Special Issue Foreword

Research Advances in Hyperspectral Remote Sensing

Prasad S. Thenkabail

Hyperspectral data provides substantially increased understanding of plant biophysical and biochemical properties relative to multispectral broadband data. Accuracies in classifying, modeling, mapping, and monitoring are substantially higher when specific hyperspectral narrowbands (HNBs) and hyperspectral vegetation indices (HVIs) are used as opposed to multispectral broadbands. Even though this is now a well-established fact, there is still significant knowledge gap in our understanding of the importance hyperspectral data in study of agricultural crops and vegetation (Thenkabail et al., 2011). Indeed, opportunities exist for making significant knowledge advances in several areas of hyperspectral study of vegetation and agricultural crops such as in: 1. Establishing specific HNBs and HVIs to quantify biophysical and biochemical properties, 2. Overcoming Hughes' phenomenon and data redundancy, 3. Building hyperspectral libraries of crops and vegetation, and 4. Developing advanced automated methods of hyperspectral data analysis. Also, increasing amounts of hyperspectral data (e.g., entire archive of ~64,000 images of EO-1 Hyperion available from USGS Earthexplorer (http://earthexplorer.usgs.gov/);see highlight article in this issue for details) are becoming available, globally, for researchers around the world to conduct specific studies on specific issues in different croplands and vegetation of the world. Given this fact, the need for focused research to better understand, model, and map specific vegetation and agricultural crop characteristics utilizing hyperspectral data is of great importance. In this context, PE&RS initiated this special issue on the topic. The special issue covers some of these advances through seven distinct peer-reviewed articles. Highlights of these seven peer-reviewed articles along with their key knowledge advancement are summarized below.

1.0 Automated Hypercor Algorithm to Process Large Volumes of Hyperspectral Data and Identify Important Bands to Model Biophysical and Biochemical Quantities and a Novel Concept of Multi-Correlation Matrices Strategy (MCMS) to Select and use Hyperspectral Narrowbands (HNBs) in Hyperspectral Vegetation Indices (HVIs)

Aasen et al. developed automated algorithms to process large volumes of hyperspectral data in order to determine which hyperspectral narrowbands (HNBs) and\or hyperspectral vegetation indices (HVIs) hold the best information. They developed an impressive algorithm called HyperCor to process large volumes of hyperspectral data and discern their information pathways. HyperCor takes hundreds or thousands of HNBs, computes their two-band or multi-band HVIs, correlates with biophysical and biochemical quantities of vegetation or agricultural crops, and shows us the windows or regions in electromagnetic spectrum providing high and low information content. This is exactly the type of tool that we need to make best and most efficient use of hyperspectral data in applications such as the vegetation, and the agricultural crops. Their study on rice crop was conducted with 5 years of solid data (3 years for model development and 2 years for validation). In selecting best HNBs, it must be noted that the HNBs which are most prominent for one biophysical quantity may not be the most prominent for another biophysical quantity. At times, it may not even be most prominent for the same biophysical quantity in another date. This phenomenon happens as a result of having narrowbands adjacent to one another providing near similar information (e.g., 680 nm and 690 nm are likely to have equally good correlation with biomass). This will require us to select the most prominent narrowband in each spectral range (e.g., a band centered at 680 nm, with 10 nm bandwidth, could be most frequently occurring narrowband in modeling biophysical and biochemical properties of vegetation and agricultural crop in 600 to 700 nm band range). Overall, they clearly established that automated algorithms like HyperCor are extremely valuable in analyzing biophysical and biochemical variables of massive volumes of hyperspectral data of agricultural crops and vegetation. They also introduced a novel concept of multi-correlation matrices strategy (MCMS) to select and use HNBs in HVIs. The idea here is to source the importance of HNBs based on their significant occurrence in different correlation matrices (CMs). They showed that MCMS provided significantly improved accuracies in studying rice biomass. MCMS is an interesting concept, but requires further development. Aasen et al. developed their models using various approaches: pooled data of various growth stages, individual growth stages, linear models, and non-linear models. It must be noted that robust models of biophysical and biochemical properties of specific crops need to be developed, ideally, taking data of across sites and across growing stages. Such models are also often nonlinear in nature as a result of saturation in reflectivity in full canopy cover scenario. Also, such models need to be developed for individual crops rather than grouping multiple crop types in models to achieve best results that are specific and targeted.

2.0 Automatic Labeling of Classes through Spectral Matching Techniques (SMTs) by using Ideal Spectra fromSpectral Library

It is well known that hyperspectral data is not panacea for addressing complex issues of cropland and vegetation classification, modeling, and characterization. This is because, even though hyperspectral data provides a quantum leap in information, discerning that information is not an easy task given the complexities of processing massively large data volumes, establishing redundant bands, and implementing methods and techniques that accurately and rapidly establish information from data. In this regard the work of Parshakov et al. is invaluable. They have implemented an innovative spectral matching technique (SMT; also see Thenkabail et al., 2007 for concept of SMTs) approach that automatically determines and labels crop types by matching the class spectral signatures with the ideal or reference spectral libraries developed using hyperspectral Hyperion data. The Z-score distance SMT that they use accounts for the variation of pixel spectra by measuring the distance between the class spectra and the reference spectra in units of standard deviation. They used this SMT method to identify and label 11 agricultural crops classified using Landsat TM data. Their study established that the accuracies of their automated SMT provided: A. 6 to 11 percent greater accuracies relative to ISODATA classification followed by manual identification of classes, and B. 12 percent greater than the spectral angle mapper (SAM) followed by manual identification of classes. However, the accuracies were 2 to 12% lower than the Maximum Likelihood Classification followed by manual class identification. Their method can be used to automatically identify and label classes classified using multispectral or hyperspectral data. The automated SMT by Parshakov et al. is novel and clearly demonstrates pathway to identify and label crop types and other land use classes automatically saving time and removing user bias. However, there will be complexities of applying automated SMTs over large and complex areas. Nevertheless, by building adequate and accurate ideal or reference spectral libraries of agricultural crops, vegetation categories, and other land use classes as well as by further development of automated SMTs for specific regions of the world, automated class labeling proposed and demonstrated by Parshakov et al. will become accurate, unbiased, rapid, and widely implementable.

3.0 Methods for Data Dimensionality Reduction and Overcoming Hughes' Phenomenon

The study by Nadiminti et al. address the important issue of high dimensionality of hyperspectral data, ways and approaches to overcome them, and the benefit of doing so to overcome Hughes Phenomenon (Note: Hughes phenomenon means that when the dimensionality of data increases, the training sample number should also increase in order to maintain precision of classification, alternatively we need to increase the number of training samples which can be often resource prohibitive. Thereby, in order to process hyperspectral data effectively, it is necessarily to reduce the dimensionality of hyperspectral data or increase the sample number of training data used in classification. Including highly correlated bands (e.g., R-square >0.9) in analysis either makes no difference to classification accuracies or, many a times, actually leads to decrease in classification accuracies. This is because highly correlated bands provide same, duplicate, information whereas the training samples remain the same in spite of increase in number of bands. Nadimiti et al. used hyperspectral Hyperion images of three seasons (Monsoon, winter, and summer) over tropical forests to classify and separate three species: Teak, Bamboo, and mixed forests. Data dimensionality reduction was explored using Kernel Principal Component Analysis (k-PCA), Independent Component Analysis (ICA), and Principal Component Analysis (PCA). Their results re-enforced the recent findings elsewhere (Thenkabail et al., 2013) that 4 to 8 % hyperspectral narrowbands (HNBs) provide optimal results, leaving the rest of the bands redundant. In kPCA, for example, 10 kernel principal components or HNBs, selected based on eigenvectors (factor loadings), explained 99% variability in data when 179 Hyperion HNBs were used in analysis. Thereby, the study establishes the fact that, often, HNBs that adjoin one another are redundant for a given application. Thereby, identifying redundant bands help overcome Hughes' Phenomenon.

4.0 Significantly Improved Vegetation Classification Accuracies by Combining Hyperspectral Data with LiDAR

Hyperspectral data is in itself a great advancement over broadband multispectral data. This is now an established fact (Thenkabail et al., 2011). Nevertheless, there is considerable scope for improvement in our understanding of agricultural crops and vegetation communities by combining multiple sources of remote sensing data. This aspect is well illustrated by Zhang et al. in their paper on studying wetland vegetation communities of Florida everglades by combining hyperspectral data with Light Detection and Ranging (LiDAR) data. They studied 13 common everglades vegetation communities using 224 band hyperspectral Airborne Visible/ Infrared Imaging Spectrometer (AVIRIS) data acquired at 12 m spatial resolution and a Leica ALS-50 LiDAR system collecting small footprint multiple returns, and intensity at 1060 nm wavelength with average point density for the study area of 1.18 pts/m^2. They showed by fusing the hyperspectral and LiDAR data and using the 3 machine learning algorithms [Random Forest (RF), Support Vector Machine (SVM), and *k*-Nearest Neighbor (*k*-NN)] it is possible to increase the overall accuracy by as much as 10% (from 76% overall accuracy using Hyperspectral data alone to 86% when both Hyperspectral and LiDAR are used) in classifying 13 everglade wetland vegetation communities.

5.0 Advances in Soil Moisture Retrieval from Irrigated-, Rainfed-, and Fallow Agricultural Farmlands by Combining Hyperspectral Data with Thermal and Radar Data

Sanchez et al. combined hyperspectral data with thermal and radar data to retrieve soil moisture from agricultural fields. They used data from Compact Airborne Spectrographic Imager (CASI 550) sensor and Thermal Airborne Spectrographic Imager (TASI 600) and combined with Airborne L-band (ARIEL-2) to retrieve soil moisture from irrigated-, rainfed-, and fallow- farmlands that include cereals, sunflower, vineyards, and fallow-farmlands. They showed that hyperspectral bands and indices (HNBs and HVIs) had significantly better correlation with observed soil moisture when integrated with the land surface temperature (LST) and brightness temperature (BT) rather than when they were used alone. They specifically recommend indices derived using hyperspectral wavebands in the red-edge and near infrared rather than visible. Even through microwave L-band data is widely used for soil moisture retrieval, using that data along with hyperspectral data has significant advantages.

6.0 Significantly Increased Classification Accuracies of Pine Forests using Hyperspectral Narrowbands as Opposed to Multispectral Broad-bands

Awad et al. clearly establish that classification accuracies can be substantially increased using hyperspectral narrowband data as opposed to multispectral broadband data. They used CHRIS PROBA hyperspectral data to classify Stone Pine forests and compare the results with Landsat ETM+ classified results. Using the 63 band spaceborne CHRIS PROBA hyperspectral data they were able to establish an increased accuracy of as much as about 30%. The producer's, User's, and overall accuracies in classifying stone pine forests using CHRIS PROBA were 90% or higher whereas using 6 non-thermal Landsat ETM+ these accuracies were around 60%.

7.0 Significantly Improved Biomass Modelling of Four Leading World Crops using Hyperspectral Narrowbands (HNBs) and hyperspectral Vegetation Indices (HVIs)

Hundreds of HNBs and HVIs were used model above-ground biomass of 4 leading world crops (rice, wheat, corn, alfalfa) based on 2 years of detailed data acquired for these crops in the irrigated agricultural fields of California by Marshall and Thenkabail. The best biomass models explained greater than 80% variability using highly selective sequential search methods (SSM) involving two-band HVIs or multi-band HVIs involving one to 3 HNBs. The key is also to select specific narrowbands (~10 nm or less) from two or three distinct portion of the spectrum: (a) green and near-infrared, (b) blue and NIR, (c) near-infrared (NIR) and short-wave infrared (SWIR), or (d) green, NIR, and SWIR. These specific HNBs may change for crop to crop and even within crop. But, what needs to be noted is that there are some very selective HNBs and HVIs derived off them (see the Table 2 and 3 in the highlight article of this issue) which consistently perform highly across different crops and their varying characteristics. The HNBs and HVIs vary because selecting one HNB versus another often makes only a slight difference (e.g., 680 nm or 690 nm are highly correlated and perform about the same; similarly 855 nm or 910 nm are highly correlated and perform similarly; a point also noted in other reported studies). But, in modelling a specific crop an HVI involving 855 nm and 680 nm (HVI855680) may perform marginally better than an HVI involving 910 nm and 690 nm (HVI910690). This performance may, at times differ for another crop. This does not mean that we need to use both the indices. It will suffice to use a single index (e.g., HVI855680) to model both crops because the two HVIs are equally good (e.g., one index may have an R-square of 0.85 with biomass and another 0.87; in which case we will select the one with 0.87 and ignore the one with 0.85). The study by Marshall and Thenkabail, re-affirms the fact that there are redundant bands as well as there are specific HNBs that are of highest importance to model specific biophysical and biochemical characteristics of crops or vegetation. These HNBs and HVIs perform significantly better than any known broadband derived indices. Readers should refer to various Tables and figures of the paper by Marshall and Thenkabail for better understanding. Further, greater, comprehensive understanding can be acquired by going through Table 2, 3, and 4 as well as Figure 4a and 4b of the highlight article in this special issue.

Hyperspectral remote sensing (or Imaging Spectroscopy) is fast moving from an era of research into an era of applications. Many spaceborne hyperspectral sensors (e.g., HypspIRI, UAV based platforms, interest from private entities; see Thenkabail et al., 2011) are planned in near future. This special issue adds to maturing knowledge of hyperspectral remote sensing in general, and hypersepctral remote sensing of vegetation and agricultural crops in particular.

Credit to this special issue on “**Hyperspectral Remote Sensing of Vegetation and Agricultural Crops**” goes to several people. I would like to thank the authors for their outstanding work. Each paper was reviewed by at least 3 reviewers. Good reviewers are few but pivotal for success of any quality journal. I am thankful to many good reviewers who helped improve the quality of each paper. I am grateful for the advice, support, and guidance of Dr. Russell Congalton, Editor in Chief of PE&RS. Ms. Jeanie G. Congalton, PE&RS manuscript coordinator, was always there with her insights. Finally, I would like to thank the U. S. Geological Survey (USGS), especially USGS Western Geographic Science Center (WGSC) and its leadership, for all the opportunities and encouragement that I have received over the years.

7.0 References:

Middleton, E.M.; Ungar, S.G.; Mandl, D.J.; Ong, L.; Frye, S.W.; Campbell, P.E.; Landis, D.R.; Young, J.P.; Pollack, N.H., 2013. “The Earth Observing One (EO-1) Satellite Mission: Over a Decade in Space,” *Selected Topics in Applied Earth Observations and Remote Sensing, IEEE Journal of* , vol.6, no.2, pp.243,256, April 2013. doi: 10.1109/JSTARS.2013.2249496

Thenkabail, P.S., Lyon, G.J., and Huete, A. 2011. Book entitled: “Hyperspectral Remote Sensing of Vegetation”. CRC Press- Taylor and Francis group, Boca Raton, London, New York. Pp. 781 (80+ pages in color).

Reviews of this book: http://www.crcpress.com/product/isbn/9781439845370

Thenkabail, P.S., Mariotto, I., Gumma, M.K.,, Middleton, E.M., Landis, and D.R., Huemmrich, F.K., 2013. Selection of hyperspectral narrowbands (HNBs) and composition of hyperspectral twoband vegetation indices (HVIs) for biophysical characterization and discrimination of crop types using field reflectance and Hyperion/EO-1 data. IEEE JOURNAL OF SELECTED TOPICS IN APPLIED EARTH OBSERVATIONS AND REMOTE SENSING, Pp. 427-439, VOL. 6, NO. 2, APRIL 2013.doi: 10.1109/JSTARS.2013.2252601.

Thenkabail, P.S., GangadharaRao, P., Biggs, T., Krishna, M., and Turral, H., 2007. Spectral Matching Techniques to Determine Historical Land use/Land cover (LULC) and Irrigated Areas using Time-series AVHRR Pathfinder Datasets in the Krishna River Basin, India. Photogrammetric Engineering and Remote Sensing. 73(9): 1029-1040.

Special Issue Editor

Dr. Prasad S. Thenkabail
Research Geographer
U. S. Geological Survey
Email: pthenkabail@usgs.gov; Thenkabail@gmail.com