Comparison of Sub-pixel Classification Approaches for Crop-specific Mapping

Yang Shao and Ross S. Lunetta

U.S. Environmental Protection Agency, National Exposure Research Laboratory,
Research Triangle Park, NC 27711, USA (e-mail: shao.yang@epa.gov)

***Abstract*—This paper examined two non-linear models, Multilayer Perceptron (MLP) regression and Regression Tree (RT), for estimating sub-pixel crop proportions using time-series MODIS-NDVI data. The sub-pixel proportions were estimated for three major crop types including corn, soybean, and wheat; throughout the entire 480,000 km^2 Laurentian Great Lakes Basin. Accuracy assessments were conducted using the cropland data layer (CDL) developed by the National Agricultural Statistics Service (NASS). The performances of the sub-pixel classifications were compared based on root-mean-square error (RMSE) and scatter plots. For MLP regression, the RMSE values at 500 m spatial resolution were 0.16, 0.14, and 0.07 for corn, soybean and wheat, respectively. The RT approach achieved slightly higher RMSE values of 0.18, 0.15, and 0.07 for corn, soybean, and wheat. The latter approach did not provide greater interpretability, because tree sizes were rather large for MODIS-NDVI sub-pixel crop estimation problems.**

Keywords- MODIS NDVI; Sub-pixel; MLP regression; Regression tree

I. INTRODUCTION

The Moderate Resolution Imaging Spectroradiometer (MODIS) data has been increasingly used for crop mapping and other agricultural applications. Phenology-based classification approaches using the NDVI (Normalized Difference Vegetation Index) 16-day composite (250 m) data product is among the most promising for automated processing. Most MODIS-NDVI crop mapping applications to date have focused on per-pixel classification methods; while the sub-pixel crop patterns and proportions have not been thoroughly exploited. Linear mixture model is one of the most commonly used approaches for estimating sub-pixel land-cover proportions [1], [2]. The spectral response of a pixel is assumed to be a linear combination from contributing sub-pixel land-cover proportions. The linear mixture assumption, however, might not be valid when NDVI values are considered as endmembers or inputs to linear mixture models. In addition, there might be high level iterations between MODIS-NDVI features or bands from different acquisition dates. The linear mixture model thus might not be appropriate for MODIS-NDVI sub-pixel analysis. The non-linear models such as neural network and regression tree are better suited for handling complex datasets with nonlinear relationships and high-order feature interactions [3]. Although these non-linear unmixing models have been widely used for many remote sensing applications [4], few studies have compared their performances for crop-specific sub-pixel estimations. The main objective of this paper was to implement and compare two non-linear unmixing models: (a) Multilayer Perceptron (MLP) neural network regression algorithm; and (b) Regression tree (RT) approach. The sub-pixel proportions were estimated for three major crop types including corn, soybean, and wheat; throughout the Great Lakes Basin (GLB).

II. METHODS

A. DATA

MODIS 16-Day composite NDVI data (MOD13Q1) for year 2007 was obtained from the USGS EROS Data Center. All available cropland data layers (CDLs) for 2007 across the US portion of the GLB were obtained from the United States Department of Agriculture (USDA). The CDLs were primarily developed from AWiFS imagery and had a high classification accuracy (>90%) for most major crop types (NASS, 2007). For the GLB, CDLs were mainly developed for the states of Michigan and Ohio. The spatial resolution of CDL is 56 m, which is finer than the 250 m MODIS-NDVI dataset.

A general cropland and non-cropland image classification was conducted first using MODIS-NDVI imagery as input. Training pixels for cropland and non-cropland were derived using the 2001 National Land Cover Dataset (NLCD) as references. A MLP neural network classification approach was employed to develop a cropland/non-cropland mask [5]. All the non-cropland pixels were labeled as zero percent for crop types of corn, soybean, and wheat. The sub-pixel estimations for individual crop types were then focused on the pixels within the cropland mask.

B. Sub-pixel estimation using MLP regression and regression tree

MLP is often used for per-pixel classification applications, but it can also serve as a regressor when the sub-pixel proportions are directly used as output targets in network training. The neural network simply approximates the regression function between the input patterns (MODIS-NDVI values) and the target values (sub-pixel crop proportions). The advantage of using MLP regression is that it does not make assumption about the relationship between the MODIS-NDVI time-series data and crop sub-pixel proportions. The relationship is simply learned from the training samples.

The sub-pixel crop proportions for the training/testing pixels were calculated using the 2007 CDLs as references. Specifically, the 2007 CDLs were overlaid onto the MODIS NDVI imagery. For each 250 m MODIS NDVI pixel, we calculated the proportions for three individual crop types of corn, soybean, and wheat, respectively. A total of 5,000 MODIS-NDVI pixels were randomly selected as the training data points for the MLP regression approach. We divided the training data points into two sets. One set was used to training the MLP neural network and the other set was used as validation data to guide the network training. This is particularly useful for building a more generalized neural network, which may reduce the overfitting problem. A three layer MLP neural network was designed for the sub-pixel crop estimation. The input layer consisted of 13 MODIS-NDVI values from Julian days of 123 to 234. NDVI values from other days were discarded because of low information contents (e.g., snow cover in winter). A total of 30 nodes were initially used in the hidden layer. There was only one output node in the output layer, representing the sub-pixel crop proportions. We used sigmoid and linear activation functions for the hidden layer and output layer, respectively. A back propagation algorithm was used to train the MLP regression model. A number of network training parameters (i.e., learning rate, momentum) were examined to obtain best performance. The trained MLP network was then employed to predict sub-pixel crop proportions for the entire GLB. It must be noted that three independent MLP networks were developed, corresponding to three individual crop types (corn, soybean and wheat).

Regression tree relies on a binary recursive partitioning on input dataset to reduce the sum of square errors [6]. We used the same training dataset as those for the MLP regression approach to train the regression tree. Using the validation dataset, a pruning process was conducted to determine optimal tree size. Three independent RTs were developed to estimate sub-pixel proportions for corn, soybean and wheat. The trained regression trees were then used to produce estimations of sub-pixel crop proportions for the GLB.

C. Accuracy Assessments

The accuracy of sub-pixel crop estimation was assessed using RMSE:

$$RMSE = \sqrt{\frac{\sum_{i=1}^{n}(E_i - A_i)^2}{n}} \qquad (1)$$

where E_i is the estimated crop proportions from MODIS NDVI pixels, A_i is the actual crop fraction cover derived from 2007 CDLs and n is the total number of pixels used in testing. To reduce the impacts of mis-registration between the MODIS-NDVI imagery and CDLs, we calculated RMSE values at two aggregated spatial resolutions of 500 and 1,000 m.

III. RESULTS

For MLP regression approach, the best sub-pixel classification performances were obtained around 150 training iterations. RMSE values at 500 m scale were 0.16, 0.14, and 0.07 for corn, soybean, and wheat, respectively. The results were slightly better than those from the RT approach (corn=0.18; soybean=0.15; wheat=0.07). The RMSE values decreased when pixels were aggregated to 1,000 m spatial resolution, especially for corn and soybean estimations. This was expected because RMSE measures typically decrease as pixel aggregation level increases. The impacts of mis-registration between MODIS-NDVI imagery and CDLs might be reduced at coarse spatial scale.

	RMSE (500 m)		RMSE (1,000 m)	
	MLP	RT	MLP	RT
Corn	0.16	0.18	0.11	0.14
Soybean	0.14	0.15	0.10	0.12
Wheat	0.07	0.07	0.04	0.05

Table 1. Comparisons of RMSE values for MLP and RT approaches.

For both 500 and 1,000 m spatial resolution, the MLP regression approach outperformed the RT approach, although the differences of RMSE values were only about 1–3% for three different crop types. Table 1 shows comparisons of RMSE values for MLP regression and RT approaches.

Sub-pixel estimations from MODIS-NDVI were plotted against actual crop proportions derived from CDLs at 1,000 m spatial scale (Fig. 1 a–f). Only 1/100 of pixels are plotted here to make the plot readable. The results from MLP regression approach show slightly less scattering, suggesting better overall performances. The main reason that the MLP regression approach outperformed the RT approach is due to its stronger interpolation capabilities, especially for large and complex ecological datasets [7]. Although many researchers suggested that RT approach has greater interpretability, it was not applicable for the MODIS-NDVI sub-pixel models, because the tree sizes were fairly large, even after the pruning procedure. The RT approach, however, is relatively easy for implementation purpose because there is no need for specifying model parameters. For the MLP regression approach, it is important to adjust network learning parameters such as the number of nodes at hidden layer, learning rate, momentum, and activation function to obtain best performance.

Both the MLP regression and RT approaches tended to underestimate wheat proportions, especially for pixels with high wheat proportions. The low RMSE values (0.05) were misleading because they are simple statistics for global measures. Scatter plots (Fig 1c and Fig 1f) show that both MLP and RT approaches performed better for pixels with very low wheat proportions (<20%) compared to those with higher wheat proportions. For corn and soybean estimations, there are also points with large differences between the MODIS-NDVI derived sub-pixel crop proportions and actual proportions from CDLs. One possible reason is that MODIS-NDVI data alone does not provide enough information for an accurate estimation of sub-pixel crop proportions. A potential solution is to include more input features such as additional MODIS bands, surface temperature, and digital elevation model.

Figure 2 shows sub-pixel crop distributions derived using the MLP regression approach. The sub-pixel crop proportions were grouped into five categories based on the abundance of crop proportions for each pixel. A majority of MODIS-NDVI crop pixels have mixed crop proportions. Further studies are needed to understand the relationship between the actual cropping units/field sizes and the spatial resolution of MODIS-NDVI imagery.

IV. CONCLUSIONS

Sub-pixel crop proportions were estimated using a MLP regression and a RT approach using time-series MODIS-NDVI data. The accuracy assessments suggest that the MLP regression approach performed better than the RT approach, although only 1–3% of RMSE differences were observed for these two non-linear unmixing models. The RT approach did not provide better interpretability, because relatively large tree sizes were needed to model complex relationships between the MODIS-NDVI data and sub-pixel crop proportions. Additional research is needed to enhance the accuracy of sub-pixel estimates for major crop types in the GLB.

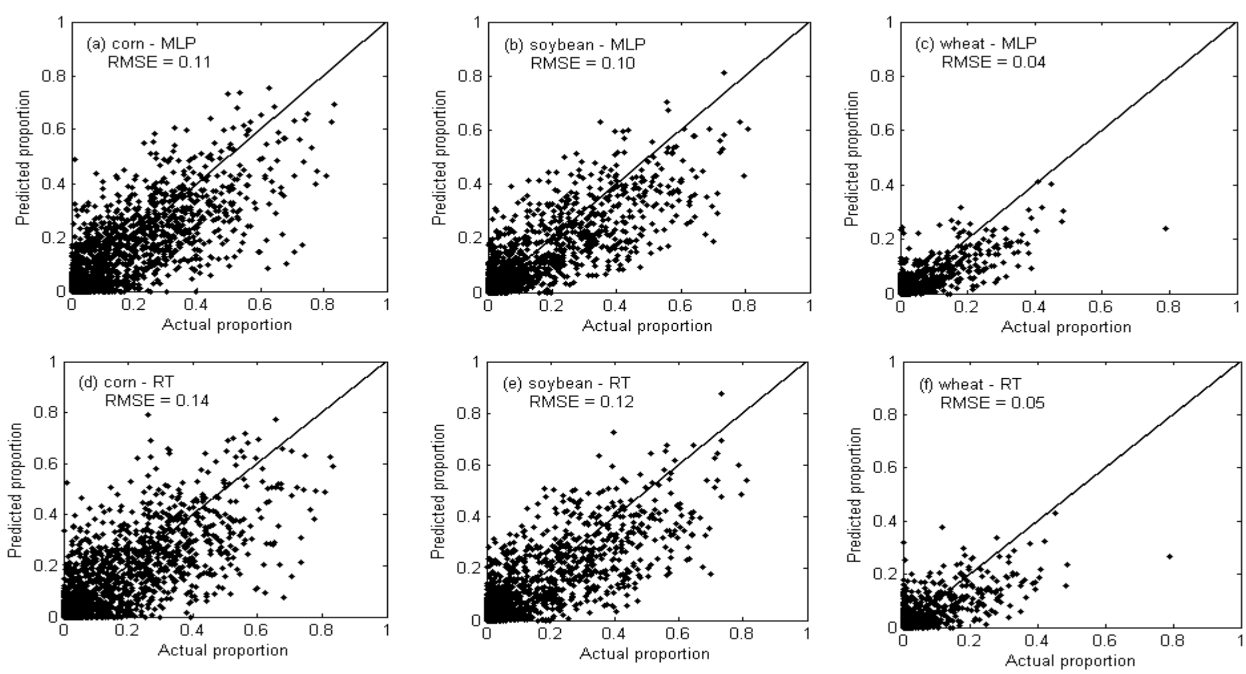

Fig. 1. Scatter plots of MODIS-NDVI derived sub-pixel crop proportions versus actual crop proportions for MLP (a-c) and RT (d-f).

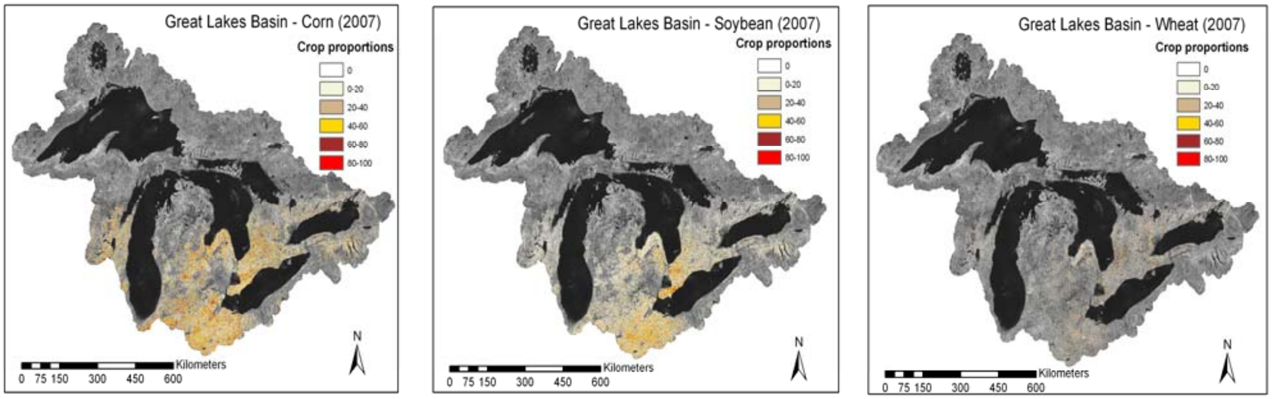

Fig. 2. Sub-pixel crop distributions for corn, soybean and wheat across the GLB.

ACKNOWLEDGMENT

The authors would like to thank David Johnson and Rick Mueller for their assistance in providing the USDA NASS data in support of this research. The U.S. Environmental Protection Agency partially funded and partially conducted the research described in this paper. Although this work was review by EPA and has been approved for publication, it may not necessarily reflect official Agency policy. Mention of any trade names or commercial products does not constitute endorsement or recommendation for use.

REFERENCES

[1] M.O. Smith, S.L. Ustin, J.B. Adams, and A.R. Gillespie, “Vegetation in Deserts.1. A Regional Measure of Abundance from Multispectral Images,” Remote Sens. Environ., vol. 31, no. 1, pp. 1-26, January, 1990.

[2] J.F. Knight, R.S. Lunetta, J. Ediriwickrema, and S. Khorram, “Regional scale land cover characterization using MODIS-NDVI 250 m multi-temporal imagery: A phenology-based approach,” GIScience & Remote Sensing, 43(1), pp. 1-23, 2006.

[3] G. De’Ath, K.E. Fabricius, “Classification and regression trees: a powerful yet simple technique for ecological data analysis,” Ecology, 81, pp. 3178–3192, 2000.

[4] M.C. Hansen, R. S. DeFries, J. R. G. Townshend, R. A. Sohlberg, C. Dimiceli, and M. Carroll, 2002b, “Towards an operational MODIS continuous field of percent tree cover algorithm: Examples using AVHRR and MODIS data,” Remote Sens. Environ., 83, pp. 303–319, 2002b.

[5] Y. Shao, R.S. Lunetta, J. Ediriwickrema, and J. Iiames, “Mapping Cropland and Major Crop Types across the Great Lakes Basin using MODIS-NDVI Data,” Photogrammetric Engineering and Remote Sensing, in press.

[6] L. Breiman, J. Friedman, R. Olshen, & C.J. Stone, Classification and regression trees. New York Chapman and Hall, 1984.

[7] W.G. Liu and E.Y. Wu, "Comparison of non-linear mixture models: Sub-pixel classification." Remote Sensing of Environment, 94(2), pp. 145-154, 20005.