MAPPING LAND COVER WITH HYPERSPECTRAL AND MULTISPECTRAL SATELLITES USING MACHINE LEARNING AND SPECTRAL MIXTURE ANALYSIS

Matthew L. Clark

Department of Geography and Global Studies, Sonoma State University, Rohnert Park, California, 94928, USA

ABSTRACT

The goal of this study was compare hyperspectral and multispectral imagery for mapping broad land-cover classes at the spatial scale of a satellite image. The study area was the San Francisco Bay Area and was roughly the size of a Landsat scene (30,000 km^2). The Random Forests machine learning and Multiple-Endmember Spectral Mixture Analysis (MESMA) classifiers were compared to predictor variables composed of simulated HyspIRI hyperspectral images, simulated Landsat 8 and Sentinel-2 multispectral images, and real Landsat 8 images. The Random Forests machine learning classifier consistently outperformed MESMA and there were significant improvements in overall accuracy with multi-temporal (spring, summer, fall) over summer-only images for all sensors tested. Hyperspectral reflectance data had no difference to less accuracy relative to comparable multispectral datasets. However, HyspIRI hyperspectral metrics that targeted key spectral features, related to chemical and structural properties, yielded significantly improved accuracy over both real and simulated multispectral datasets.

Index Terms— Land cover, hyperspectral, multispectral, multi-temporal, Random Forests, Multiple-endmember Spectral Mixture Analysis

1. INTRODUCTION

Maps of land cover are essential to global climate modeling and can be used to better understand underlying natural and socio-economic processes driving land change. Most land-cover maps at regional (>10,000 km^2) to global scales are produced with remote sensing techniques applied to multispectral satellite imagery. Hyperspectral, or imaging spectrometer, sensors measure reflected solar radiance over hundreds of narrow, contiguous spectral bands. Applications using these data have shown encouraging results in mapping the properties of both terrestrial and aquatic ecosystems (reviewed in [9]); however, due to cost constraints, most studies cover relatively small spatial extents (<1,000 km^2) with single-season airborne imagery. A hyperspectral satellite is needed to fully scale and compare results to multispectral satellites.

NASA's planned Hyperpectral Infrared Imager (HyspIRI; [9]) will also have 30-m, visible through shortwave infrared (VSWIR) imagery with relatively high SNR and a similar image extent and revisit time to Landsat. To advance the satellite mission's science, algorithms and processing readiness, NASA funded the HyspIRI Preparatory Science Campaign. This project provided simulated HyspIRI imagery using images acquired from NASA's Airborne Visible Infrared Imaging Spectrometer (AVIRIS) "Classic" sensor flown over large regions in California in spring, summer and fall from years 2013 to 2015 [9].

The goal of this study was to compare the performance of hyperspectral and multispectral satellite sensors for mapping broad land-cover classes at the regional scale of the San Francisco Bay Area, California. Specifically, HyspIRI Preparatory Campaign simulated images were compared to multispectral imagery simulated using Landsat 8 and Sentinel-2 broadband characteristics, but at the same spatial scale and radiometric properties as HyspIRI. Map accuracy from simulated imagery was also compared to actual Landsat 8 imagery. This study was designed with a method that could scale to broader regional and global scales.

2. METHODS

2.1. Study area

Imagery in this study covered a box centered (37° 52' 45.73" N, 122° 13' 8.54" W) on the San Francisco Bay Area in northern California, USA. The region has a mostly Mediterranean climate with wet winters and dry summers.

2.2. Land-cover reference data

Land-cover reference data collection was facilitated with an automated Web-based system called the Virtual Interpretation of Earth Web-Interface Tool (VIEW-IT; [2]). This tool provides a platform for human interpretation of land-cover components from natural color high-resolution imagery using the Google Earth browser plugin. Interpretation of land cover was in 100-, 250- or 500-m square polygons in areas of well-mixed cover. Percent cover for each polygon was classified into a Land Cover Classification System (LCCS; [7]) class using modified rules

Authorized licensed use limited to: Nirma University Institute of Technology. Downloaded on September 17,2022 at 06:02:08 UTC from IEEE Xplore. Restrictions apply.

leading to 12 map classes. For forest and shrubland classes, only samples with >65% cover (closed-canopy) we considered in this analysis. Final classes include: Evergreen needleleaf trees (conifer), Evergreen broadleaf trees (hardwood), Deciduous broadleaf trees (hardwood), Shrubslands, Herbaceous, Dune vegetation, Tidal salt marsh, Annual crops, Perennial crops, Built-up, Urban Vegetated, and Bare.

2.3. HyspIRI Preparatory Science data products

Hyperspectral imagery used in this study were collected as part of the NASA HyspIRI Preparatory Science Campaign and included year 2013 spring, summer and fall scenes from the AVIRIS Classic sensor [9]. The AVIRIS-C sensor collects spectral radiance data over 224 bands from 370 to 2500 nm with 10 nm sampling. The sensor was flown on a NASA ER2 jet at 65,000 ft (~20 km) over a ~30,000 km^2, an area roughly the size of a Landsat satellite image. There were 11 nominal flight runs over the study area. Each run had a ~12 km swath width and flight paths provided 20% image overlap among runs.

Orthorectified surface reflectance simulated HyspIRI products were downloaded from the NASA Jet Propulsion Laboratory website (http://aviris.jpl.nasa.gov). In simulated HyspIRI, JPL spatially resamples at-sensor radiance data and adds a noise function to approximate the potential noise equivalent delta radiance function of HyspIRI. Apparent surface reflectance was retrieved by JPL using an improved version of ATREM [12].

2.4. Hyperspectral metrics

Simulated HyspIRI VSWIR reflectance products were processed to create a suite of 86 "hyperspectral metrics" that respond to vegetation chemistry and structure. Metrics were based on narrowband ratio (indices), spectral derivative, and absorption-feature fitting techniques that spanned the VSWIR spectral regions. See [3, 4] for a more thorough description of these metrics.

2.5. Landsat-8 and Sentinel-2 imagery

Hyperspectral-based mapping accuracy was tested against that from of multispectral, Landsat 8 OLI and Sentinel-2 MSI satellite sensors. To control for geometric, radiometric and bidirectional reflectance distribution function (BRDF) factors, Landsat 8 and Sentinel-2 data were simulated from HyspIRI radiance and reflectance data (includes noise) using sensor-specific filter functions. Real Landsat 8 images were also analyzed. These data were calibrated to surface reflectance using an empirical line correction method.

2.6. Land-cover classification

The following eight predictor variable groups were used in classification analyses:

- Simulated HyspIRI: reflectance bands and hyperspectral metrics.
- Simulated Sentinel-2: radiance and reflectance bands.
- Simulated Landsat 8 OLI: radiance and reflectance bands.
- Real Landsat 8 OLI: radiance and reflectance bands.

All groups were tested using summer and stacked multi-temporal (spring, summer, fall) data. Predictor variables were extracted for pixels within VIEW-IT polygons. Extracted pixel data were randomly split in half at the polygon level for training and testing for each class.

2.6.1. Random Forests classification

The machine learning algorithm Random Forests [1] was one classifier chosen for mapping. The training dataset for each predictor variable group was first filtered at the PSU level to remove outliers using Euclidean distance from the Euclidean mean. For HyspIRI-like data, filtered training data were processed with the VSURF routine to select an optimal subset of predictor variables [8]. A final RF classier was trained with parameters: 999 decision trees, node size of 10, and SQRT(number of variables) randomly sampled as candidates at each split. Confidence maps indicate the percentage of votes for a given class out of the 999 total RF votes. Classified scenes were mosaicked together using an overlap rule that selected the class with the highest confidence.

2.8.2. MESMA classification

The Multiple Endmember Spectral Mixture Analysis (MESMA) classifier [10] was compared to the RF classifier. MESMA was run with fractions constrained to a minimum (-5%) and maximum (105%), and the fit required to meet a maximum RMSE (2.5% reflectance). Photometric shade was not allowed to exceed 80%. Training spectra were optimized in order to obtain a parsimonious MESMA endmember library that were spectrally-pure for each class, while being the most distinct from other classes. First, starting with 12-class training data, polygons with <90% cover removed from analysis. Summer and multi-temporal spectra were reduced at the polygon level with Endmember Average RMSE (EAR; [6]) and Minimum Average Spectral Angle (MASA; [5]), respectively, and the final library was further optimized with the Iterative Endmember Selection (IES) process [11]. In MESMA, one to three endmembers were combined with photometric shade to create 2- to 4-EM mixture models.

For each scene, MESMA produces a map of fractional abundance for each land-cover class (i.e., 12 classes) and a model RMSE. Fractions were shade normalized and runs were mosaicked together by selecting pixels with the lowest RMSE in areas of overlap. Each set of shade-normalized

Authorized licensed use limited to: Nirma University Institute of Technology. Downloaded on September 17,2022 at 06:02:08 UTC from IEEE Xplore. Restrictions apply.

fraction images was then mapped to 12 discrete cover classes using LCCS rules.

2.9. Accuracy assessment

The accuracy assessment of maps for each predictor variable group was assessed using independent test samples. Test pixels were sampled from the mosaicked map products (i.e., no duplicates in scene overlap).

This section focuses on classifications with RF to investigate the methodological factors of temporal resolution, atmospheric correction and spectral resolution.

Temporal resolution: There were significantly higher OA with multi-temporal (spring, summer, fall) than single-season (summer) datasets, for all multispectral and hyperspectral sensor comparisons (Table 1, Fig. 1).

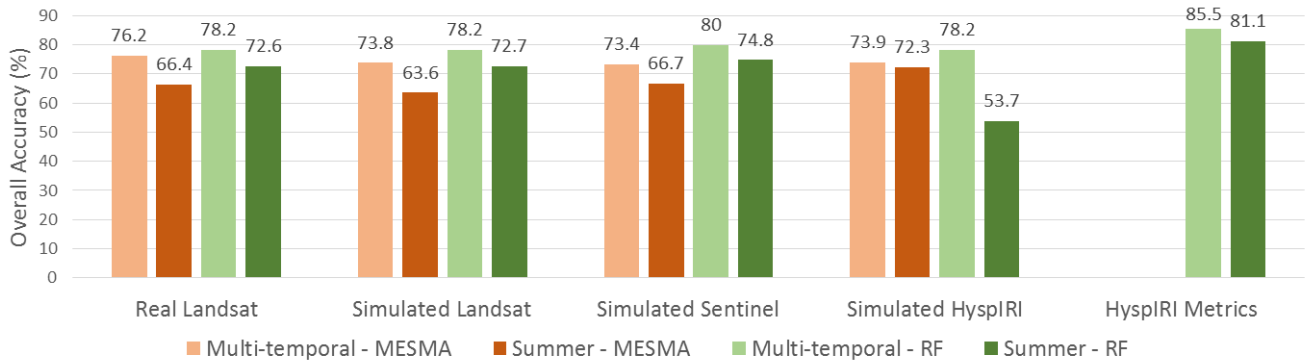

Fig. 1. Overall accuracy with real Landsat OLI and simulated HyspIRI, Landsat OLI and Sentinel-2 data. Comparison includes maps produced with the Random Forests and MESMA classifiers with summer or three-season imagery, respectively.

3. RESULTS AND DISCUSSION

3.1. Random Forests vs. MESMA classifications

With the exception of the summer simulated HyspIRI comparison, RF overall accuracy (OA) was 2.0 to 9.1% higher than MESMA for equivalent predictor variables (Fig. 1). The HyspIRI hyperspectral metrics classification with RF had OA that was 8.8% and 11.6% higher than with the HyspIRI reflectance MESMA with summer and multi-temporal data, respectively. Based on these results, the RF machine learning algorithm is a preferable classifier than MESMA, a more common technique in the hyperspectral remote sensing community.

With the exception of the summer simulated HyspIRI comparison, RF overall accuracy (OA) was 2.0 to 9.1% higher than MESMA for equivalent predictor variables (Fig. 1). The HyspIRI hyperspectral metrics classification with RF had OA that was 8.8% and 11.6% higher than with the HyspIRI reflectance MESMA with summer and multi-temporal data, respectively. The MESMA classifier was also involved a much longer processing time than equivalent RF processing. For example, a single multi-temporal reflectance run in MESMA could take three days to process but on 7 hours in RF. Based on these results, the RF machine learning algorithm is a preferable classifier than MESMA for broad land-cover mapping of satellite imagery.

3.2. Random Forests classification comparisons

Table 1. Accuracy of classified mosaicked maps from the Random Forests classifier. Results are organized by predictor variables used (Landsat, Sentinel), temporal resolution (Summer or Multi-Temporal), and data type (Radiance or Reflectance).

	Real Landsat		**Simulated Landsat**		**Simulated Sentinel**	
Data type	*S*	*MT*	*S*	*MT*	*S*	*MT*
Radiance	72.6	78.2	70.9	77.6	76.4	79.9
Reflectance	72.5	80.5	72.7	78.2	74.8	80.0

Atmospheric correction: We expected atmospheric correction and surface reflectance retrieval to improve accuracy over using raw radiance data. For classifications with multispectral sensors, there were -1.6% to +1.9% differences in OA when comparing reflectance to radiance bands for equivalent sensors and temporal resolution combinations (Table 1). Although these differences are mostly statistically significant ($p < 0.05$), there was no clear evidence that atmospheric correction improved accuracy.

Spectral resolution: There was a 1.8% to 21.1% significant decrease ($p < 0.05$) in OA when using HyspIRI reflectance bands relative to Landsat or Sentinel bands simulated from HyspIRI reflectance, with the exception of multi-temporal comparisons against real and simulated Landsat, which had no difference. In contrast, there was 5.5% to 8.5% greater OA with HyspIRI hyperspectral metrics relative to multispectral bands with comparable temporal resolutions ($p < 0.00001$).

Authorized licensed use limited to: Nirma University Institute of Technology. Downloaded on September 17,2022 at 06:02:08 UTC from IEEE Xplore. Restrictions apply.

The advantage of hyperspectral imagery was thus only found when pre-processing the reflectance bands to metrics, which isolate chemical and structural features in the spectra.

Variable selection: As found in [3], HyspIRI hyperspectral metrics had significantly better accuracy than using reflectance bands in RF. In this case, OA with metrics was 7.3% improved for multi-temporal data and reached 85.5%. There were 35 of 258 variables selected by VSURF for prediction in the final RF. In [3], no VSURF variable selection was used prior to training the RF and OA was only slightly less, at 85.1% (no significant difference). Furthermore, the VSURF routine took 3 days to run for multi-temporal metrics, despite multi-threaded processing. In summary, there was no clear advantage in a variable optimization step prior to training the RF.

5. CONCLUSION

The goal of this study was compare hyperspectral and multispectral imagery for mapping broad land cover classes at the spatial scale of a satellite image. The San Francisco Bay Area study area was roughly the size of a single Landsat scene (30,000 km^2). The following conclusions came from this analysis:

- The Random Forests machine learning classifier consistently outperformed MESMA, a favored algorithm in the hyperspectral community.
- There were significant improvements in overall accuracy with multi-temporal (spring, summer, fall) over summer-only images for all sensors tested. This result indicates that repeat passes from satellite sensors are important for land-cover classification, regardless of sensor spectral resolution.
- Multispectral imagery was simulated from hyperspectral imagery so as to control for all factors not related to spectral differences. Atmospheric correction and surface reflectance retrieval from hyperspectral imagery prior to simulation did not offer clearly improved accuracy over using raw radiance data.
- Hyperspectral reflectance data had no difference to less accuracy relative to comparable multispectral datasets. However, hyperspectral metrics had significantly improved accuracy over both real and simulated multispectral datasets.
- Land-cover maps from multi-temporal, HyspIRI hyperspectral metrics had overall accuracy of 85.5%. This represented a 7.3% significant improvement in accuracy over real, multi-temporal Landsat 8 OLI imagery.

Acknowledgements

Funding for this project was provided by NASA HyspIRI Preparatory Airborne Activities and Associated Science Research, grant NNX12AP09G.

7. REFERENCES

[1] Breiman, L. “Random Forests”, Machine Learning, 45, 5-32, 2001.

[2] Clark, M.L., and Aide, T.M. “Virtual Interpretation of Earth Web-Interface Tool (VIEW-IT) for collecting land-use/land-cover reference data”. Remote Sensing, 3, 601-620, 2011.

[3] Clark, M.L. and Kilham, N.E.. “Mapping of land cover in Northern California with simulated hyperspectral satellite imagery”. ISPRS Journal of Photogrammetry and Remote Sensing, *revised*.

[4] Clark M.L. and Roberts D.A. “Species-level differences in hyperspectral metrics among tropical rainforest trees as determined by a tree-based classifier”. Remote Sensing, 4, 1820-1855, 2012.

[5] Dennison, P.E., Halligan, K.Q., and Roberts, D.A. “A comparison of error metrics and constraints for multiple endmember spectral mixture analysis and spectral angle mapper”. Remote Sensing of Environment, 93, 359-367, 2004.

[6] Dennison, P.E., and Roberts, D.A. “Endmember selection for multiple endmember spectral mixture analysis using endmember average RMSE”. Remote Sensing of Environment, 87, 123-135, 2003.

[7] Di Gregorio, A. “Land cover classification system: classification concepts and user manual: LCCS (No. 8)”. Food and Agriculture Organization, 2005.

[8] Genuer, R., Poggi, J.M. and Tuleau-Malot, C., “Variable selection using random forests”. Pattern Recognition Letters, 31, 2225-2236, 2010.

[9] Lee, C.M., Cable, M.L., Hook, S.J., Green, R.O., Ustin, S.L., Mandl, D., et al. “An introduction to the NASA Hyperspectral InfraRed Imager (HyspIRI) mission and preparatory activities”. Remote Sensing of Environment, 167, 6-19, 2015.

[10] Roberts, D.A., Gardner, M., Church, R., Ustin, S., Scheer, G., and Green, R.O. “Mapping chaparral in the Santa Monica Mountains using multiple endmember spectral mixture models”. Remote Sensing of Environment, 65, 267-279, 1998.

[11] Schaaf, A.N., Dennison, P.E., Fryer, G.K., Roth, K.L., and Roberts, D.A. "Mapping plant functional types at multiple spatial resolutions using imaging spectrometer data”. GIScience and Remote Sensing, 48(3), 324-344, 2011.

[12] Thompson, D.R., Gao, B.C., Green, R.O., Roberts, D.A., Dennison, P.E., & Lundeen, S.R. “Atmospheric correction for global mapping spectroscopy: ATREM advances for the HyspIRI preparatory campaign”. Remote Sensing of Environment, 167, 64-77, 2015.

Authorized licensed use limited to: Nirma University Institute of Technology. Downloaded on September 17,2022 at 06:02:08 UTC from IEEE Xplore. Restrictions apply.