis to reduce the potential effects of crop canopy interference and/or field patterning. All of the studies had relatively small study areas (located in a single county) and low numbers of fields (6–51), with one exception (Thoma et al., 2004, which covered 13 counties and 468 fields). The diversity in crops was limited to corn and/or soybean, however. Restricting studies to non-cropped fields, small study areas, and few number of farm fields likely does not capture the potential variability in regional farm management including crop types, seeding dates, and equipment used.

Variability in farm management can be substantial locally (i.e., within counties) and more so regionally (i.e., multiple counties). Not all farmers employ the cultural practice of fallow, row width and row spacing configurations vary with farm managers, and timing of management operations vary dependently on local weather conditions, for example. There are also substantial variations in both NT and tillage equipment that can influence the proportion of soil disturbed.

Classification tree analysis (CTA) is becoming a popular method of classifying remotely sensed data (Lawrence et al., 2004), while regression trees are applied to continuous data analyses. Boosted classification tree analysis (BCTA), including stochastic gradient boosting, is a variant of standard CTA that has the potential for greater prediction accuracy, although the results are more difficult to interpret (DeFries and Chan, 2000; Lawrence et al., 2004). Methods of boosting CTAs are also commonly called voting or ensemble methods and operate by generating multiple CTA trees with each subsequent tree "boosted" based on classification errors from the previous tree. Each new tree thus focuses on the more difficult classifications in the previous tree (Freund and Schapire, 1996; Lawrence et al., 2004). The final classification is the result of a plurality "vote" of the multiple classification trees. Previous research has shown that BCTA can achieve substantial improvements in prediction accuracy over single classification trees (Freund and Schapire, 1996; Lawrence et al., 2004), although it can also lead to reduced accuracies depending on the data (Lawrence et al., 2004).

The objectives of this research were to predict in the presence of crop canopy in a spatially large, management diverse study area (1) tillage system (NT versus tilled) and (2) soil disturbance, calculated as the proportion of the soil surface disturbed by seeding and fertilizer application. To meet these objectives, we compared the accuracy of our previous method of logistic regression (LR) for identifying no-till fields to traditional CTA and BCTA. We also compared linear regression (LM), regression tree analysis (RTA), and stochastic gradient boosting (SGB) for estimating soil disturbance using Landsat ETM+ imagery.

2. Methods

2.1. Study area

The study area was located in the dryland wheat growing region in north central Montana, roughly bound by Great Falls to the south, Cut Bank to the northwest, and Havre to the northeast (Fig. 1). Primary crops grown in the study area

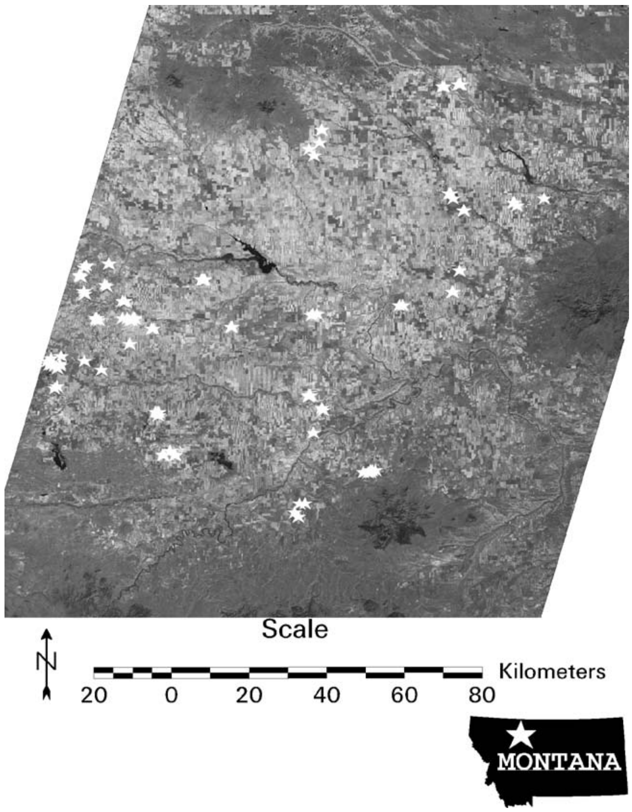

Fig. 1. 26 June 2002 Landsat ETM+ image of study area. White stars represent location of fields included in the study.

are spring wheat, winter wheat, and barley (*Hordeum vulgare* L.). Study area soils are dominated by deep, medium to fine textured ustolls, orthents, and argids with a frigid temperature regimes formed in nearly flat to strongly rolling glacial till plains. Elevation of cropland area ranges from 600 to 1400 m, and average annual precipitation ranges from 250 to 375 mm, the majority of which occurs in spring and early summer.

2.2. Data collection

A survey was used to obtain field-level farm management data. Farmers in the region were initially contacted by phone to determine if they would be interested in participating in the study and, if 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). CRP fields were not used in this analysis. Using legal descriptions to locate fields on satellite imagery, true-color Landsat ETM+ subset images from a 26 June 2002 image were created for each participant. The 26 June image was chosen because it was the earliest cloud free image. The farm field survey was built such that farmers identified the field(s) of interest on the subset image of their farm. The survey asked a series of farm management questions that corresponded to the fields identified on the subset satellite image. 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.

Two general management classes (NT and tilled) and soil disturbance were defined using the survey responses. The NT class was defined as fields where crops were directly seeded into the previous crop's standing stubble and weeds were managed strictly with herbicide. The tilled class was defined as fields that were managed using any type of equipment that employs soil inversion to manage weeds prior to seeding the crop. Soil disturbance was calculated as the proportion of soil disturbed as a function of row width, row spacing and broad-field tillage. Row widths ranged from 2.5 to 15 cm and row spacing ranged from 18 to 30 cm:

$$\mathrm{SD} = \frac{\mathrm{rw}}{\mathrm{rs}} \tag{1}$$

where SD is the soil disturbance, rw the row width, and rs is the row spacing.

Fields with 2.5 cm row width and 25 cm row spacing, for example, had a soil disturbance factor of 0.10. At the extremes were broad-field tillage with a soil disturbance factor: 1.0, and NT fallow fields having soil disturbance: 0.0.

Digital brightness values (DNs) for all seven Landsat ETM+ spectral bands for each pixel within surveyed fields were extracted from a mosaic of two contiguous, georeferenced 26 June 2002 Landsat ETM+ images. The thermal band of Landsat was resampled using the nearest neighbor method to account for differences in spatial resolutions between bands. Pixels from the center portions of fields were selected by field corresponding to the 116 fields delineated by the farmers in the survey responses. Pixels, grouped by field, were randomly separated into training and validation datasets by management and soil disturbance in order to capture representative variability associated with seeding dates and soil properties, which varied widely across the surveyed fields.

2.3. NT versus tillage analysis

LR, CTA, and BCTA models were used to classify NT and tilled management. LR analysis was performed using the S-Plus statistical software (Insightful Corp., 2001) and methods described in Bricklemyer et al., 2002. LR is an appropriate method for binomial (0–1) questions, thus when using non-cropped fields, discerning NT and tilled fields is a binomial question. Once cropped fields are included in the analysis the question might be no longer binomial due to the influence of crop canopy. CTA and BCTA can be used for both binomial and multiple class questions. CTA and BCTA models were built using the See-5 data mining statistical software (Quinlan, 1992). Accuracy was determined on a pixel basis, but all accuracy assessment pixels came from different fields than were used for training to ensure independence of the accuracy assessment. Producer's accuracy, user's accuracy, overall accuracy, and Kappa analysis assessed classification results (Congalton and Green, 1999). A Z-test of the Kappa statistics for each classification method determined if (1) the classification method was significantly better than random chance and (2) the classification methods were significantly different from one another.

2.4. Soil disturbance analysis

Soil disturbance was estimated using LM, RTA, and SGB. All analyses were performed using the R 2.0.1 statistical software (R, 2004). All models were built using 90% of fields with 10% of fields held out for independent validation. The regsubset, an all subsets regression model building procedure in the leaps package (Lumley, no date), 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). The rpart package (Therneau and Atkinson, 2005) built our regression trees and the gbm package (Ridgeway, 2004) performed the SGB analysis. Values predicted by each method were compared to known values from the independent validation dataset using a paired t-test comparison of means to determine if the difference was significantly different than zero and produce 95% confidence intervals (CIs). A significant difference suggested that the model was not accurately predicting the mean of known values; however, statistical differences are a function n, and n was large in this analysis, thus statistical significance might be less important in this instance.

3. Results

3.1. NT versus tilled analysis

Overall accuracy and NT class accuracy of the three methods used to predict tillage were encouraging; however tilled class accuracy was substantially lower. LR had 94%, CTA 87%, and BCTA 89% overall accuracy (Tables 1–3). The final logistic model was:

$$\begin{aligned}\text{logit}(\pi) = {} & -46.9 + 0.215(\text{band } 4) + 0.275(\text{band } 6) \\ & + 0.651(\text{band } 3) - 0.559(\text{band } 2) \\ & - 0.028(\text{band } 5) - 0.228(\text{band } 1) \\ & - 0.150(\text{band } 7)\end{aligned}$$

NT class producer's and user's accuracies were consistently >90%, however tilled class accuracies ranged from 18 to 80% with LR user's accuracy performing best. Logistic regression unexpectedly outperformed CTA and BCTA, however BCTA out performed CTA as expected. LR for classifying tillage systems using remote sensing works by first choosing the best predictor variables (i.e., ETM+ bands) to predict the probability of tillage, then by applying a cutoff value to find the best single split in the data to optimize classification accuracy. CTA and BCTA start with the best single split for classifying the data, and then use the next best predictors to find multiple paths to the same classification. We expected that multiple paths to the same correct response would be a more accurate method than using a single split in the data given the diversity present in our data. Crop types included winter wheat (WW) seeded in the fall of 2001 and spring wheat (SW), barley Bly), lentils (*Lens culinaris Medik*), and peas (*Pisum sativum*) planted between 7 April and 28 May, which consequently caused canopy cover to range from 0% in fallow fields to canopy closure (100%) in winter wheat fields. The majority of fields were in cereal crops (37 SW, 19 WW, and 19 Bly) with two lentil and pea fields included.

K_{hat} values were 0.4 for LR and 0.2 for both CTA and BCTA. Z-test of significance found K_{hat} values were highly significant (p-value < 0.0001) (Tables 1–3). Significant

Table 1
Logistic regression confusion matrix and accuracy assessment for no-till and tillage classes

	Reference		
	No-till	Tillage	Total
Class			
No-till	36984	2109	39093
Tillage	217	871	1088
Total	37201	2980	40181
Producer's accuracy (%)	99	29	
User's accuracy (%)	95	80	
Overall accuracy (%)	94		
K_{hat}	0.4		
Z	42		
p-Value	<0.0001		

K_{hat} is estimated Kappa statistic and Z is the K_{hat} Z-score used to test for significance.

Table 2
Non-boosted classification tree analysis confusion matrix and accuracy assessment for no-till and tillage classes

	Reference		
	No-till	Tillage	Total
Class			
No-till	34157	2156	36313
Tillage	3044	824	3868
Total	37201	2980	40181
Producer's accuracy (%)	92	28	
User's accuracy (%)	94	21	
Overall accuracy (%)	87		
K_{hat}	0.2		
Z	24		
p-Value	<0.0001		

K_{hat} is estimated Kappa statistic and Z is the K_{hat} Z-score used to test for significance.

Table 3
Boosted classification tree analysis confusion matrix and accuracy assessment for no-till and tillage classes (99 boosts)

	Reference		
	No-till	Tillage	Total
Class			
No-till	35367	2448	37815
Tillage	1834	532	2366
Total	37201	2980	40181
Producer's accuracy (%)	95	18	
User's accuracy (%)	94	23	
Overall accuracy (%)	89		
K_{hat}	0.2		
Z	19		
p-Value	<0.0001		

K_{hat} is estimated Kappa statistic and Z is the K_{hat} Z-score used to test for significance.

differences occurred between all possible combinations of the three methods in pairwise Z-test comparisons (Table 4).

3.2. Soil disturbance analysis

LM, RTA, and SGB over-estimated known soil disturbance. The best LM equation for predicting soil disturbance was:

$$\text{Soil disturbance} = -6.7 + 0.005(\text{B4}) - 0.018(\text{B5}) + 0.087(\text{B6}) + 0.089(\text{B7}) + 0.00011(\text{B4} \times \text{B7}) - 0.000023(\text{B4}^2) - 0.00027(\text{B6}^2) - 0.000025(\text{B7}^2)$$

The intercept and all predictor variables were significant (p-value < 0.05), and the model had an adjusted $R^2 = 0.16$. Bands 4, 6, and 7 were important bands in RTA and were used in six of the eight regression tree nodes. Bands 4 (53%) and 6 (32%) contributed to 85% of the relative influence for predicting soil disturbance in the SGB analysis. The remaining bands contributed to 15% combined.

A paired t-test determined if the there was a statistically significant difference between the predicted and known means for soil disturbance. A p-value > 0.05 meant that we were not able to detect a significant difference between the means, thus we could conclude on average accurate predictions. Soil disturbance predicted using LM and SGB were significantly greater than the known mean soil disturbance (p-value < 0.001). LM and SGB overestimated mean soil disturbance by 36 and 13%, respectively (Table 5). We were not able to detect significant differences between soil disturbance predicted using RTA and the known soil disturbance at $\alpha = 0.05$ (p-value = 0.08), although the 95% CI was narrow and included 0 (Table 5). This suggested that RTA was the best method for predicting for soil disturbance in this study.

Table 4
Pairwise comparison of logistic regression, classification tree analysis, and boosted classification tree analysis for classifying no-till and tilled management

Pairwise comparison (Z-test)	Z-score	p-Value
LR vs. CTA	21	<0.0001
LR vs. BCTA	19	<0.0001
CTA vs. BCTA	3	0.003

LR: logistic regression; CTA: classification tree analysis; BCTA: boosted classification tree analysis; alpha: 0.05.

4. Discussion

4.1. NT versus tilled analysis

LR was the better method of differentiating NT fields from tilled fields for our data as compared to CTA and BCTA. Ninety-four percent accuracy is very encouraging and is comparable to the previous study in the same region (Bricklemyer et al., 2002), however, the tilled class accuracies were unacceptable. A study area of this magnitude has tremendous variability in soils, seeding dates, fallow management, farming equipment, and timing of operations that could lead to low accuracy values. The LR equation included bands 3 and 4. These bands are sensitive to green leaf biomass, suggesting that vegetation was impacting the model's predictions. NDVI calculated for all fields used in the study area averaged 0.35 with a standard deviation of 0.18. Variability in crop canopy cover was the most likely reason for low tilled class accuracy in all methods used, because many of the tilled fields in the analysis were currently being cropped. Crop canopy coverage ranged from 0% in fallow fields to 100% in winter wheat fields. Classification of the NT class was likely affected by crop canopy cover as well, however accuracy was less affected because of the substantially greater number of NT pixels in the analysis. Tables 1–3 show that there were a large number of NT pixels misclassified. LR was better able to take into account this variability and produce highly accurate classifications overall and for the NT class.

We expected the CTA and BCTA to outperform LR. LR is appropriate for strictly binary responses, and the variability in our data, including, for example, fallow and cropped fields, was expected to result in a response that was not truly binary. Using a combination of NT and tilled fields in fallow and cropped states, for example, might have resulted in four levels of response. Our results, however, indicated that LR was able to account for these distinctions. BCTA had a higher overall accuracy compared to CTA; however, tilled class accuracies were lower than CTA. Our data were heavily skewed in favor of NT, which was evident in pixel totals and the low tilled class accuracies in all three methods (Tables 1–3). Unbalanced datasets are known to cause problems with CTA methods, including BCTA (Lawrence and Wright, 2001). We also analyzed our data with a reduced NT dataset, and the results were not improved. Although the

Table 5
Two sample t-test comparisons of known validation data and linear regression (LM), regression tree analysis (RTA), and stochastic gradient boosting (SGB)

Method	d.f.	Known mean	Predicted mean	p-Value	95% CI	Over-prediction (%)
LM	11582	0.135	0.183	0	−0.05, −0.04	36
RTA	11582	0.135	0.142	0.08	−0.01, 0.001	5
SGB	11582	0.135	0.153	0	−0.02, −0.01	13

reduced dataset was balanced, it likely did a poorer job of representing NT variability as a result. BCTA can achieve substantial improvements in prediction accuracy over single classification trees (Freund and Schapire, 1996; Lawrence et al., 2004), however it can also lead to reduced accuracies depending on the data, as seen in this study. LR nonetheless was the better predictor than both CTA and BCTA, shown by higher overall accuracy and class accuracies.

4.2. Soil disturbance analysis

Classification of soil disturbance was best achieved using RTA. LM and SGB were not useful methods for predicting soil disturbance. Data suggest that the boosting algorithm in SGB better dealt with the variability and unbalanced nature of the dataset than LM. The training data included soil disturbance values from 0 to 1.0 with the majority of data being skewed closer to 0, which is evident in the mean soil disturbance value of 0.135, calculated across all fields in the survey. We used this known mean soil disturbance value as a means to estimate regional soil disturbance, which is more useful information for land managers and for use as inputs into biophysical carbon models. RTA had the ability to more accurately estimate soil disturbance on a regional scale even though the data, as discussed previously, were complex and highly variable in this study.

RTA should theoretically predict local soil disturbance (pixel/field-level) based on the narrow 95% CI. Visual inspection of known versus predicted values, however, showed that although mean values were similar, pixel estimates of soil disturbance ranged greatly. This could be a function of a number of factors. The effect of crop canopy and time since disturbance cannot be ignored. Fields with crops or that were disturbed many months prior to the image acquisition date would be less likely to be accurately estimated. NT fields planted with winter wheat, for example, would have been disturbed as long as eight months before the image date. Conversely, a tilled field planted with spring wheat on 25 May would be expected to be more accurately estimated. Fields managed with broad-field tillage (soil disturbance = 1.0) were not well estimated in this study, nor were NT fallow fields (soil disturbance = 0.0). This could be a function of the regression analyses that was a part of all methods and did not produce a 0 outcome given any combination of predictors, although some values approached 1.0.

Soil disturbance is an important input for the Century biogeochemical carbon model, which is used to predict carbon sequestration in agricultural soils, grasslands, forests, and savannas (Parton et al., 1988). Century more recently has become the process model behind the Natural Resources Conservation Service's (NRCS) Voluntary Reporting Carbon Management Online Tool (COMET-VR) (http://www.cometvr.colostate.edu/, last visited 6 April 2005). Monitoring/verification of management practices that store carbon will likely be a part of any domestic carbon credit program, whether market-based or a national effort. Classification of remotely sensed images has the potential for highly accurate and rapid monitoring/verification of soil disturbance and changes in agricultural management practices.

5. Conclusion

The models developed in this study are scene specific empirical models that would not be transferable to different regions or acquisition dates. The methods employed are relatively easily accomplished, however, given reference data (i.e., survey data). Reference data could be difficult to obtain directly from farmers as in this study, however local governmental offices may be able to supply large amounts of farm/field specific data.

Predicting NT and tillage management in the presence of a crop canopy and in a spatially large, management diverse study area proved to be challenging. Logistic regression (94% accurate) outperformed both CTA (87% accurate) and BCTA (89% accurate) for discriminating NT and tilled fields. Tilled class accuracy was unacceptable and unusable in all three methods, however and was likely a result of the variability in management operations and, most notably, due to crop canopy interference. Using LR and Landsat ETM+ imagery has potential for mapping the extent of NT and tilled management practices across large areas, but with some considerations. Studies now suggest that LR is applicable in both semi-arid, lower production regions such as Montana (present study; Bricklemyer et al., 2002) and moister, more productive regions such as the mid-west (vanDeventer et al., 1997; Gowda et al., 2001; Vina et al., 2003). Previous studies predicting tillage practices used non-cropped fields in relatively small study areas, which could have been relatively homogenous with respect to farm management operations. Timely image acquisition will remain critical for useful classification. Selecting images when fields have minimal green biomass and without crop canopy, such as in previous studies, resulted in higher accuracy compared to including cropped fields in this study. Intrinsically, this makes sense because a crop canopy could mask the tillage signature.

The problem remains of monitoring and verifying management practices in fields that are continuously cropped. CTA and BCTA did not perform as well as LR in this study, however the results using these methods are encouraging and warrants continued exploration. The ability to map management practices has applications for resource managers looking to focus mitigation activities in areas with high non-point source pollution potential and/or high carbon sequestration potential. Determining the best method for predicting agricultural land management practices needs further investigation.

Estimation of soil disturbance was best accomplished using RTA (5% over prediction of the mean). Using SGB (13% over prediction of the mean) also shows promise for estimating a continuous variable using Landsat ETM+ spectral bands. The ability to estimate soil disturbance has potential applications to farm program eligibility, carbon sequestration modeling, and monitoring/verification of practices for carbon credit reporting. Some of the new farm programs, such as the Environmental Quality Incentive Program (EQUIP) and the Conservation Security Program (CSP), have specific eligibility criteria related to historic and present land management. Landsat ETM+ and TM images are nationally archived and available for analysis of past practices.

The effects of agricultural management on non-point source pollution can be reduced by targeting mitigation efforts in high risk areas. Carbon sequestration efforts must be monitored and verified by documenting management practices and/or modeling soil carbon change using biophysical models. Accurate classification of Landsat-based images for documenting tillage practices and soil disturbance has the potential to detect high risk non-point pollution areas, monitor management practices, and estimate inputs for carbon models.

Acknowledgements

The authors would like 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 survey.

References

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, 267–271.

Campbell, C.A., Zentner, R.P., Liang, B.C., Roloff, G., Gregorich, E.C., Blomert, B., 2000. Organic C accumulation in soil over 30 years in semiarid southwestern Saskatchewan—effect of crop rotations and fertilizers. Can. J. Soil Sci. 80, 179–192.

Congalton, R.G., Green, K., 1999. Assessing the Accuracy of Remotely Sensed Data: Principles and Practices. Lewis Publications, Boca Raton.

CTIC, 2004. National survey shows more farmers choose conservation. Conservation Technology Information Center, West Lafayette, IN (News release, 9 Nov).

DeFries, R.S., Chan, J.C.W., 2000. Multiple criteria for evaluating machine learning algorithms for land cover classification from satellite data. Remote Sens. Environ. 74, 503–515.

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, 91–96.

Insightful Corp., 2001. S-Plus 6 for Windows Guide to Statistics, vol. 1. Insightful Corp., Seattle, WA.

Lal, R., 1998. Mulching effects on runoff soil erosion, and crop response on alfisols in western Nigeria. J. Sustain. Agric. 11, 135–154.

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, 331–336.

Lawrence, R.L., Wright, A., 2001. Rule-based classification systems using classification and regression tree (CART) analysis. Photogram. Eng. Remote Sens. 67, 1137–1142.

Liang, B.C., McConkey, B.G., Campbell, C.A., Curtin, D., Schoenau, J., Lafond, G.P., Brandt, S.A., Moulin, A., 1999. Effect of tillage and crop rotations on the light fraction organic carbon and carbon mineralization in Chernozemic soils of Saskatchewan. Proc. Soils Crops 552–569.

Lumley T, no date available. Using Fortran Code by Alan Miller. Leaps: Regression Subset Selection. R package version 2.7.

Mickelson, S.K., Boyd, P., Baker, J.L., Ahmed, S.I., 2001. Tillage and herbicide incorporation effects on residue cover, runoff, erosion, and herbicide loss. Soil Till. Res. 60, 55–66.

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, 109–131.

Peterson, G.A., Halvorson, A.D., Havlin, J.L., Jones, O.R., Lyon, D.J., Tanaka, D.L., 1998. Reduced tillage and increasing cropping intensity in the Great Plains conserves soil C. Soil Till. Res. 47, 207–218.

Potter, K.N., Jones, O.R., Torbert, H.A., Unger, P.W., 1997. Crop rotation and tillage effects on organic carbon sequestration in the semiarid southern Great Plains. Soil Sci. 162, 140–147.

Quinlan, J., 1992. C4.5: Programs for Machine Learning. Morgan Kaufmann Pub., San Mateo, CA.

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 (URL: http://www.R-project.org).

Ridgeway, G., 2004. gbm: Generalized Boosted Regression Models. R package version 1.3-5 (http://www.i-pensieri.com/gregr/gbm.shtml).

Therneau, T.M., Atkinson B., 2005. R Port by Brian Ripley ⟨ripley@stats.ox.ac.uk⟩. rpart: Recursive Partitioning. R package version 3.1-22. 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, 224–233.

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. Photogram. Eng. Remote Sens. 63, 87–93.

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