RICHNESS AND EQUITABILITY MEASURES APPLIED TO A K-MEANS CLASSIFICATION OF ALL FOUR MARS ROVERS' APXS OXIDES AND ELEMENTS DATA. C.M. Rodrigue[1] , [1]Geography, California State University, Long Beach, CA 90840-1101, rodrigue@csulb.edu.

Introduction: [1] describes the combination of APXS data for all 4 Mars rovers into a common target classification system using K-means clustering of standardized weight abundance data for 16 oxides and elements. The goal here is to compare and contrast the 4 rover work areas in terms of their geochemical "personalities" using established measures of "biodiversity" and 1 new measure.

Biogeography and ecology have long attempted to quantify "biodiversity" to analyze ecosystems, identify biodiversity hotspots, and create metrics for conservation and restoration projects. Biodiversity, while intuitively straightforward, is challenging to operationalize. It can refer to simple counts of species observed in an area, that is, species richness. Species richness can mislead by granting equal weight to a large resident population and to the occurrence of a single vagrant individual outside its normal range. Despite this, species richness is easy to count and does register fluctuations in rare species, so its use persists [2].

Species richness has been expanded to represent scale-dependent patterns in a nested spatial hierarchy. Species counts can be done for local study areas or small patches (*alpha* diversity) or for coarser-scale landscapes (*gamma* diversity) [3] though the exact scales for each level vary among workers [4]. This scale-dependent biodiversity system is coupled with measures of species contrast among pairs of local patches (*beta* diversity, counting only species in the pair but not shared between them) [5].

Nested measures of species counts and contrasts still embody the same problem of false equalization of species weights, only at multiple scales. Other metrics have been devised to penalize inequity and reward evenness of species representation [2]. Two of the most common are designed to form indices ranging from 0 for complete dominance by a single species and 1 for complete evenness among several species:

Simpson's Inverse D (D_{INV}):

$$D_{INV} = 1 - \Sigma p_i^2 \qquad \text{(Eq. 1)}$$

where: i = a particular species

p = proportion of each species

Shannon's Equitability measure (E_H):

$$E_H = [-1(\Sigma p_i(\ln(p_i)))]/\ln(\alpha) \qquad \text{(Eq. 2)}$$

where: pi = proportion of each species

α = target diversity at a given study area

As reported in [1], APXS oxide and element standardized relative abundances for MPF, MER, and MSL were classified into 15 clusters using K-means clustering, and the distribution of these clusters among the 4 rover study areas was compared. The present paper applies richness and equitability metrics to these data.

Data and Methods: Counts of APXS targets by cluster are here given in *alpha, beta,* and *gamma* diversity formats. Proportions of each cluster by rover site were then calculated from these to produce the evenness measures.

Results:

Cluster richness and contrast. Gamma diversity is here used as the global diversity of clusters, given by the K=15 request. In terms of local or *alpha* diversities, Chryse Planitia sampled by MPF unsurprisingly has the lowest *alpha* diversity at 2. limited by the 11 APXS targets the Sojourner rover was able to take. The other 3 sites are similar to one another, Meridiani Planum having the highest *alpha* diversity at 13.

In terms of *beta* diversity contrasts, MPF again unsurprisingly shares the fewest clusters with the other sites, since it only had 2 to share. With regard to the more productive rovers, Gusev Crater most resembles Meridiani Planum, given that MER-A and MER-B failed to share only 3 target clusters. Gale Crater and Gusev Crater are the most divergent, MER-A and MSL failing to share more than half their clusters, fully 7 out of MSL's 11 and MER-A's 12 cluster counts. Meridiani serves as a bridge overlapping Gusev and Gale, MER-B failing to share only 3 cluster types with MER-A and only 4 with MSL.

Diversity	α	β MPF	β MER-A	β MER-B	β MSL
MPF	2	---			
MER-A	12	10	---		
MER-B	13	11	3	---	
MSL	11	9	7	4	---
γ	15				

APXS target richness or target contrast at the 4 Mars rover sites	
α alpha	# clusters found at a given site: richness
β beta	# clusters not shared between two sites: contrast
γ gamma	# clusters at all four rover sites: richness

Table 1: Cluster richness and contrast by rover site

Cluster equitability. The Simpson Inverse Diversity measures show similar high levels of cluster equitability for the MSL, MER-A, and MER-B sites (D_{INV}= 0.83 for the first 2 and 0.77 for MER-B), while MPF shows a lower D_{INV} at 0.46, which, even so, is pretty high for a site with 11 readings and 2 clusters.

The Shannon Equitability Index shows similarly high equitability figures for MSL, MER-A, and MER-B (E_H = 0.84, 0.81, and 0.71, respectively). For MPF, however, it provides the absurdly high figure of 0.95: A site with 2 clusters among 11 readings becomes the APXS diversity rock star!

Equitability measures are easily skewed by low numbers of species [6]. In this case, the bloated figure for Chryse Planitia is an artifact of the conversion of Shannon Diversity (H', the numerator in Eq. 2) into the equitability measure, E_H, by dividing H' by the natural logarithm of *alpha* diversity. This denominator excessively prioritizes evenness over richness in situations with low richness.

I experimented with using ln *gamma* diversity as an alterative denominator in Eq. 2 for a variety of real and artificial biogeographical datasets, and the new measure, E_R, asymptotes at 1.0, like E_H, as desired. It also penalizes equitability in situations of low richness.

Applied to the APXS data, the results are far more reasonable: While not as high as D_{INV} or E_H, E_R shows high values for MLS, MER-A, and MER-B (0.74 for the first two and 0.67 for MER-B). The MPF site's relative evenness is appropriately offset by its small *alpha* diversity.

	MPF	MER-A	MER-B	MSL
# of targets	11	220	370	300
cluster richness (α div.)	2	12	13	11
total richness (γ div.)	15			
Simpson's D_{INV}	0.46	0.83	0.77	0.83
Shannon's E_H	0.95 ?	0.81	0.71	0.84
Rodrigue's E_R	0.24	0.74	0.67	0.74

Table 2: Cluster diversity and equitability measures by rover site

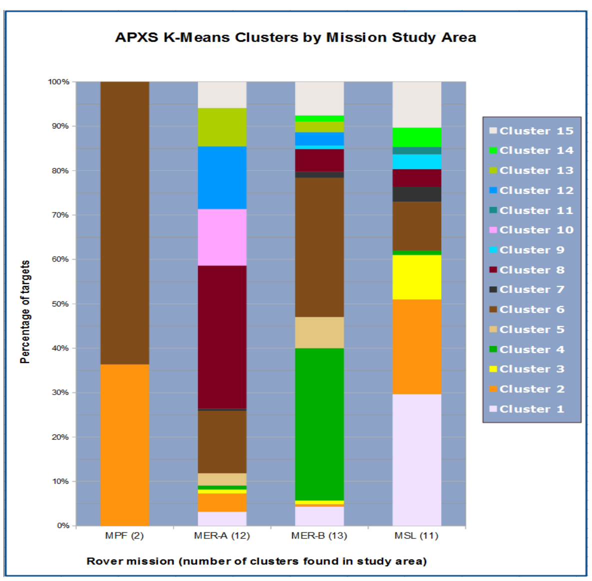

Figure 1: Clusters by rover site (from [1])

Discussion and Conclusions. All of the equitability measures slightly deprecate the Meridiani cluster mix in comparison with Gusev and Gale, even though it had a slightly higher *alpha* diversity than Gusev and a noticeably higher *alpha* diversity than Gale. Even though MER-B encountered more of the different APXS clusters than any of the other rovers, it encountered 7 of them quite rarely, while 2 were over-represented: basaltic soils (Cl. 6) and lightly acid aqueous altered basaltic soils and rocks (Cl. 4).

Similarly, all 3 of the equitability measures elevate the diversities of Gusev and Gale. About a third of MER-A's targets are picritic basalt (Cl. 8) covering the floor of Gusev. Even with that level of abundance in 1 cluster, 5 other clusters are quite abundant, so Gusev not only features high *alpha* diversity but a relatively balanced array of clusters, including neutral and acid aqueous altered materials (Cl. 12 & 13) and evolved tephrites (Cl. 10, unique to Gusev, on Husband Hill).

Gale, too, receives high scores on all 3 of the equitability indices, virtually matching the Gusev scores, this despite having a noticeably lower *alpha* diversity. A lot of the balanced profile of Gale Crater has to do with the room afforded by the relative under-representation of basalts (esp. Cl. 6, 8, & 5) in favor of evolved magma materials (Cl. 3 & 2) and materials altered by neutral-alkaline aqueous processes (Cl. 1, 9, & 12).

Chryse Planitia, as sampled by MPF, had only 2 clusters (the andesitic Cl. 2 and the Mars-typical soils in Cl. 6). With this tiny *alpha* diversity and relative balance between its 2 clusters' relative abundance, this site had the greatest disagreement in equitability-based measures of diversity. Simpson's D_{INV} did mark it down but not proportionately, and Shannon E_H can be blinded by relative evenness as it was, to an absurd degree, here. Only the new adjustment to Shannon's H' produced an appropriate mark-down for MPF. This measure may be of utility in continuing analyses of spectrometry on Mars or in other geoscience applications, such as paleontology and palynology.

Acknowledgments: Mars Exploration Rover Oxide Abundance Data and MSL APXS data were obtained from the PDS Geosciences Node: https://pds-geosciences.wustl.edu/. Mars Pathfinder APXS Oxides Data came from USGS PDS Imaging Node Server: https://pdsimage2.wr.usgs.gov/. I wish to thank Fred Calef (JPL) for MER-A UTM data.

References: [1] Rodrigue, C.M. (2020) *LPS* LI, 1262. [2] Morris, E.K. et al. (2014) *Ecol. & Evol.*, 4, 3514-3524. [3] Whittaker, R.H. (1960) *Ecol. Monogr.*, 30, 279-338. [4] Tuomisto, H. (2010) *Ecography,* 33, 2-22. [5] Bynum, N. (2019) *Biodiversity*, Biol. LibreTexts, 7, 17392. [6] Sheldon, A.L. (1969) *Ecol.*, 50, 466-467.