Correlation Among Simple Band Ratios in a Semi-Arid Environment

Alyssa Farnes and Keith T. Weber, ISU GIS Training and Research Center

Abstract

Simple Band Ratios (SBR) calculated from multispectral imagery provide quantitative indices of environmental factors, and when collected in a time series, can be used to monitor landcover change. Landsat 5 and Landsat 8 imagery were obtained to produce Normalized Difference Vegetation Index (NDVI), Modified Soil-Adjusted Vegetation Index Two (MSAVI2), and Normalized Difference Moisture Index (NDMI) time series data. These data were compared across two study areas in Idaho, USA. Sce3nes captured two semi-arid environments with differing vegetation types across four nonadjacent growing seasons. Correlation was strong between NDVI and MSAVI2 ($\bar{x}$ R^2= 0.96; P≤ 0.01) in two of four study years with the sagebrush steppe-dominated landscape of eastern Idaho, while the cheatgrass-dominated landscape of western Idaho showed much lower correlation. There was no significant correlation between NDVI and NDMI or MSAVI2 and NDMI (P > 0.01). These results indicate that while NDVI or MSAVI2 may be used effectively to understand vegetation characteristics in semi-arid environments, using both SBRs is redundant and introduces statistical pseudoreplication. Further, NDMI describes an altogether different component of the ecosystem, namely moisture content.

Introduction

Remote imagery acquired over time at a meaningful spatial resolution can be useful to land managers for monitoring changes in vegetation productivity and health. This can be accomplished by developing an index derived from remotely sensed data. The Normalized Difference Vegetation Index (NDVI, Eq. 1) is a popular Simple Band Ratio (SBR) for photosynthetically-active biomass monitoring (Tucker, 1979). In addition, the Modified[1] Soil Adjusted Vegetation Index Two (MSAVI2, Eq. 2) considers the influence of soil reflectance on the spectral model (Qi, 1994) and is particularly useful in arid/semi-arid environments where there is

[1] The Soil-Adjusted Vegetation Index (SAVI) was developed by A.R. Huete in a paper published in 1988. Six years later in 1994, J. Qi published the MSAVI paper with Huete, et al. The modification comes from replacing the constant *L* with a variable L function to vary inversely with the amount of vegetation present.

pronounced soil exposure. The Normalized Difference Moisture Index (NDMI, Eq. 3) quantifies moisture in vegetation and is useful to monitor wildfire fuels (Wilson, 2002). Because each SBR has complimentary uses, the purpose of this paper was to test correlation and potential redundancy among these SBRs.

The equation used to obtain NDVI: $NDVI = \frac{NIR-RED}{NIR+RED}$ (Rouse, 1974). (Eq. 1)

The equation used for MSAVI2: $MSAVI_2 = \frac{2\rho_{nir} + 1 - \sqrt{(2\rho_{nir} + 1)^2 - 8(\rho_{nir} - \rho_{red})}}{2}$ (Qi, 1994). (Eq. 2)

The equation used to obtain NDMI: $NDMI = \frac{NIR-SWIR1}{NIR+SWIR1}$ (Wilson, 2002). (Eq. 3)

Methods

Study Area

Two study areas (Figure 1) were chosen based on a Burn Probability dataset (Scott, et. al., 2020). Burn Probability was selected as a determining factor because of this study's overarching focus investigating interactions between the power grid and wildfire. The High Burn Probability Study Area (High BP) was an area with ≥ 5% annual burn probability located in western Idaho (115.886°W 43.299°N). This area is dominated by the invasive annual grass, cheatgrass (*Bromus tectorum*). The Low Burn Probability Study Area (Low BP) has an annual burn probability of ≤ 1% and was located in eastern Idaho (112.118°W 43.312°N). The vegetation in this area is considered sagebrush-steppe (*Artemisia tridentata* and grasses).

For this study, irrigated agricultural areas were identified using National Agricultural Imagery Program (NAIP) data and removed from further analysis. This was done to focus on the natural areas which are the primary site of wildfires. As a result, the High BP study area covered 1,048 km^2 while the Low BP study area covered 578 km^2.

Data Sources

Four recent fires burning in 2011, 2013, 2014, and 2016 were identified that intersected the study areas (Weber 2022). Two fires were chosen for each study area. Between four and five scenes were acquired for each year/study area based on availability of quality imagery (Table 1). Landsat 5 or Landsat 8 imagery were acquired for each area (Path 041 Row 030 (High BP) and Path 039 Row 030 (Low BP)) during the growing season of March through August in all four years (2011, 2013, 2014, and 2016). Level one (L1) imagery was downloaded from USGS Earth Explorer providing a total of 44 scenes of data. Scenes with >25% cloud cover over the wildfire areas, identified visually, were excluded. Only those scenes with a temporally corresponding imagery pair for High and Low BP were used in the analysis (*n* = 36 scenes or 18 pairs). This was done to eliminate a potential data bias where study sites had a different number of scenes analyzed.

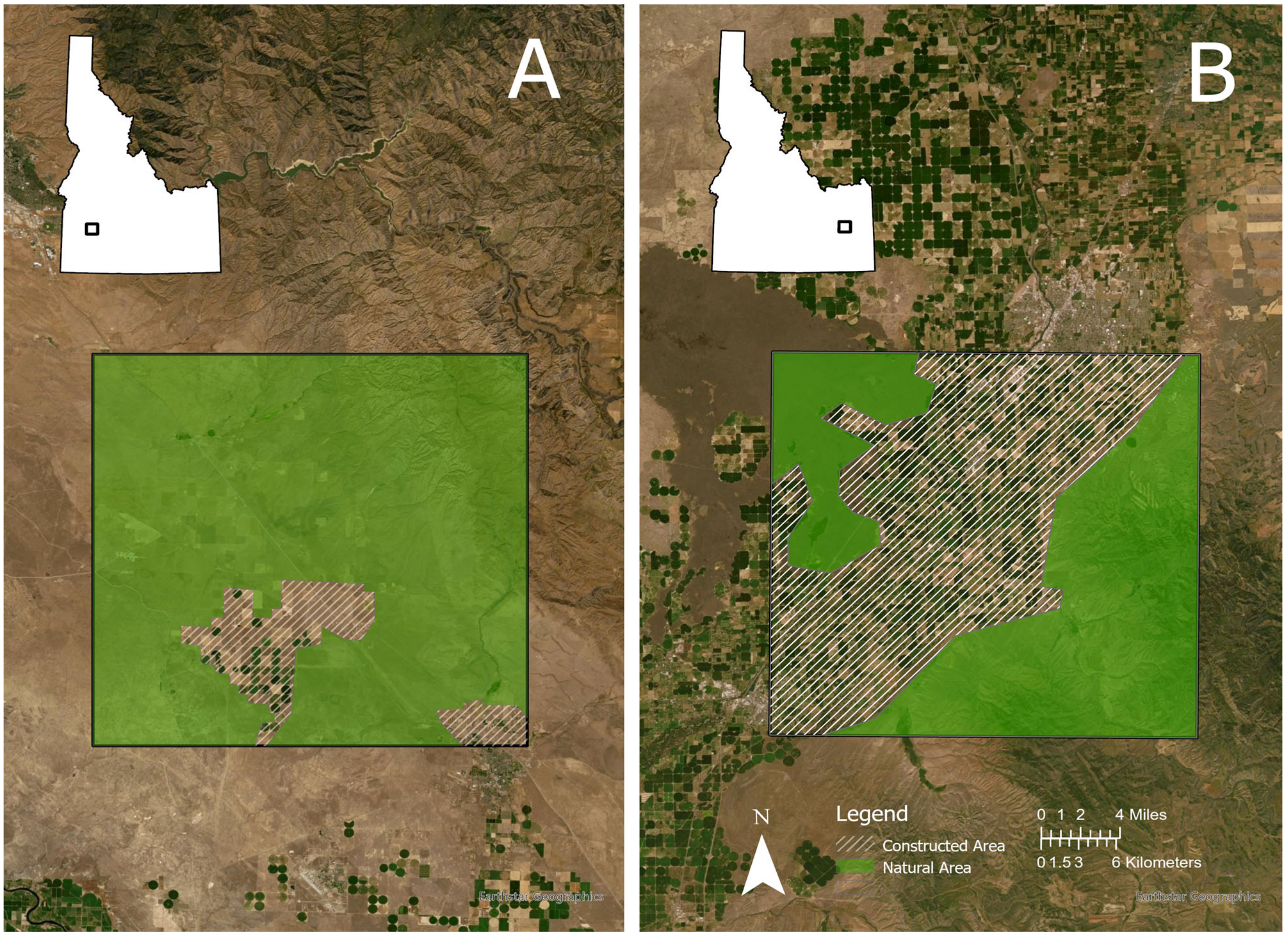
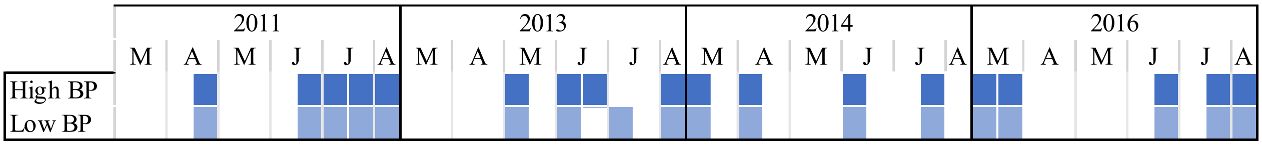


Figure 1. *Study areas within Idaho, USA. High (A) and Low (B) burn probability study areas are shown within the boxes. Green is used to designate the natural areas used for analysis in this study.*

Sampling and statistical analysis

Each Landsat scene was imported into Idrisi TerrSet using the Landsat archive import tool. Scenes were atmospherically corrected by converting each multispectral band (DN) to reflectance using the Cos(t) reflectance correction model. The VegIndex tool was used to calculate both NDVI and MSAVI2 SBRs for each scene. Both equations used their respective Red and Near-Infrared (NIR) bands: Bands 3 and 4 for Landsat 5; Bands 4 and 5 for Landsat 8. The Overlay tool used the NIR and Short-Wave Infrared (SWIR) bands to calculate NDMI. Bands 4 and 5 were used for the Landsat 5 equation; Bands 5 and 6 for the Landsat 8 equation. Each SBR was exported using the GEOTIFF/TIFF export tool for further analysis in ArcGIS Pro.

Table 1. *Imagery coverage by year and study area (high and low burn probability). Shaded boxes depict the months during the growing season (March through August, abbreviated by first letter) for which Landsat data was acquired. All 2011 imagery is from the Landsat 5 sensor; all other imagery is from Landsat 8. All imagery pairs were captured within two days of each other (note: mid-summer 2013 imagery was acquired on June 30 and July 2 for the High BP and Low BP study areas, respectively).*



In ArcGIS Pro, the Zonal Statistics as Table tool was used to generate descriptive statistics for each SBR layer (NDVI, MSAVI2, and NDMI) within each study area polygon. All resulting tables were merged by SBR type. The combined tables were then exported to Excel. Analysis of Variance (ANOVA) was calculated comparing NDVI to MSAVI2, NDVI to NDMI, and MSAVI2 to NDMI using the Excel Data Analysis tool pack to determine the significance of correlation between each SBR pair. A confidence interval where $P \leq 0.01$ was used to ensure a conservative approach to interpretation.

Results/Discussion

A pattern of high correlation between NDVI and MSAVI2 was observed (Table 2). There were no instances where either the NDVI and NDMI or the MSAVI2 and NDMI analyses resulted in a significant correlation (Table 2), suggesting these SBRs provide very different information to the user. When the mean of each SBR was averaged across all four years, the Low BP area exhibited higher mean values (NDVI = 0.37, MSAVI2 = 0.22, NDMI = 0.01) than that found in the High BP area (NDVI = 0.26, MSAVI2 = 0.19, NDMI = -0.10) (Figure 2) suggesting higher overall productivity in the Low BP area. In addition, sagebrush-steppe areas typical of the Low BP study area tended to exhibit less bare ground exposure relative to the cheatgrass dominated High BP area. This further helps explain the higher overall values found in the Low BP area.

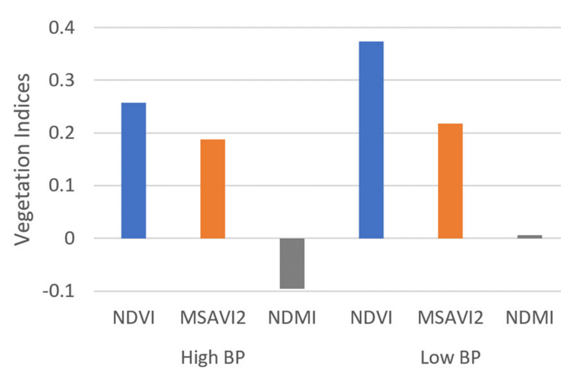


***Figure 2**. Grand mean of indices across all study years by SBR and study area (high and low burn probability (BP))*

The lack of significant correlation between NDVI (or MSAVI2) and NDMI compared to NDVI and MSAVI2 is likely due to the design or intent of each SBR. In other words, NDVI and MSAVI2 are vegetation productivity indices and should be expected to agree whereas NDMI is a moisture index.

Table 2. Coefficient of determination (R^2) from Analysis of Variance (ANOVA) tests between SBRs. Significant results are bolded and based on a P-value of ≤ 0.01.

	2011		2013		2014		2016	
	High BP	Low BP	High BP	Low BP	High BP	Low BP	High BP	Low BP
NDVI/MSAVI2	**0.976**	0.985	0.980	0.999	0.999	0.944	0.994	**0.771**
NDVI/NDMI	0.160	0.942	0.921	0.466	0.344	0.017	0.126	0.686
MSAVI2/NDMI	0.231	0.968	0.979	0.495	0.314	0.129	0.121	0.956

Another noted relationship was that NDVI typically had higher values than MSAVI2. The only time this was not true was a single instance early in the growing season for the High BP Area (Figure 3). This was likely due to actively growing cheatgrass resulting in MSAVI2 values that exceeded those of NDVI. Throughout the rest of the season, these SBRs tended to trend well with very similar residuals seen between NDVI and MSAVI2 across a given year. This is most likely due to the influence of bare ground reflectance. Large residuals between these VIs indicate, to a degree, the amount of bare ground. High residuals may also indicate a lack of fuel continuity and the development of fine fuels across the study area. The SBRs explored here focus on vegetation reflectance values and do not explicitly define different soil types; rather, this study compares the results of VIs that are unadjusted and adjusted for soil brightness respectively.

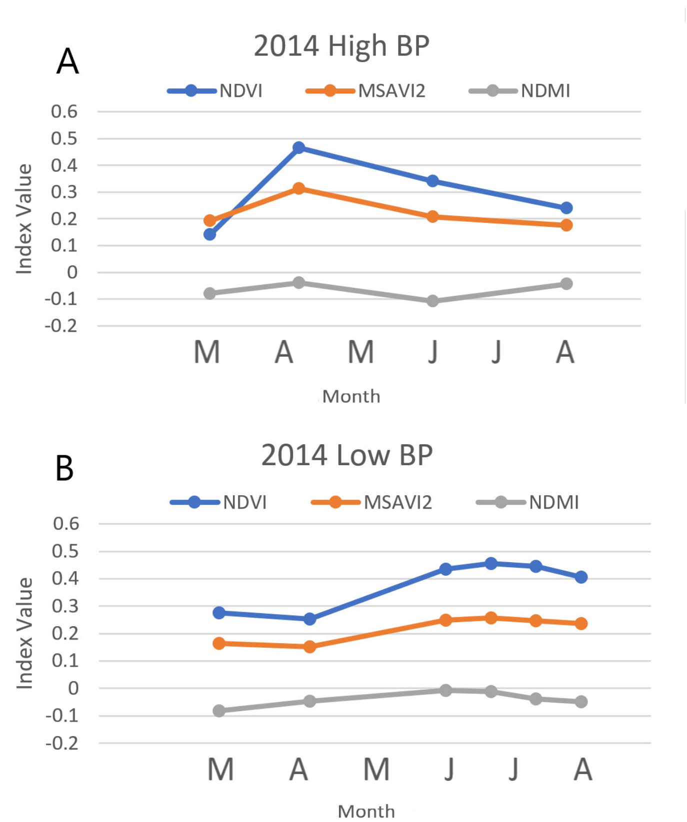


***Figure 3**. SBRs and their index values plotted across the growing season for both the High (A) and Low (B) burn probability (BP) study areas in 2014 (months of the growing season are abbreviated by first letter from March through August).*

Conclusions

The results of this study indicate the two VIs (NDVI and MSAVI2) are strongly correlated in semi-arid environments with lower bare ground exposure, but less well correlated --and with large residuals-- in regions with high bare ground exposure. It is noteworthy to consider the potential to leverage the differences between NDVI and MSAVI2 to aid in identifying areas of high bare ground and develop a model of fine fuels for use in wildfire susceptibility modeling. Further research and field validation is merited to fully explore this concept.

Because there is a notable difference observed between the various SBRs used in this study, it is evident that choosing the proper SBR to be used in a study is important. This determination should be based on the sensitivity of the SBR to a study area's land cover type and biophysical setting, the percent of bare ground exposure found in the study area, and the albedo of the soils (light colored soils have a greater tendency to saturate sensors). Studies in mesic areas with consistently high precipitation may find little difference between vegetation indices and should find NDVI to be most useful as it is available as a prepared data product from various sources such as NASA MODIS (Didan, 2015). Arid and semi-arid regions with high soil exposure may find it advantageous to use MSAVI2 as it tends to correlate better with measured biomass production (Theau et al 2010; Chen et al 2011).

Arid and semi-arid ecosystems cover 37% of land (Barakat, 2011), making the use of accurate vegetation indices important as future studies are conducted. Additional work might also include comparing additional VIs, such as the Transformed Soil Adjusted Vegetation Index (TSAVI), which also minimizes soil brightness influences by a different equation. For example, there may be correlation between NDVI, MSAVI2, and TSAVI because they are all characterized by a slope rather than a distance (Baret, 1991).

Because NDVI and NDMI did not correlate in either study area, future work could explore other variables that might influence correlation besides soil reflectance. Further analysis of this study in general should observe variables that might contribute to the differences in grand mean between each study area such as precipitation, elevation, or fire frequency. Applications of this research in wildfire management could include determining which SBR is most appropriate for monitoring fuel continuity or fuel load in landscapes of varying vegetation density and soil reflectance.

Acknowledgements

Funding for this study was provided through an Idaho State University Center for Advanced Energy Studies (ISU CAES) grant to the GIS Training and Research Center. Landsat 5 and 8 imagery were acquired through the U.S. Geological Survey online Earth Explorer.

References

Barakat, H.N. 2009. Arid lands: Challenges and hopes. In Earth System: History and Natural Variability-Vol. III; EOLSS publishers/UNESCO: Oxford, UK.

Baret, F., & Guyot, G. 1991. Potentials and limits of vegetation indices for LAI and APAR assessment. Remote Sensing of Environment, 35, 161-173. https://doi.org/10.1016/0034-4257(91)90009-U

Chen, F., K. T. Weber, and B. Gokhale. 2011. Herbaceous Biomass Estimation from SPOT 5 Imagery in Semiarid Rangelands of Idaho. GIScience & Remote Sensing 48:2, 195-209. DOI: 10.2747/1548-1603.48.2.195

Didan, K. 2015. MOD13Q1 MODIS/Terra Vegetation Indices 16-Day L3 Global 250m SIN Grid V006. NASA EOSDIS Land Processes DAAC. Accessed 2022-10-20 from https://doi.org/10.5067/MODIS/MOD13Q1.006

Rouse, J. W. Jr., Haas, R., H., Deering, D. W., Schell, J. A., and Harlan, J. C., 1974. Monitoring the Vernal Advancement and Retrogradation (Green Wave Effect) of Natural Vegetation. NASA/GSFC Type III Final Report, Greenbelt, MD., 371.

Scott, Joe H.; Gilbertson-Day, Julie W.; Moran, Christopher; Dillon, Gregory K.; Short, Karen C.; Vogler, Kevin C. 2020. Wildfire Risk to Communities: Spatial datasets of landscape-wide wildfire risk components for the United States. Fort Collins, CO: Forest Service Research Data Archive. Updated 25 November 2020. https://doi.org/10.2737/RDS-2020-0016

Theau, J., T. T. Sankey, and K. T. Weber. 2010. Multi-sensor Analyses of Vegetation Indices in a Semi-arid Environment. GIScience & Remote Sensing 47:2, 260-275. DOI: 10.2747/1548-1603.47.2.260

Imagery in Semiarid Rangelands of Idaho. GIScience & Remote Sensing 48:2, 195-209. DOI: 10.2747/1548-1603.48.2.195

Tucker, C. J. 1979. Red and Photographic Infrared Linear Combinations for Monitoring Vegetation. In Remote Sensing of Environment (Vol. 8).

Qi, J., Chehbouni, A., Huete, A. R., Kerr, Y. H., & Sorooshian, S. (1994). A modified soil adjusted vegetation index. *Remote Sensing of Environment*, *48*(2), 119–126. https://doi.org/https://doi.org/10.1016/0034-4257(94)90134-1

Weber, K. T., Idaho State University GIS Training and Research Center. 2020. Historic Fires Database (HFD) version 3.0. Downloaded from https://giscenter.isu.edu/research/Techpg/HFD/ October 20, 2022

Wilson, E.H. and Sader, S.A., 2002, "Detection of forest harvest type using multiple dates of Landsat TM imagery." Remote Sensing of Environment, 80, pp. 385-396.