Retrieving Ecosystem Light Use Efficiency Using Hyperion

K. Fred Huemmrich
Elizabeth Middleton
Petya Campbell
David Landis
Fluxnet Members
Introduction

We ask the question:

Can a single algorithm driven by hyperspectral satellite data provide an estimate of carbon flux variables globally?

Why use a single algorithm?

Some MODIS products (such as LAI, $f_{\text{PAR}}$, or GPP) use a land cover classification as a first step to decide the algorithm/coefficients used in the retrieval. Accuracy of MODIS Collection 5 Land Cover Type product (MCD12Q1) is estimated to be 75% http://landval.gsfc.nasa.gov/ProductStatus.php?ProductID=MOD12

Algorithms requiring land cover classification will have questionable results for the 25% mis-classified pixels
Remote Sensing of Fluxes: Hyperion and Fluxnet

• Can statistical approaches use spectral information to adjust for site differences but capture seasonal changes in flux variables?

• To address this question we examined a number of different sites with different vegetation types throughout the year.
  - Hyperion on EO-1 can provide consistent repeated observations of widely distributed sites
  - Spectra are averages of uniform regions around flux tower

• We combined satellite imagery with carbon flux data from the AmeriFlux and CarboAfrica networks
Light Use Efficiency Model

This study looks at the Light Use Efficiency ($\varepsilon$) at midday (11:00 AM to 1:00 PM local time)

\[
\text{GEP} = \varepsilon f_{\text{APAR}} \text{PAR}_{\text{in}} \quad \text{or} \quad \varepsilon = \frac{\text{GEP}}{(f_{\text{APAR}} \text{PAR}_{\text{in}})}
\]

Where:
GEP is the gross ecosystem production
PAR$\text{in}$ is the incident Photosynthetically Active Radiation (PAR)
$f_{\text{APAR}}$ is the fraction of PAR absorbed by vegetation
$\varepsilon$ is the light use efficiency, the conversion factor between energy and absorbed carbon

- In existing models $\varepsilon$ is assigned a maximum value based on cover type and downregulated based on responses to meteorological variables such as temperature and humidity
Multitemporal Data for 5 Towers

Data collected in 2008-2009 (n = 47 total)

- Duke Forest, NC, USA
  - Hardwood – n=5
  - Loblolly pine – n=5

- Mongu, Zambia
  - Miombo woodland – n=23

- Konza Prairie, KS, USA
  - Tallgrass Prairie
  - Two towers, each n=7
Seasonal Data

Seasonal change is described by Hyperion’s repeated observations

- Black points – NDVI from MODIS N-bar
- Red and Green points – NDVI from Hyperion bands convolved to MODIS bands
Spectral Vegetation Indices

- Calculated 101 SVIs from Hyperion surface reflectance
- Compared with midday LUE using data from all sites
  - TCARI performed best, $R^2=0.69$
  - Provides a baseline to compare with statistical approach

**Transformed Chlorophyll Absorption Ratio Index (TCARI)**

TCARI = $3[(R700-R670)-0.2(R700-R550)(R700/R670)]$

RMSE = 5494, $R = 0.83$

Kim et al. 1994
Partial Least Squares Regression

PLSR uses information from all spectral bands (129 bands)
- Trained using random subsets from all sites
- Produces coefficients for every spectral band

Black line is mean reflectance of all spectra
Partial Least Squares Regression

- PLSR produces great results for the training subsets
- R for all the test subsets are less than the best SVI (TCARI R=0.83)
  - Poorer RMSE for the test subset data (TCARI RMSE = 5494)
- The average of the coefficients from all the subsets does slightly better than TSAVI

Blue Bars – PLSR applied to training data subset, n=23
Red Bars – PLSR applied to test data subset, n=24
How extendable are the PLSR results?

- Use all the data for training PLSR
- Use Multisite dataset from previous study to test
  - Multisite dataset – 33 sites, \( n=79 \), only mid-growing season observations

Training Data: Multitemporal
Test Data: Multitemporal

Training Data: Multitemporal
Test Data: Multisite
What caused the differences?

- PLSR coefficients failed because of differences in Hyperion reflectances due to atmospheric correction
  - Multisite data were processed with ATREM, Multitemporal data were processed with ACORN
  - An important consideration for spectral libraries

Average reflectance of all observations of each type
PLSR can get reasonable results for all points combined

- Trained using all points from both Multitemporal and Multisite datasets

![Graph showing the relationship between Observed LUE and Modeled LUE. The equation of the linear regression line is $y = 1x + 0.0213$ with $R^2 = 0.78863$. The data points are differentiated by color: red for Multitemporal, black for Multisite, and the entire dataset combined is denoted as 'All Pts'.]
Conclusions

• A general approach for retrieval of biophysical variables without classification may yet be possible
  - Information from multiple spectral bands is required

• Statistical approaches like PLSR can be powerful tools for data analysis and product generation
  - Critically important to have good training data that represents the full range of cases
  - Differences in processing approaches may result in significant errors

• Hyperion’s ability to provide multitemporal and multisite observations makes it an important tool for pre-HyspIRI algorithm development and testing