Change Detection in Repeat Imagery Using Principle Component Analysis

Reid A. Parsons[1,2]
(rparson4@fitchburgstate.edu)
Hirdy Miyamoto[3]

[1]University Museum, University of Tokyo, Japan
[2]Earth and Geographic Sciences, Fitchburg State University, U.S.A.
[3]Department of Systems Innovation, University of Tokyo, Japan

[Links to Abst. 2381]

Takeaway:

PCA, traditionally used in terrestrial remote sensing to distinguish between different landcovers in a multi-/hyper-spectral image, can also be used to find interesting changes in a multi-temporal image. PCA provides the advantage of separating changes into a hierarchical set based on statistical significance which helps isolate variations in lighting from more subtle, geologically interesting changes.

For More: **Parsons, R.A. & Miyamoto, H. (2018) J. of Physics: Conf. Series 1036 (012004)**

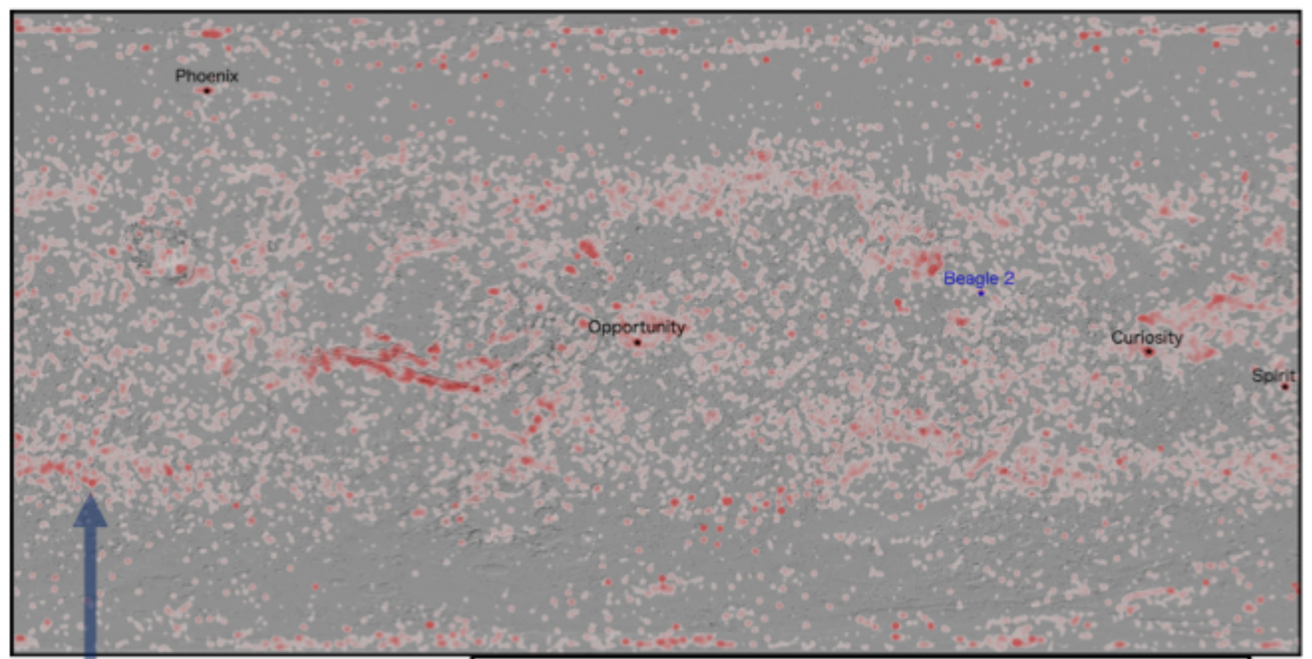


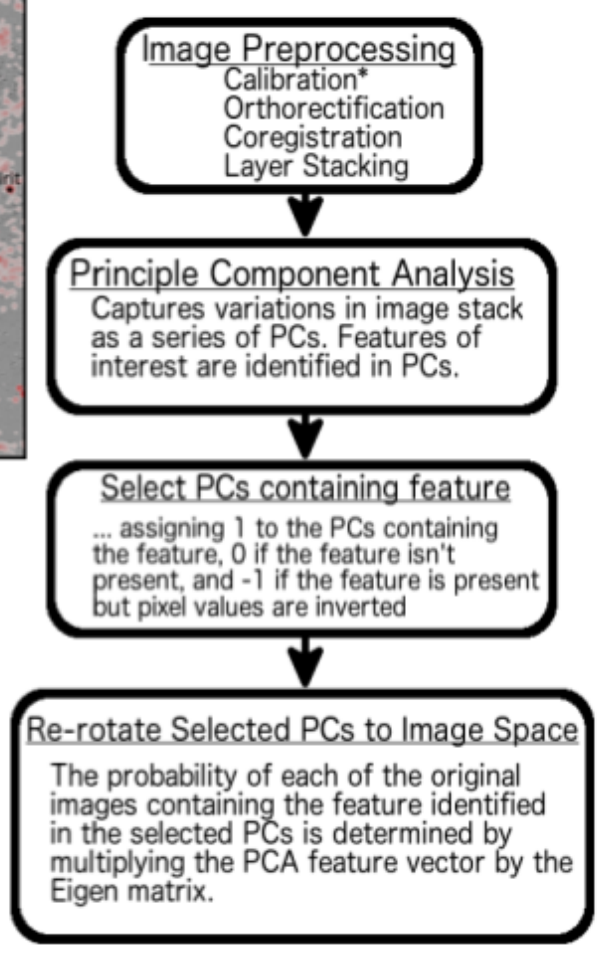
Figure 1: Heat map of 117,408 RDR HiRISE images with deeper reds indicating more repeat images.

Problem:

The large and fast-growing satellite image database for Mars [adjacent poster] requires a more efficient means for change detection than manual image "flipping." The current reduced data record (RDR) of HiRISE images with resolution better than 1.1 m/pixel is 50 TB with over 110,000 images (Fig.1). Even once images from the same location are collected, the large image size (Fig.2) and subtlety of change (e.g. dune migration) may mean that ongoing geologic processes are not detected.

The preferred, Earth-based method of using multi-spectral image classification does not apply to the single-band, hi-res data covering the largest area of Mars.

Method:



Application of the Principle Component Analysis (PCA) data reduction technique to repeat imagery provides a means of highlighting changes. PCA transforms a multi-dimensional dataset to a new coordinate system in order to better depict variations between data points. The orientation of the new coordinate system is defined by the longest axes (maximum variance) of the data point cloud in order to capture the dataset's "principle components." Once repeat images have been co-registered and stacked, each pixel has a value associated with each image in the time-series (analogous to spectral bands). The value of that pixel in each image give its coordinates in a multi-dimensional image space. All pixels in the time series produce a cloud of points (Fig.3) which can be analyzed with PCA.

An example of a PCA transformation using 2 of the 9 HiRISE images from Palikir Crater (Fig.4) is shown in Fig.3. Image (g) has RSL (highlighted in red in Fig.3a). These RSL pixels were highlighted by selecting pixels associated with the bulge in the scatter plot (Fig.3b) of (f) vs. (g) pixel values. These pixels were brighter in (f) than (g) due to RSL activity. Rotation of the pixel cloud via PCA plots the data against axes which run through the pixel point cloud and capture the maximum variance in the data (Fig.3c).

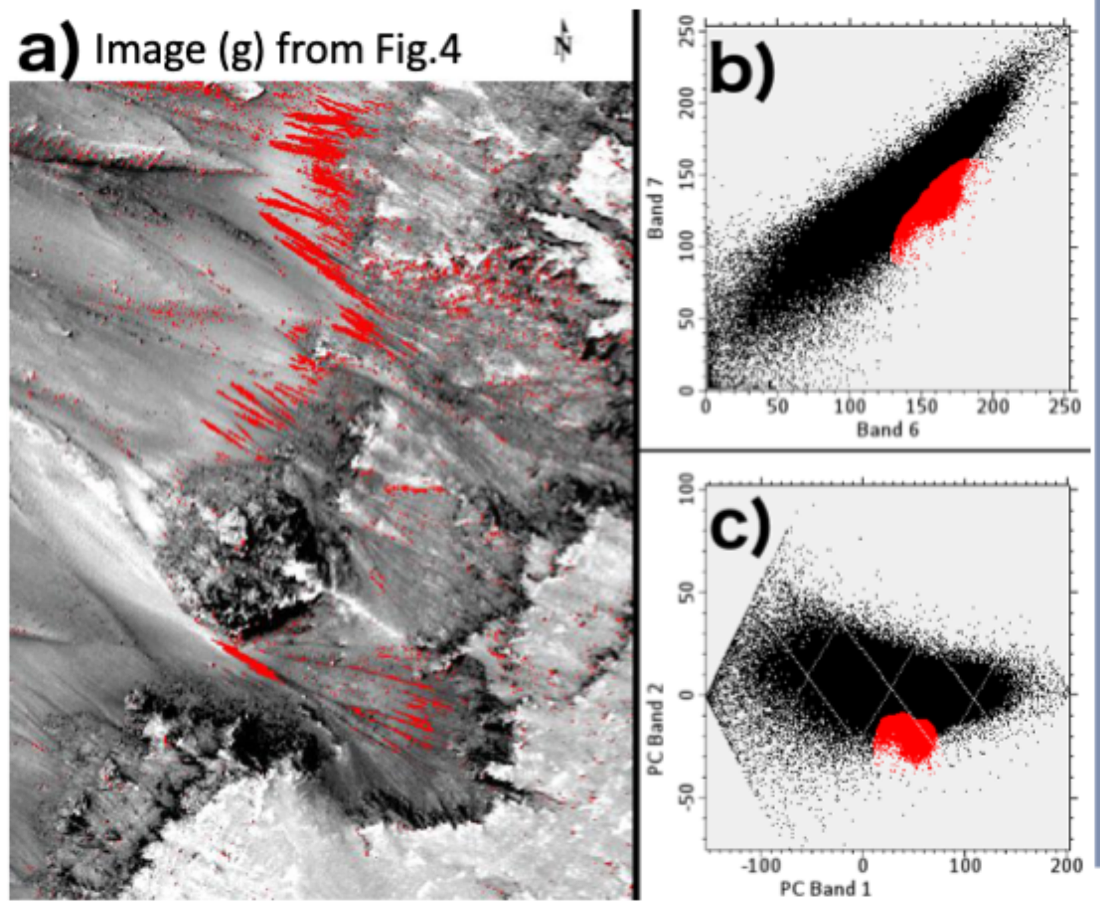


Figure 3: Illustration of PCA using 2 of the 9 images in the Palikir HiRISE time-series.

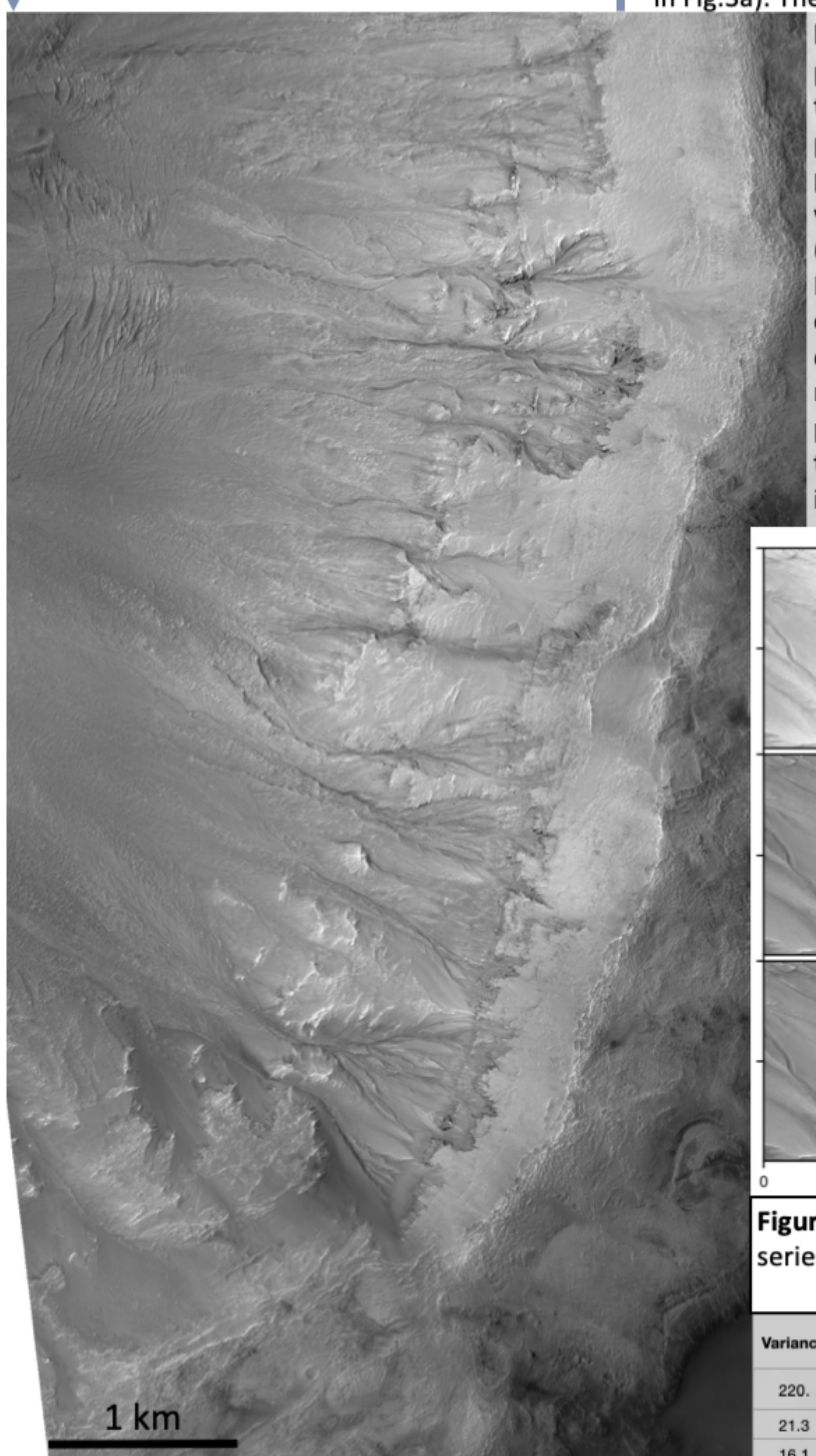


Figure 2: Example HiRISE image from Palikir Crater. At full resolution, this projected, RDR image contains 25,647 x 42,912 = 1.1 billion pixels – complicating efforts to locate changes within repeat images. Can you find the location of the Fig.4 time-series?

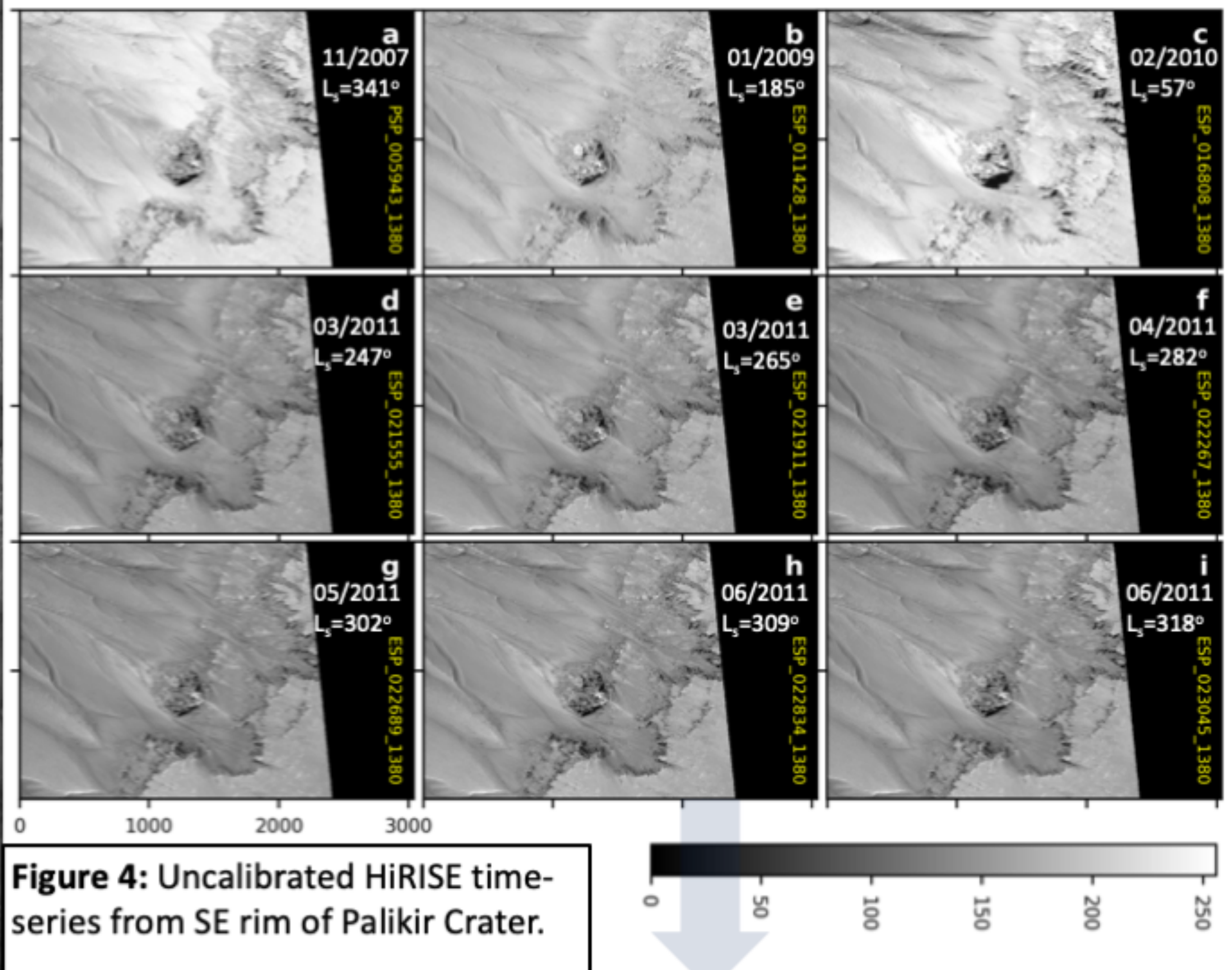


Figure 4: Uncalibrated HiRISE time-series from SE rim of Palikir Crater.

Variance	Eigen Vectors	(a)	(b)	(c)	(d)	(e)	(f)	(g)	(h)	(i)	
220.	PC1	0.41	0.37	0.41	0.29	0.31	0.30	0.29	0.30	0.30	Weights imgs nearly evenly (average)
21.3	PC2	-0.07	-0.05	-0.83	0.25	0.29	0.30	0.18	0.17	0.09	- Favors img (c) (bright) ->lighting/frost(?) differences
16.1	PC3	0.89	0.01	-0.27	-0.09	-0.06	-0.14	-0.19	-0.18	-0.19	- Favors img (a) (bright) capturing more lighting
9.76	PC4	0.20	-0.79	0.09	-0.26	-0.13	0.14	0.18	0.32	0.32	- Img combination capturing RSL (in blue)
6.33	PC5	-0.01	-0.34	0.21	0.34	0.33	0.36	-0.59	0.07	-0.37	- More RSL (in red)
4.23	PC6	0.02	-0.33	0.13	0.20	0.49	-0.30	0.48	-0.50	-0.15	- Incoherent variations
2.95	PC7	-0.03	0.14	0.01	-0.73	0.52	0.05	0.06	0.26	-0.33	- More incoherent variations
2.74	PC8	-0.02	-0.03	-0.06	0.17	0.28	-0.74	-0.29	0.45	0.24	- RSL both blue and red ?
2.51	PC9	-0.03	0.03	-0.02	-0.26	0.32	0.14	-0.38	-0.47	0.66	- RSL (red)

Why not just subtract images?:

Although useful, image differencing (Fig.6) or ratioing cannot isolate the lighting differences between images as effectively as using PCA (Fig.5)

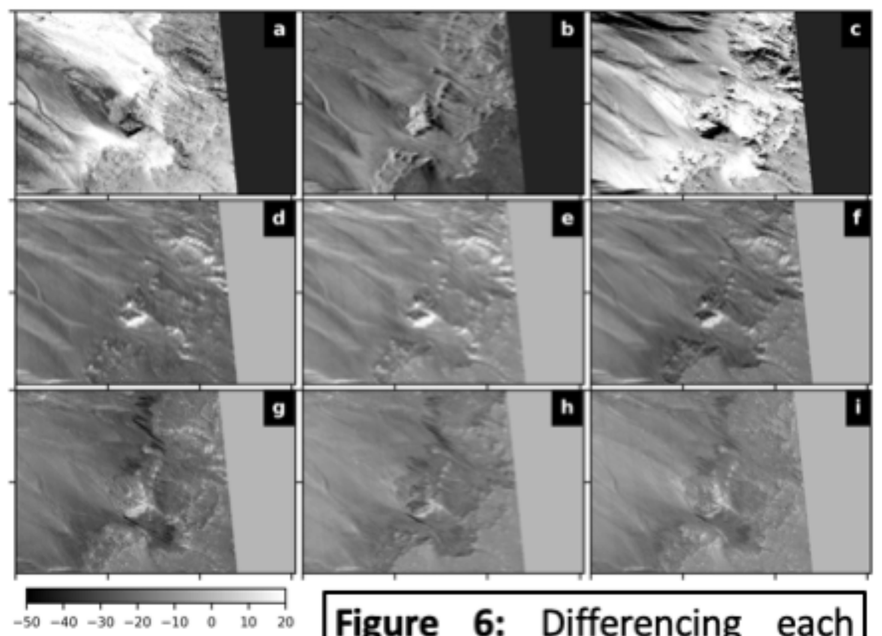


Figure 6: Differencing each image in the Palikir time-series from an average image.

Result:

Fig.5 shows the result of the PC transform (performed in ENVI®) on the full time-series. Because most of the variation in the data is captured in PCA1, representing an average image, it uses a color scale associated with Fig.4. PCs 2 and 3 are dominated by differences in lighting and possibly dust/frost deposition/erosion. RSL features are identifiable in PCA4 and dominate PCA5 with subsequent components capturing smaller variations in lighting, RSL shape (PCA8 & 9), or camera noise.

It is much easier to identify features of interest in the PCA images than the original time-series (or in differenced images) – essentially highlighting locations which have experienced change. This makes finding changes easier over large areas with multiple, overlapping images. Once locations are identified, it is fairly trivial to find the images responsible for the change. However, if a location accumulates dozens of images, the PCs can be designated with a 0, 1 , or -1 if the PC doesn't contain the feature, contains the feature, or contains the feature in the negative, respectively. This "feature vector" is then multiplied by the variance vector and the matrix involved in the PCA transformation (see table). Although not 100% reliable, this vector product gives the probability of the original image containing the feature. (see Fig.7 inset).

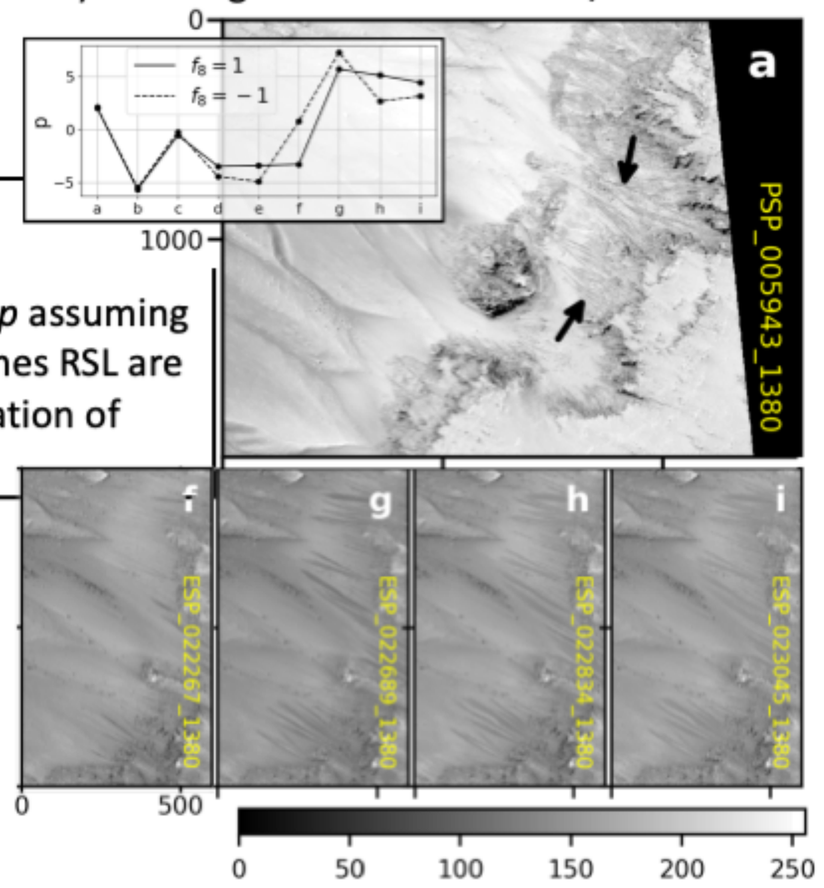


Figure 7: Comparison of images containing RSL. Inset gives the "feature probability" value, p, for all images in the time-series. The solid line gives p assuming RSL are blue in PCA8 while the dashed line assumes RSL are red in PCA8. The dashed line gives a better indication of which images contain RSL.

Conclusion:

PCA is a valuable tool for detecting change in single-band, time-series images collected on Mars or other terrestrial objects. If images of the same resolution can be accurately co-registered, temporal changes can be extracted using PCA. We encourage that PCA images be included as a PDS data product for locations with high (~5+) image counts to aid in the detection of ongoing geologic activity.

Figure 5: PCs which result from plotting the image pixels on axes given by the eigen vectors of the image covariance matrix (see table).