Supplemental Data for Manuscript <u>AntSci-2019-SP-1331</u>, *Remote Characterization of Photosynthetic Communities in the Fryxell Basin of Taylor Valley, Antarctica*

M. R. Salvatore[1]*, S. R. Borges[1], J. E. Barrett[2], E. R. Sokol[3], L. F. Stanish[3], S. N. Power[2], and P. Morin[4]

[1]Department of Astronomy & Planetary Science, Northern Arizona University, NAU Box 6010, Flagstaff, AZ, 86011, USA, mark.salvatore@nau.edu, *Corresponding author

[2]Department of Biological Sciences, Virginia Tech, 2125 Derring Hall, Mail Code 0406, Blacksburg, VA, 24061, USA

[3]National Ecological Observatory Network, Battelle Memorial Institute, 1685 38th Street, Suite 100, Boulder, CO, 80301, USA

[4]Polar Geospatial Center, University of Minnesota-Twin Cities, 1954 Buford Avenue, St. Paul, MN, 55108, USA

Methods

The primary methods used in this manuscript are associated with (1) the calibration of orbital WorldView-2 and WorldView-3 (WV2 and WV3, respectively) data to surface reflectance and the derivation of spectral parameters, and (2) the use of linear spectral unmixing techniques to estimate the contribution of different spectral endmembers to the observed spectral signature. Both of these techniques are described in more detail below. In addition, **Figure S1** shows a flowchart of the methods used in this work to help orient the reader to the steps that are necessary to derive our results.

Dark Object Subtraction and Regression (DOS-R) Atmospheric Correction
The Dark Object Subtraction and Regression (DOS-R) atmospheric correction technique is, at the moment, the most robust atmospheric correction technique that can be performed on images in the absence of ground validation or synoptic atmospheric measurements. In the absence of regular atmospheric observations, scene-derived atmospheric correction is most feasible in the McMurdo Dry Valleys (MDV), where atmospheric conditions are known to change on short timescales and weather forecasting is near impossible (Doran *et al.* 2002). In addition, even if atmospheric characterization in the MDV were possible, the long solar path lengths through the atmosphere due to the high latitudes in the MDV create additional problems in trying to remove atmospheric effects. As a result, scene-derived atmospheric correction is most appropriate for the MDV.

Dark Object Subtraction (DOS) (and minor improvements to this technique) is one of the most common methods of atmospheric correction used globally with multispectral remote sensing data (e.g., Chavez 1988). The traditional DOS technique uses the darkest observed pixels (typically shadowed water) and removes that spectrum directly from all other pixels in the scene through simple subtraction. The spectral signature from shadowed deep water consists almost entirely of contributions from atmospheric scattering, which is why this technique is an effective means of atmospheric removal. In the MDV, however, where deep water is oftentimes not a part of the scene, the darkest pixels are typically shadowed areas of interesting geologic units. These geologic units exhibit their own unique spectral signatures, and so removing this spectrum from every other pixel in the scene can result in overprinting each pixel with spectral contributions from geologic surfaces. For this reason, traditional DOS techniques are not appropriate for regions like the MDV where ideal pixels for atmospheric removal do not exist.

Instead, the DOS-R technique uses shadowed regions to predict the atmospheric spectrum regardless of the nature of the shadowed surface. This technique uses surface radiance spectra from a landscape under differing illumination conditions (e.g., a hillslope). By varying the intensity of solar illumination across a geologically homogeneous surface, it is possible to observe how each spectral band varies as a function of solar illumination. Assuming that the longest wavelength spectral band is not significantly influenced by Rayleigh scattering within the atmosphere, it is then possible to determine the relationship between the last spectral band and every other spectral band and to derive a mathematical regression to predict values for each spectral band where solar illumination is equal to zero. **Figure S2** demonstrates these regression lines from Salvatore *et al.* (2014). Where the regression predicts that the longest wavelength band has a radiance value equal to zero, it is assumed that solar illumination is also equal to zero and, in the absence of atmospheric scattering at this longest wavelength, any predicted value in other bands is the result of atmospheric scattering. These predicted values are then subtracted

from all pixels within the image. In this study, we used at least four locations within each image to derive an average atmospheric scattering spectrum that was then removed from each pixel.

Three primary assumptions are made when using the DOS-R atmospheric correction method. First, top-of-atmosphere radiance values must be collected from a region that exhibits significant differences in solar illumination (e.g., topographic slopes) but one that is assumed to be compositionally homogeneous. Compositional homogeneity is required to ensure that any observed spectral variation is due to variations in solar illumination and not the result of differences in surface composition. Second, it is assumed that atmospheric scattering is negligible in the longest wavelength band of the orbital data (~0.91 µm). Through this assumption, it is possible to fit regressions between the longest wavelength band and all other spectral band to derive the synthetic radiance values where the radiance in the last band is equal to zero. These synthetic radiance values then serve as a prediction of the atmospheric radiance at each wavelength. Rayleigh theory predicts that atmospheric scattering at 0.91 µm is approximately 1.1% of downwelling radiance, which suggests that our assumption is not perfectly satisfied. Nonetheless, a 1.1% error in atmospheric correction is well within the limits of our remote sensing and spectral characterization capabilities. The final assumption is that contributions from the atmosphere are uniform across the entire scene. This assumption allows us to remove a single synthetic atmospheric spectrum from the entire scene. Once subtracted from each band, the resultant image then represents derived surface radiance, which can be further calibrated to surface reflectance.

Linear Spectral Unmixing

We utilize the linear least-squares spectral unmixing algorithm developed by Ramsey & Christensen (1998). Similar linear unmixing algorithms have been previously used for identifying the fraction of surface endmembers using visible/near-infrared (VNIR) datasets (e.g., Adams *et al.* 1995, and references therein). These techniques are rooted in the same fundamental principle that observed spectra can be modeled using a linear combination of provided library endmembers. The model used here linearly combines the provided library endmembers to minimize the total spectral offset between the measured and modeled spectra. Our selection of library endmembers is explained in detail in the main manuscript text.

The linear unmixing model used in our analyses utilizes two important constraints (Ramsey & Christensen 1998). First, while spectra can be modeled at negative abundances in each pixel, negative values are interpreted to indicate their absence from the spectrum. Second, we utilize a unity constraint in our unmixing model, which requires the model to sum to 100% abundances. The result of the unmixing model is an image cube containing 6 bands, five bands corresponding to the modeled percentages of each spectral endmember and one corresponding to the calculated root mean square (RMS) error indicating the "goodness of fit" of the model to the input spectrum. A good model fit does not necessarily indicate an accurate modeled surface composition, but instead that the spectral shape can be accurately determine using the provided endmembers.

The primary reason why good model fits cannot be directly related to areal surface abundance is because of the non-linear properties of VNIR reflectance data. In the thermal infrared (TIR), where linear spectral unmixing models are commonly used, the absorption coefficients of most geologically relevant materials are substantially higher than in the VNIR spectral region. This dictates that observed radiation is sourced from interactions with only the uppermost surface and not from underlying materials. As a result, the observed spectrum is

representative of the areal mixture of endmembers in the scene and can, therefore, be linearly unmixed. At VNIR wavelengths where absorption coefficients are low relative to in the TIR, radiation is able to interact with materials below the uppermost surface as a result of increased refraction and scattering. Take a mixture of cinnamon and refined sugar as an example. Mixtures of 50% cinnamon and 50% sugar appear to be dominated by the cinnamon component, as cinnamon has a much higher absorption coefficient relative to refined sugar. The result is that radiation is able to non-linearly pass through the sugar to observe the underlying cinnamon, resulting in a darker appearance than the even 50%-50% mixture would suggest.

Fortunately, where care is taken to select optically opaque endmembers, it can generally be assumed that linear unmixing provides a first-order approximation of the relative areal abundance of the different endmembers. For example, our endmember library consists of a black microbial mat spectrum and two soil spectra. These spectral components are relatively opaque and, where mats are widespread across the surface, are typically optically thick. These assumptions allow us to use our unmixing model results to make first-order estimates of surface abundances. While certainly not a perfect assumption in all locations (e.g., where thin photosynthetic films overly geologic materials, or potentially where photosynthetic microbial materials are intimately mixed with soils outside of stream channels), we are applying this assumption to the majority of surfaces observed in the Fryxell basin for the purposes of this investigation.

The requirement for optically opaque endmembers is why water was not included in our endmember library. Unless the water is optically thick (e.g., greater than several meters), shallow water has nonlinear influences on the observed spectral signature. For example, when viewed from above, clear water that is only a few centimeters deep more closely resembles the substrate than it does optically thick water. Fortunately, shallow water is not completely transparent and it has significant wavelength-dependent influences on the observed spectral signatures. As water becomes progressively deeper, it preferentially absorbs radiation starting with the longest wavelengths and encroaching into the shorter visible wavelengths. Therefore, the resultant spectral signature varies when water is present at variable depths. The influences of steadily increasing water depth also result in progressively poorer fits between the observed spectra and the resultant unmixing models, which can be observed as increases in RMS error. It is possible, therefore, to identify where water of spectrally significant depth is present in an unmixed scene by identifying the areas within stream channels and lakes with the greatest RMS error.

RMS error is not only a valuable indicator of where the best-fit model does not appropriately match the measured spectrum. It also provides useful information regarding which spectral endmembers have been erroneously (or purposefully) omitted from the unmixing model. In the preceding paragraph, we described how the influence of liquid water can increase the RMS error in the model results. This is because liquid water at all depths and with all relevant substrates cannot be reasonably included in our unmixing model. Similarly high RMS errors can be observed in areas with snow or ice, which were not modeled in our unmixing efforts because they were not of interest for the purposes of our investigation. As a result, the high RMS errors associated with these areas indicate that an important spectral endmember is missing, resulting in poor model fits.

Selection of Spectral Endmembers

Our five spectral endmembers were carefully selected for their utility in this investigation. These five endmembers were (1) healthy black microbial mat, (2) representative soil from Taylor Valley measured using a field spectrometer, (3) a representative scene-derived soil spectrum from within Taylor Valley, (4) a "dark" endmember that exhibits 0% reflectance at all wavelengths, and (5) a "bright" endmember that exhibits 100% reflectance at all wavelengths.

Black mats were the only microbial communities included in our endmember library for several reasons. Black mats are, by far, the most spatially and areally extensive microbial community in and near the ephemeral stream channels in the Fryxell basin (Alger *et al.* 1997; Kohler *et al.* 2015). For this reason, black mat spectra were the only data collected in previous efforts to spectrally characterize microbial communities from orbit and, therefore, were the only microbial mat spectral data available to the authors at the time of writing this manuscript. Subsequent field work and spectral characterization is ongoing to identify and characterize the full range of microbial and moss communities throughout the Fryxell basin for more comprehensive spectral investigations (e.g., Borges *et al.* 2019). Nonetheless, the predictable spectral nature of photosynthetic materials and the broad utility of the NDVI parameter allow for us to identify the presence of photosynthetic communities to a first order.

This black mat spectrum was acquired by collaborator Dr. Joseph Levy (Colgate University) using a contact probe attachment for an Analytical Spectral Devices (ASD) FieldSpec4 field spectrometer, which uses artificial illumination for better signal in the region of water and atmospheric absorption bands. The spectrometer was calibrated using a Spectralon white reference and standard techniques to derive surface reflectance. The mat was collected from Taylor Valley and shipped frozen back to the United States, where it was thawed and rehydrated for analyses including spectral data acquisition. The spectrum consists of a stack of ten spectral scans and was then downsampled to WV2 and WV3 spectral resolutions using instrument bandpass filters. The representative Taylor Valley soil was measured using the same FieldSpec4 field spectrometer, although it was measured under direct solar illumination using its bare fiber optic cable and not using the contact probe attachment. Similar white reference calibration, spectral stacking, and mathematical downsampling was employed as for the black mat spectrum. This soil spectrum is described in further detail in Levy *et al.* (2014).

References

ADAMS, J. B., SABOL, D. E., KAPOS, V., FILHO, R. A., ROBERTS, D. A., SMITH, M. O., & GILLESPIE, A. R. 1995. Classification of multispectral images based on fractions of end-members: Application to land-cover change in the Brazilian Amazon. *Remote Sensing of Environment* **52**, 137-154.

ALGER, A. S., McKNIGHT, D. M., SPAULDING, S. A., TATE, C. M., SHUPE, G. H., WELCH, K. A., EDWARDS, R., & HOUSE, H. R. 1997. Ecological processes in a cold desert ecosystem: the abundance and species distribution of algal mats in glacial meltwater streams in Taylor Valley, Antarctica. *Institute of Arctic and Alpine Research Occasional Paper* **51**, 108 pp.

BORGES, S., SALVATORE, M., POWER, S., SOKOL, E., STANISH, L., & BARRETT, J. 2019. Spectral archive and remote characterization of terrestrial microbial mats in the Fryxell basin of Taylor Valley, Antarctica. *Geological Society of America Annual Meeting* **51**, abstract T41-8-2, doi:10.1130/abs/2019AM-340576.

CHANDER, G., MARKHAM, B. L., & HELDER, D. L. 2009. Summary of current radiometric calibration coefficients for Landsat MSS, TM, ETM+, and EO-1 ALI sensors. *Remote Sensing of Environment* **113**, 893-903, doi:10.1016/j.rse.2009.01.007.

CHAVEZ, P. S. 1988. An improved dark-object subtraction technique for atmospheric scattering correction of multispectral data. *Remote Sensing of Environment* **24**, 459-479.

DORAN, P. T., McKAY, C. P., CLOW, G. D., DANA, G. L., FOUNTAIN, A. G., NYLEN, T., & LYONS, W. B. 2002. Valley floor climate observations from the McMurdo dry valleys, Antarctica, 1986-2000. *Journal of Geophysical Research* **107**, doi:10.1029/2001JD002045.

KOHLER, T. J., STANISH, L. F., CRISP, S. W., KOCH, J. C., LIPTZIN, D., BAESEMAN, J. L., & McKNIGHT, D. M. 2015. Life in the main channel: long-term hydrologic control of microbial mat abundance in McMurdo Dry Valley streams, Antarctica. *Ecosystems* **18**, 310-327, doi:10.1007/s10021-014-9829-6.

LEVY, J., NOLIN, A., FOUNTAIN, A., & HEAD, J. 2014. Hyperspectral measurements of wet, dry and saline soils from the McMurdo Dry Valleys: Soil moisture properties from remote sensing. *Antarctic Science* **26**, 565-572, doi:10.1017/S0954102013000977.

RAMSEY, M. S., & CHRISTENSEN, P. R. 1998. Mineral abundance determination: Quantitative deconvolution of thermal emission spectra. *Journal of Geophysical Research* **103**, 577-596.

SALVATORE, M. R. 2015. High-resolution compositional remote sensing of the Transantarctic Mountains: application to the WorldView-2 dataset. *Antarctic Science* **27**, 473-491, doi:10.1017/S095410201500019X.

SALVATORE, M. R., MUSTARD, J. F., HEAD, J. W., MARCHANT, D. R., & WYATT, M. B. 2014. Characterization of spectral and geochemical variability within the Ferrar Dolerite of the McMurdo Dry Valleys, Antarctica: weathering, alteration, and magmatic processes. *Antarctic Science* **26**, 49-68, doi:10.1017/S0954102013000254.

UPDIKE, T., & COMP, C. 2010. *Radiometric use of WorldView-2 imagery: technical note*. Longmont, CO: DigitalGlobe, 16 pp.

Table and Figure Captions

Figure S1: A flow chart detailing the methods used in this manuscript. Black rounded boxes indicate generated products, while gray rectangles indicate procedures, processes, or additional input data derived from other sources. References are superscripted and correspond to the following:

[A]Image metadata provided by DigitalGlobe, Inc., and the Polar Geospatial Center.
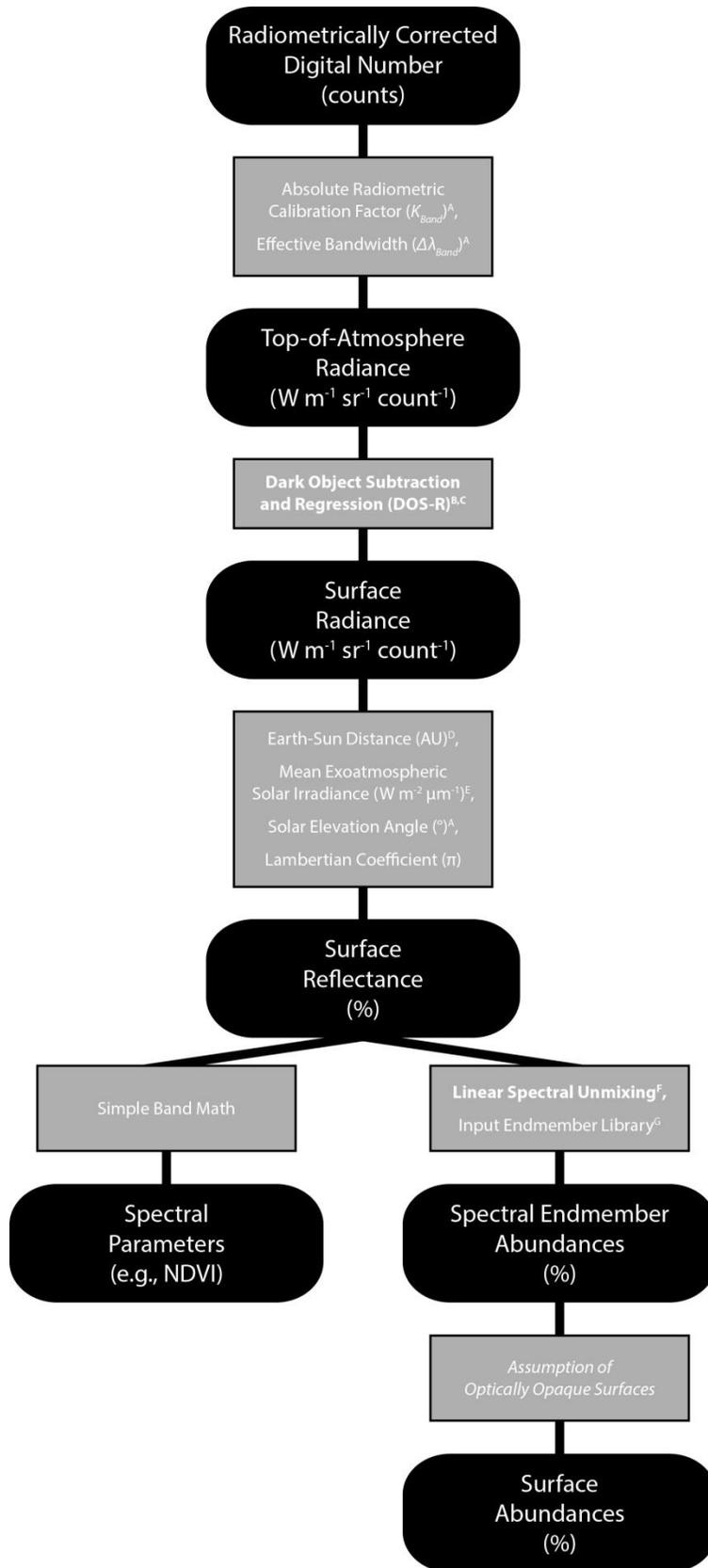[B]Salvatore et al. (2014)
[C]Salvatore (2015)
[D]Chander et al. (2009)
[E]Updike & Comp (2010)
[F]Ramsey & Christensen (1998)
[G]Input endmembers are derived either from scene-derived spectra or laboratory spectra that are downsampled to WV2 or WV3 resolutions.

Figure S2: Dark object subtraction and regression (DOS-R) scatterplots derived for Advanced Land Imager (ALI) data over the McMurdo Dry Valleys of Antarctica. From Salvatore *et al.* (2014), Figure 4c. Fitting linear regressions to these spectral data relative to the reflectance observed in the last band can be used to derive radiance associated with atmospheric scattering within the image and can then be removed from each pixel through simple subtraction.
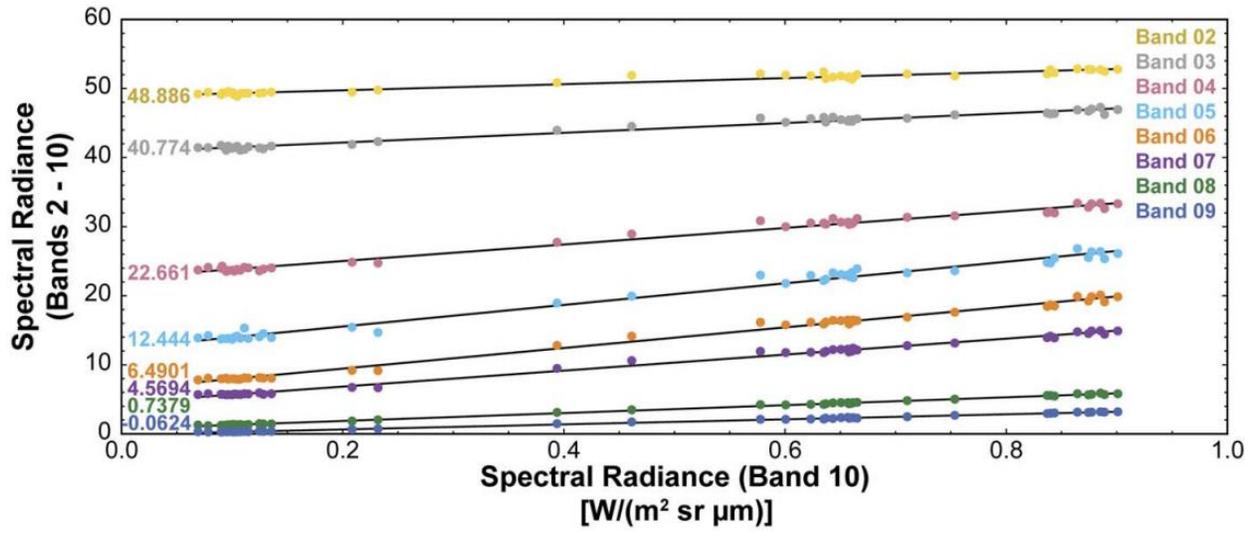
Figure S1.

Figure S2.