Classification Techniques for Mapping Biophysical Parameters in the U.S. Southern Great Plains

Paul C. DORAI$^{SWAMY}$
U.S. Department of Agriculture
ARS, Bldg. 007, Room 008, Beltsville, MD 20705
Tel.: 301-504-6576; Fax: 301-504-5031; E-mail: pdoraisw@asrr.ars.usda.gov

Alan J. STERN
Science Systems Applications, Inc.
5900 Princess Garden Parkway, Suite 300, Lanham, MD 20706
Tel.: 301-504-5272; Fax: 301-504-5031; E-mail: stern@rsrlimg.ars.usda.gov

Paul W. COOK
USDA/NASS Spatial Analysis Research Section
3251 Old Lee Hwy., Room 305, Fairfax, VA 22030
Tel.: 703-235-5218; Fax: 703-235-3386; E-mail: pcook@nass.usda.gov

ABSTRACT

This research was part of a 1997 NASA-USDA collaborative research study for regional mapping of soil moisture in the U.S. Southern Great Plains. The study was conducted during a transition from winter to summer crops in June and July. Classification techniques were developed using Landsat TM data to separate the transitional vegetation from the natural vegetation. Areal photography and associated ground truth data for selected areas were used in developing the spectral signatures for each vegetation class. Classification of the entire region used these signatures. This research suggests techniques to be used in mapping vegetation classes in a very mixed land use area and suggests ways of using multi temporal satellite imagery for separating vegetation classes and minimizing cloud cover problems.

INTRODUCTION

Leaf Area Index (LAI) and the fraction of absorbed photosynthetically active radiation (fPAR) are important biophysical variables required in models that simulate plant growth, hydrology and ecology. These variables describe vegetation canopy structure and relate to functional process rates of energy and mass exchange in plants. An important component of this process of determining these variables is ensuring that they relate to specific land cover types. The method of doing this determination is that of classification.

This paper focuses on the procedures developed in improving the accuracy of multi temporal classification. Often, clouds can obscure specific dates of imagery and cause classification errors. However, use of even a single date for clouded areas can improve the accuracy of classification over use of multi temporal imagery [1].

After finding land cover types through classification of the multi temporal Landsat TM scenes, the evaluation of both LAI and fPAR for use as satellite derived parameters in models for calculating surface photosynthesis, evapotranspiration, and net primary production can follow. Current procedures for retrieving these parameters from satellite data are based on statistical relationships with spectral vegetation indices (VI) derived from combinations of spectral reflectance measurements. The magnitude of VI is strongly affected by the canopy architecture with its optical properties, background spectral reflectance and sun/sensor geometry. The background spectral properties produced the largest effect, resulting in large uncertainties in the VI-fPAR relationships. Thus, the use of statistical correlation between VI and fPAR at regional scales is still a matter of uncertainty.

APPROACH

An approach to retrieval and mapping of LAI and fPAR was developed by coupling remote sensing data with a crop reflectance model (CR). Crop reflectance models were developed to understand the interaction between biophysical features of the canopy, the geometry of radiometric interaction and the resulting alteration to the reflected radiation [2]. Using model inversion procedures allows calculating surface properties from the given remote sensing reflectance values.

The SAIL model [3] was used in developing the retrieval algorithm. This model has done satisfactorily predicting crop reflectance in rigorous testing during several studies. The SAIL model requires information on the leaf area index (LAI), average leaf inclination angle (ALA), single leaf reflectance and transmittance, soil reflectance, solar zenith and azimuth angles; proportions of direct and diffuse solar radiation; sensor and target geometry and date and time of satellite overpass.
This study used an iterative approach in solving for LAI values given the reflectance as compared to mathematically inverting the CR model. The CR model generates reflectance values at specific wavelength bands of the satellite. Reflectance in two wavebands corresponding to the red and near infrared of the satellite bands were simulated and the normalized difference vegetation index (NDVI) calculated from the simulated reflectance. If the generated NDVI disagrees with the satellite derived NDVI, the CR model is rerun with a new value of LAI. This process continues until an acceptable agreement with the observed reflectance is reached.

A geographical information system (GIS) permits using the algorithm for regional assessment of LAI and FPAR. An accurate digital land cover classification and soil background information are needed as input to the model. This spatially distributed information may be generated using GIS. The resulting model output can be imported back to GIS for mapping and further spatial analysis. Each land cover type is assigned its own crop and soil parameters in the model.

MODEL TESTING

The site of the Southern Great Plains 1997 (SGP97) field experiment [4] was selected to test and validate the retrieval algorithm. The procedure developed was applied on a 60x45 km area in the southern tip of the SGP97 site to generate fields of LAI and FPAR. The test area included the Little Washita Watershed. Data obtained or developed from this experiment includes the following: LAI, Biomass, and FPAR from 150 sample points, digital soils data from STATSGO, and, Land use/Land cover Map derived from classification of multi temporal Landsat TM data.

The land cover map was developed from multi spectral and multi date Landsat TM imagery using supervised classification technique. Eleven land cover types were delineated. Forage and Pasture were the dominant cover type with about 60% combined. Winter wheat (15%), trees (10%), alfalfa (5%) and summer crops (corn and legume) made up the remaining about 3%.

DATA

Landsat TM images for three dates: April 20th, June 7th and July 25th for Path 28/Rows 34-36 were collected and imported into ERDAS/IMAGINE. Ground truth was collected in mid-June for three areas in Oklahoma as follows: ARM/Central Facility, El Reno, and the Washita River basin. Crop fields within each area were chosen from aerial photography to cover a wide range of cover types. During the ground truth trip, each field was placed into a crop or land cover category. Major categories chosen were the following: alfalfa, bare soil, corn, forage, legume, pasture, trees and wheat. Most of the ground truth data was collected within the Washita River basin.

METHODOLOGY

The first analysis step was that of importing the multi temporal Landsat TM scenes into ERDAS/IMAGINE. Then, three bands were dropped from each image to leave the following bands: band 3 (0.63 - 0.69 um), band 4, band 5 and band 7 (2.08 - 2.35 um). The resulting images were mosaiced and stacked to create a single image containing 12 bands and consisting of the three rows of scenes. The crop fields from the ground truth trips were split into two groups. Approximately 2/3rds of the crop fields within each land type were used in developing signatures for classification with the remaining 1/3rd of the crop fields for verifying the classification accuracy. Additional water bodies and urban areas were chosen from maps. The training classification within the training crop fields used ISODATA reiterations. The reiterations' procedure consisted of initializing means within the bands along a diagonal axis with standard deviations set to one. Maximum iterations were set to 24 with convergence thresholds to .995, except wheat and pasture that required a convergence threshold of 0.95 due to memory limitations. As a result, "signatures" for each land type were created. The resulting land cover signatures were combined into one large signature file for use with maximum likelihood classification.

Initial classification of the multi temporal image showed some problems in areas that had clouds for more than one date. These clouded areas were filled using a single date classification. First, digital counts within Band 1 (0.45 - 0.52 um) greater than 100 were considered to be clouds and screened. In the July 25th image, this was not sufficient and digital counts within Band 1 greater than 100 and where Band 5 (1.55 - 1.75 um) values were less than 175 were screened. Shadows in all images were deemed to have digital counts less than 75 in Band 4 (0.76 - 0.90), values between 25 and 50 in Band 5 and values less than 40 in Band 3. Masks for each image were created and applied in preparation for merging in the single date classification.

The July 25th TM scene was the clearest available date for creating a single date classification using the same classification technique as in the multi temporal classification. Areas within the original classification image were masked if they had clouds either on April 20th or June 7th. The July 25th classified areas replaced the corresponding clouded areas from the classified image.

Some areas appeared to have had crops grown after wheat had been harvested. While no ground truth was available to verify this occurrence with certainty, this possibility needed to be investigated. So the assumption was made that only legumes (peanuts and soybeans) and corn (for forage) were likely to be grown after the harvest of wheat. The last two dates of TM images, July 9th and July 25th, were combined to locate these areas. The classification procedure was much the same as before, though no wheat fields were used. The
resulting classification was masked so that only corn or legumes in fields following wheat would be replaced.

Output resulting from this process consisted of two classified images (Figure 1). The first classified image was that of the original classification using three dates with 12 TM bands of data. The second image is the summer-merged classification and contained those wheat areas replaced if either corn or legumes had been grown after the wheat harvest.

RESULTS

The remaining ground truth fields (that is, the remaining 1/3rd of fields not used for classification) were converted into an image so that each land type became a zone. Thus, those fields containing alfalfa became zone one. Bare soil fields became zone two, and so forth. The summary report within ERDAS/IMAGINE created an accuracy assessment for the three classifications.

The remaining ground truth fields (that is, the remaining 1/3rd of fields not used for classification) were converted into an image so that each land type became a zone. Thus, those fields containing alfalfa became zone one. Bare soil fields became zone two, and so forth. The summary report within ERDAS/IMAGINE created an accuracy assessment for the three classifications.

The summary assessment provided the calculation of two accuracy measures: the producer's accuracy and user's accuracy. The producer's accuracy is calculated by dividing the area in the reference data correctly classified into a particular land cover by the total area in the reference data for that particular land cover. Users' accuracy for each land cover is calculated by dividing the area correctly classified within the reference data by land cover with the total area in the classified data within the reference data. So the producer's accuracy tests how well the reference data represents the classified image, while the user's accuracy tests how well the classified image matches the reference data.

The summer-merged image produced improved accuracy wheat classification in comparison with the original multi-temporal classification. Since there was no ground truth to determine if areas of wheat were harvested and corn or legumes were planted afterwards, testing the accuracy was not possible. Accuracy changes in individual crops depended upon whether the particular fields were affected by clouds. Alfalfa and legumes remain unchanged between the original classification and the merged classification, which would imply that these areas were unaffected by clouds. However, corn was obviously affected since its accuracy rose from 10% to 50%. In addition, land cover types that did not cover a large area, were more likely to change because of the changes in classification. For example, there were 28 hectares of corn, so that a change from three to 10 hectares increased the accuracy from 10% to 50%. However, land cover with large areas such as pasture were more resistant to changes in the classified image.

Land cover/landuse maps play an important role since they provide the basis for many other projects. For that reason, they must be made as accurate as possible. Other parameters such as Leaf Area Index (LAI) or fPAR are critical for crop models. Often, simply classifying an area as cloud covered is not an option. Figure 2 is a map of the land use, LAI and NDVI for the Little Washita River Basin in the southern end of the area studied gives an example of these parameters.

REFERENCES


Figure 1. Classified images for the El Reno Area (a) Original Classification with three dates, (b) Merged classification with summer crops replacing wheat
Figure 2. Parameters derived from classified image in Washita area for July 9, 1997: (a) Classified image, (b) Normalized Vegetation Index, (c) Simulated Leaf Area Index