DEEP REDATUMING OF CHANDRAYAAN-2 LARGE AREA SOFT X-RAY SPECTROMETER (CLASS) DATA FOR CHEMICAL MAPPING OF THE LUNAR SURFACE. Karpoor, P.R.[1], Bharadwaj, P.[1], Narendranath, S.[2], Pillai, N.S.[2],[1]Indian Institute of Science, Bengaluru, India preethi.karpoor@gmail.com, [2]U R Rao Satellite Centre, ISRO, Bengaluru, India .

Introduction: Chandrayaan – 2 Large area Soft X-ray Spectrometer (CLASS) [1] is an X-ray fluorescence experiment onboard the second Indian lunar mission Chandrayaan 2 [2]. The primary science objective of CLASS is to map the major elemental abundances of the lunar surface globally. From the polar circular orbit of 100 km altitude of Chandrayaan-2, the best spatial resolution [3] achieved with CLASS is 12.5 km x 12.5 km, which is an order of magnitude better than such experiments in the past.

CLASS has been providing a large amount of lunar X-ray fluorescence (XRF) data. However, the inverse map from the observed spectrum to the elemental abundances is far from straightforward—complications arise due to 'nuisance' or 'irrelevant' parameters. For example, the XRF data lines that describe the elemental abundances are heavily dependent on the phenomena related to the incident solar spectrum. The solar spectrum's strength and shape rapidly change during the flares. The XRF data also depends on other factors like the geometry of observation, matrix effect, and grain size.

As opposed to the current inversion methodologies [4,5,6,7] that rely on the forward modeling of the underlying physics, this paper aims to develop deep-learning-based methods to redatum the nuisance effects of the CLASS data. Towards deriving elemental compositions, we show that redatuming allows accurate identification of the XRF signal from the background. High spatial resolution datasets such as CLASS result in millions of spectra over the years. Conventional methods are tedious and time-consuming to execute. Their accuracy is questionable when the physics of the forward modeling is incomplete. It becomes imperative to adapt machine learning algorithms to utilize big data in astronomy and learn the underlying signal structure in this context. Such algorithms can potentially filter out nuisance variations towards a reliable identification of weaker signals from minor elements [8].

Solar variability is one of the critical factors that obscure the XRF lines. For example, C1XS of Chandrayaan 1[9] reported this issue, where the solar spectrum undergoes a non-linear variation due to solar flares. Even in the NEAR (Near Earth Asteroid Rendezvous)[10] mission of NASA that aimed to map the surface chemical composition of Asteroid Eros observed rapid fluctuations in solar activity that led to their models being less reliable for further analysis. Finally, in any space mission where data gathered is dependent on solar variability, there is always a good margin for inaccuracy in modeling.

Challenges of these sorts lead to great difficulty and loss of accuracy in the determination of chemical abundances. In our novel approach, which we describe in the forthcoming sections, we aim to avoid the difficulties mentioned earlier almost entirely by not resorting to modeling. Instead, we produce XRF spectra as if virtual solar radiation illuminates a particular region of interest, allowing us to do a more accurate elemental estimation.

Deep Redatuming:

Redatuming is a crucial step for correcting experimental data in earth and planetary sciences. It accounts for the nuisance variations due to complexities and irregularities in the data acquisition (e.g., during complex space missions), which is near impossible to control. For example, measured data in seismology from uncontrollable ambient-noise sources (similar to the sun) are redatumed to generate "virtual sources" [11]. The idea is that the data from these virtual sources can aid subsequent imaging without any hindrance due to the nuisances. In contrast to physics-based redatuming commonly used in geophysics, we have adopted a recently developed symmetric auto-encoder architecture named 'SymAE' [12] to perform data-driven redatuming.

SymAE is a semi-supervised model that relies on identifying a group of instances, which share coherent information related to the phenomena of interest. In our case, we consider the recorded XRF data from CLASS under various conditions. These data are calibrated and available through the Indian Space Science Data Centre (in PDS4 standard). The data instances are scoured through to identify a consistent or coherent parameter (for example, data that may be recorded from a particular geographical location on the Moon, etc.) that comes coupled with incoherent parameters, like solar nuisances. Such datasets are then grouped together into bags of

instances to generate labeled datapoints. Multiple bags are processed using SymAE to extract the information on the surface chemical composition of the Moon. Specifically, SymAE is trained to represent these instances while disentangling the coherent information from the remaining nuisance variations. Physical symmetry is embedded into one of SymAE's encoders to achieve this disentanglement. Finally, SymAE's repres entation can then be used to redatum nuisances from one instance to another to generate virtual instances for further analysis.

Finally, this method has tremendous scope to be generalized to a wider variety of similar inverse problems.

References:

[1]Vanitha et al, 2021, LPSC abstract 1994
[2]Radhakrishna et al, 2020, Current Science, 118,29.
[3]Netra et al, 2021, Icarus, 363, 114436
[4]Narendranath et al, 2011, Icraus, 214, 53.
[5]Nittler et al, 2020, Icarus, 345, 113716.
[6]Ogawa et al, 2008, Earth Planets Space, 60, 283.

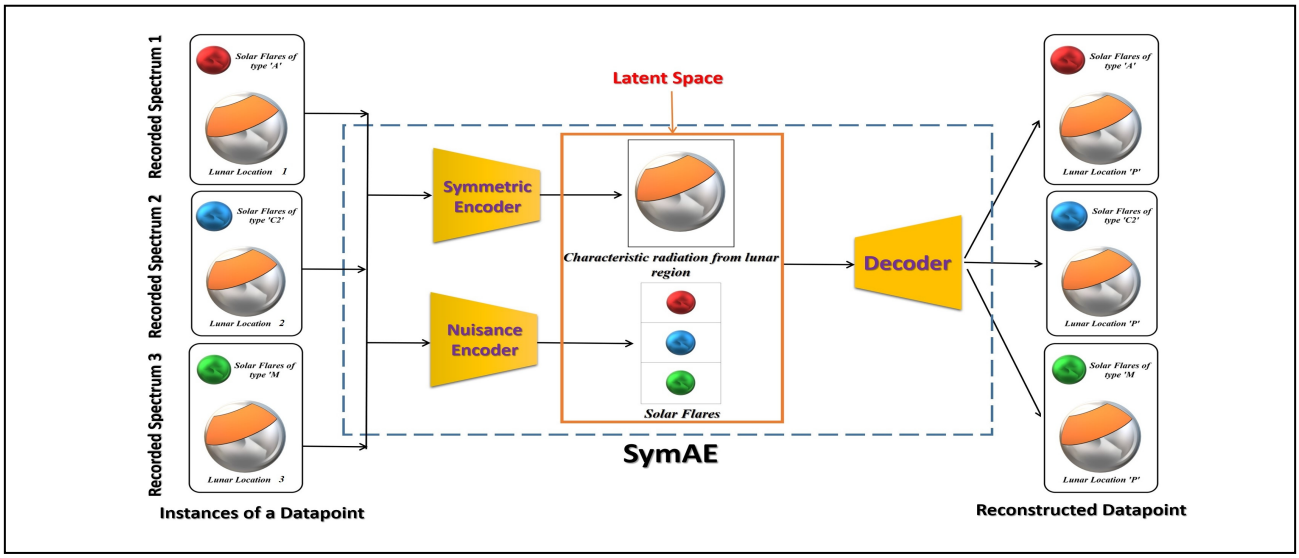

Fig 1 : SymAE assisted Workflow for Deep Redatuming of CLASS Data for Chemical Mapping of Lunar Surface

Fig 1 illustrates SymAE's network architecture used for the nuisance disentanglement. Here, the coherent information is due to the fluorescence from a lunar region in orange is separated in the latent space. We train SymAE to learn the redatuming of the solar spectrum from one instance to another. As a result, we show that any two instances that describe the chemical compositions of different lunar regions can now be compared even when the solar spectrum is variable. Such a comparison allows us to comment on the accuracy of the solar-spectrum modeling and helps us better constraint the subsequent inversion algorithms. Also, the chemical maps generated by modeling can be validated, verified for their accuracy, and enabled to serve as training data for future machine learning implementations.

To conclude, this abstract presents an application of semi-supervised deep redatuming towards an accurate inference of the chemical composition of the Lunar surface. Our approach undertaken in this work is novel and complementary to the existing physics-based redatuming practices.

[7]Weider et al, 2011, PSS, 59, 1393.
[8]Narendranath et al, 2021, LPSC abstract 2230
[9]Grande et al, 2009, PSS, 57, 717
[10]Nittler et al, 2001, Meteorit. Planet. Sci., 36, 1673
[11]Snieder, R & Wapenaar,K, 2010, Physics Today,63.9,44
[12]Bharadwaj P. et al. (2020), Society of Exploration Geophysics Annual Meeting Abstracts.