**A BIG DATA AND AI-DRIVEN APPROACH FOR ANOMALY DETECTION ON THE LUNAR SURFACE**

V. T. Bickel[1,2,3], J. Burelbach[1], B. Moseley[1,4], N. Relatores[1,5], [1]NASA Frontier Development Lab, USA, [2]Max Planck Institute for Solar System Research, Germany (bickel@mps.mpg.de), [3]ETH Zurich, Switzerland (valentin.bickel@erdw.ethz.ch), [4]University of Oxford, UK, [5]Carnegie Observatories, USA.

**Introduction:** Anomalies on the lunar surface are indicative of the turbulent endo- and exogenic evolution of the Moon. In addition to their scientific value, specific anomalies could enable a permanent and sustainable human presence on the Moon. For instance, magnetic anomalies could provide shelter from solar and cosmic radiation [1]. Other anomalies may indicate the abundance of certain materials that could boost future ISRU efforts, such as metallic deposits, potentially emplaced by impacting iron-nickel meteorites [e.g.2]. To support the search for and the classification of lunar material-related anomalies using the existing big data archives, we 1) apply unsupervised machine learning to create a new type of thermal anomaly map and 2) construct a global, ML-ready data stack to classify anomalous regions by their corresponding physical properties.

**Methods and results:**

**1) New anomaly maps:** To search for thermal anomalies, we use temperature measurements recorded by LRO's Diviner instrument (channel 7) from 2010-2019. We extract Diviner temperature profiles from 42 areas of interest (AOI), including craters, pyroclastic deposits and magnetic swirls, with areas ranging from 1 - 40,000 km$^2$. Thirty-eight of these AOIs have anomalies in their average annual temperature or magnetic field potentially indicative of metallic deposits [e.g. 3,4,5]; five are background locations which do not show any anomalies. We train a variational autoencoder (VAE) to encode the measured variations in lunar surface temperature over the lunar day at the selected AOIs. We find that the VAE needs only 4 latent values to accurately reconstruct each profile, and furthermore the latent values correspond to physically-interpretable factors of variation in the temperature measurements. We use these latent values to generate new types of global, thermal anomaly maps. Examples of these maps are shown in Fig. 1a-c for Tycho crater. In particular, Figs. 1.1-1.4 show how each latent variable influences the reconstructed temperature profile while keeping the others fixed. We observe that latent variable 1 responds to the onset of the profile, latent 2 to the peak temperature, latent 3 to the evening/night temperature sidelobe and latent 4 to a residual. The first three learned latent variables therefore have a clear physical interpretation, respectively correlating to the Sun's relative position, the amount of absorbed radiation and the thermal inertia of the surface. We verify this interpretation with numerical simulations based on a thermal physics model [6]. We note that latent 3 reaches high values on the western side and center of the crater, which could indicate high thermal conductivity, potentially related to specific materials. For comparison, three existing maps of opto-thermal surface properties are shown in Figs. 1d-1f.

**2) Anomaly classification:** Alongside the VAE-driven analysis, we create a ML-ready dataset by aligning and fusing 14 publicly available lunar datasets spanning a wide range of physical measurements. Among others, these include surface brightness over 7 UV-visible wavebands and 4 IR bands, rock abundance, lunar albedo and magnetic field strength [e.g. 7,8]. To further support our search for material-related anomalies, we apply dimensionality reduction (PCA) and clustering to this dataset. Our clustering analysis is performed in a reduced 3D space whose principal components are closely related to the overall brightness, the relative visible/IR brightness and the magnetic field strength of an area. Fig. 2 shows our physical interpretation of each of the 7 identified clusters and the spatial distribution of each cluster at 6 AOIs. We observe that the clusters are spatially correlated; e.g. Marius-A, Hell-Q and Giordano Bruno show physical anomalies in their centers, but not in their surroundings.

**Preliminary conclusion:** We have shown that unsupervised learning can aid the detection of physical anomalies on the lunar surface. Using a VAE has improved our understanding of the variational factors in surface temperature. Our clustering method allows fast identification and classification of anomalous craters. Our AI- and data-driven method could help to refine existing model-driven approaches, e.g. for the determination of rock abundance, CF, thermal inertia, and other.

**Acknowledgements:** The initial results have been produced during NASA's Frontier Development Lab (FDL). We would like to thank FDL and its sponsors (Luxembourg Space Agency, Google Cloud, Intel AI, HP Enterprise, Element AI and USGS); the SETI Institute for hosting the FDL; Jean-Pierre Williams for his help in working with Diviner; as well as Dennis Wingo, Allison Zuniga and our other FDL mentors.

**References:** [1] Bamford et al. (2014). *Acta Astro*., 105 (2), pp. 385-394. [2] Wieczorek et al. (2012), *Science, 335, 6073, pp. 1212-1215*. [3] Williams et al. (2018), *JGR:Planets, 123(9):2380–2392.* [4] Zheng et al. (2014), 45th *LPSC*. [5] Hu et al. (2018), *TGRS*, 56:1–10. [6] Hayne et al. (2017), *JGR:Planets*, 122(12):2371-2400. [7] Bandfield et al. (2011), *JGR:Planets*, 116(E12). [8] Williams et al. (2017), *Icarus*, 283:300–32

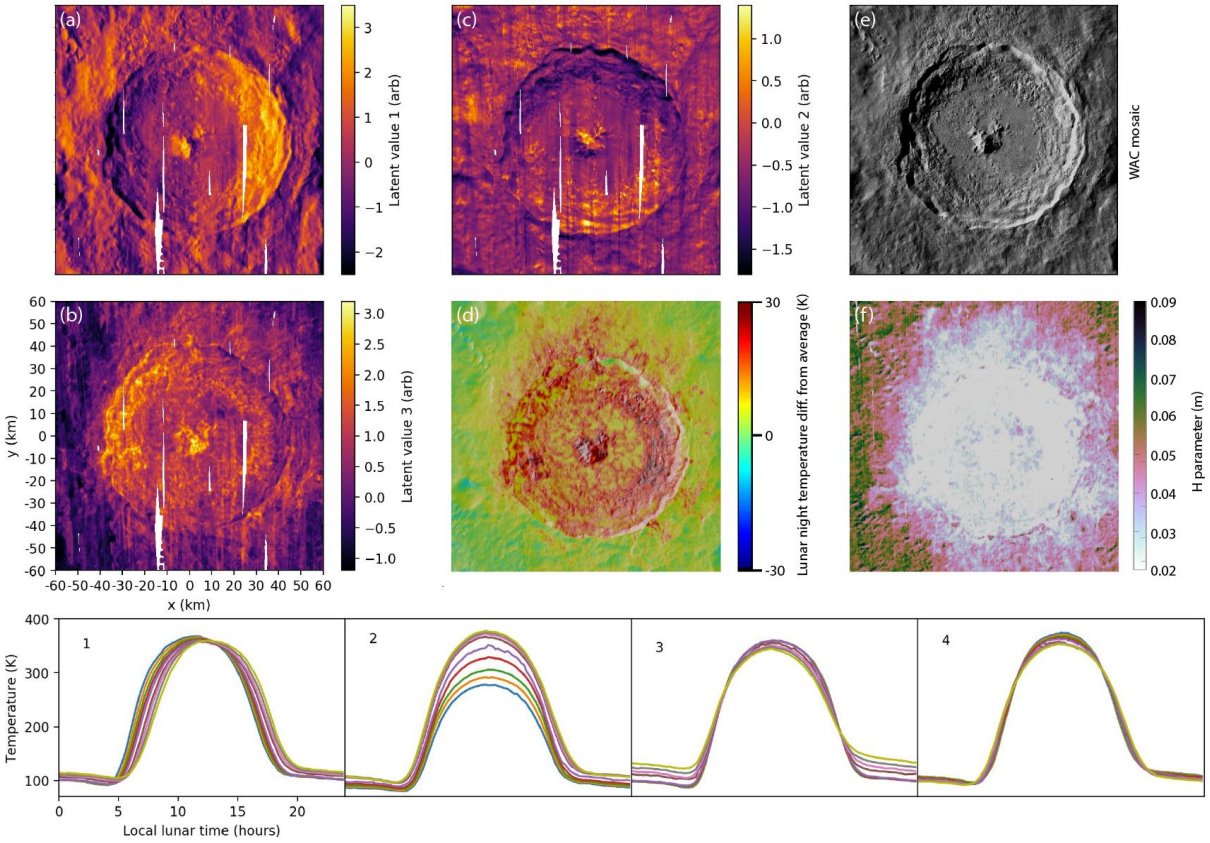


**Fig. 1. | New anomaly maps:** (a, b, c) VAE latent variables 1, 3, and 2 over the Tycho crater. (d) Diviner lunar night temperature difference from average, (e) Tycho crater WAC mosaic (visible range), (f) Diviner H parameter map [3]. The VAE latent variable 3 (b) potentially indicates different physical properties and spatial pattern than temperature and thermal inertia maps (d and f). Bottom plots show the reconstructed profiles generated from the VAE when sampling each latent variable independently, (1 and (a), 2 and (c), 3 and (b); 4 is a residual). Local lunar time is in local lunar "hours" where 12 is lunar noon and 0 is lunar midnight.

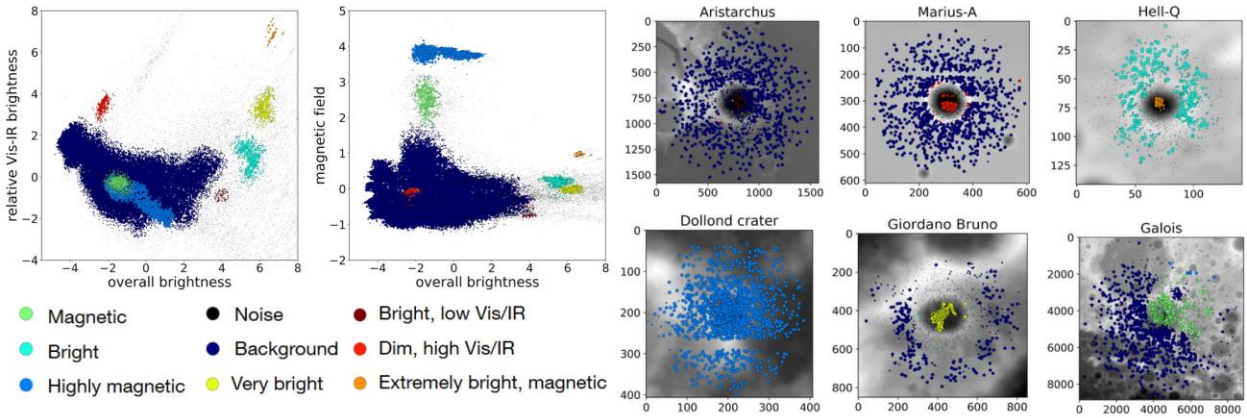


**Fig. 2. | Anomaly classification:** Left: Scatter plots of relative visible/IR brightness (PC 2) against overall brightness (PC 1), and magnetic field strength (PC 3) against overall brightness, plotting feature vectors over all crater AOIs. Points are color coded by their predicted cluster; legend shows our physical interpretation of each cluster. Right: Elevation maps of six craters, with spatial overlays of their feature vectors, colored by cluster (scale in hectometers).