Multiphysics property prediction from hyperspectral drill core data

Akshay Kamath[1], Samuel Thiele[1], Moritz Kirsch[1], and Richard Gloaguen[1]

[1]Helmholtz-Zentrum Dresden-Rossendorf, Helmholtz Institute Freiberg for Resource Technology, 09599 Freiberg, Germany.
Correspondence: Akshay Kamath (a.kamath@hzdr.de)

Abstract. Hyperspectral data provides rich quantitative information on both the mineralogical and fine-scale textural properties of rocks, which, in turn, largely control their petrophysical characteristics. We therefore developed a deep learning model to predict petrophysical properties directly from hyperspectral drill core data. Our model learns relevant features from high-dimensional hyperspectral data and co-registered sonic, gamma-gamma density and gamma-ray logs to infer slowness, density, and gamma-ray values. We demonstrated the performance of this approach on data acquired in the Spremberg region of Germany. Our results demonstrate that with meticulous pre-processing steps and thorough data cleaning, one can overcome the difference in capturing resolution and learn the relationship between hyperspectral data and petrophysics. Using a test dataset from a spatially independent borehole, we generate a pixel-resolution (≈ 1 mm^2) model of the petrophysical properties and resample it to match the measured logs. This test indicates substantial accuracy, with R^2 scores and root-mean-squared errors (RMSE) of 0.7 and 16.55 μs.m^{-1}, 0.86 and 0.06 g.cm^{-3} and 0.90 and 15.29 API for the slowness, density and gamma-ray readings respectively. Overall, our findings lay the groundwork for building deep learning models that can learn to predict physical and mechanical rock properties from hyperspectral data. Such models could provide the high-resolution but large-extent data needed to bridge the different scales of mechanical and petrophysical characterisation.

Some food for thought... If I'm using gamma logs say as input parameters to predict formation, and I also have VNIR-SWIR-MWIR-LWIR data for the same holes and find that it can accurately predict gamma values, why not directly predict lithology?

Same logic for sonic logs... If the reason for acquiring sonic logs is to log porosity/permeability, why not directly predict that rather than travel times?

This is just a thought exercise, to spur you to think about end-user applications, it in no way invalidates your work.

1 Introduction

15 Hyperspectral imaging provides detailed insights into the mineralogical composition of rocks by capturing minute spectral signatures across extensive wavelength ranges. This technology enables rapid, non-invasive characterisation and spatial analysis, with diverse applications for geological mapping, minerals exploration and geometallurgy. These applications typically leverage the unique combination of millimetre-scale spatial resolution and kilometre-scale extent possible with hyperspectral drill core or outcrop scanning methods (Thiele et al., 2021, 2024; Laukamp et al., 2021).

20 Geophysical well logging has been employed to acquire petrophysical data from boreholes utilising a suite of tools to measure various properties in situ. This process includes the collection of gamma-ray, density, and sonic logs, among others, which provide continuous and high-resolution records of the subsurface properties. Gamma-ray logs are often used to identify shale-rich layers by measuring the natural radioactivity of the formations, primarily from potassium, thorium, and uranium content (Serra, 1984). Density logs, which measure bulk density through the attenuation of gamma rays, can be used for the

25 estimation of porosity and lithological differentiation (Schlumberger, 1972). Sonic logging is used to measure the travel times of acoustic waves travelling through the formations, providing insights into both the mechanical properties and porosity of the rock, through the measurement of slowness (inverse of P-Wave velocity) (Bourbié et al., 1987).

While hyperspectral data are primarily known for their sensitivity to mineralogy (Clark, 1999), they are theoretically also sensitive to textural attributes such as grain size, especially in the mid- and long-wave infrared regions. Initial work by Keresz-

30 turi et al. (2023) and Lee et al. (2023) has built on these theoretical connections to successfully predicting petrophysical and geomechanical properties, albeit on relatively small sample sets. In this contribution, we build on this work to test if 'hyperspectral upscaling' workflows that were developed to predict mineralogy (e.g. Thiele et al., 2024) can also be used to map petrophysical properties along drill cores at high (mm scale) spatial resolution. In doing so, we aim to both enhance the spatial resolution of down-hole petrophysical logs and work towards potentially generalisable methods that could, one day, be used to

35 predict important petrophysical and mechanical properties across large drill core libraries and hyperspectral scans of outcrops. Our objectives are thus twofold: first, to assess the feasibility and accuracy of predicting petrophysical properties directly from hyperspectral drill core data, and second, to use the extensive spatial coverage and high resolution of hyperspectral data to upscale petrophysical measurements acquired from geophysical well logging.

2 Geological Setting

40 The Central European Kupferschiefer is Europe's largest Cu and Ag-resource and one of the most prolific sediment-hosted copper districts globally (Borg et al., 2012). The strata-bound mineralisation occurs in a thin but laterally persistent carbonaceous shale unit along the southern margin of the Permian-aged central European Basin, with known deposits stretching from the Rhine area of western Germany to south-central Poland.

The Spremberg–Graustein deposit in Lusatia, southeastern Germany, was discovered in 1953 and intensively explored be-

45 tween 1953 and 1974 (Kopp et al., 2012) Renewed exploration interest led to another deep drilling campaign from 2008 to 2012, producing the drill core material and geophysical logging data used in this study. These drill cores sample a large part

of the Central European Basin's stratigraphy, from the basal Rotliegend terrestrial conglomerates, sandstones, shales deposited in the Permian, overlying Zechstein carbonates and evaporites, and lower to middle Triassic Buntsandstein formations. The Kupferschiefer sensu stricto is a bituminous calcareous shale usually less than a metre thick (Spieth, 2019) at the base of the
50 Zechstein Formation, and is overlain by several cycles (Werra, Stassfurt and Leine) of marine sedimentary rocks, including dolostone and limestones, anhydrite and gypsum, and evaporites.

The orebody at Spremberg–Graustein occurs at a depth of ca. 850–1580 m, and hosts at least 100 Mt of Cu–Ag ore with average copper grades of 1.5 to 2.0. For this study, two drill holes, KSL133 and KSL131, were chosen to provide training data, due to their comprehensive coverage of the key lithological units and availability of downhole geophysical well logging data.
55 A third, spatially separated borehole, KSL136, was used exclusively as test data to assess the accuracy of our predictions.

3 Methodology

3.1 Hyperspectral data acquisition

Hyperspectral data capturing the three drill holes (KSL133, KSL131 and KSL136) were acquired using a SPECIM SiSuRock drill core scanner, which is equipped with three hyperspectral cameras. These sensors capture a broad range of hyperspectral
60 data across different spectral regions: the visible-near to short-wave infrared (VNIR-SWIR; 380–2500 nm), mid-wave infrared (MWIR; 2700–5300 nm), and long-wave infrared (LWIR; 7700–12300 nm). The extensive spectral coverage captures spectral features that are diagnostic of most common minerals, including those expected at Spremberg (quartz, anhydrite, carbonates, clays, and feldspars) (Géring et al., 2023). An in-depth review of the data acquisition process, as well as the details of the Kupferschiefer hyperspectral dataset can be found in Thiele et al. (2024).

65 ### 3.2 Petrophysical data acquisition

Sonic logger data were acquired using the USBA-21 acoustic logger that measures the travel time of acoustic waves between two magnetostrictive transducers (sources) and a piezoelectric receiver. Two sources are used to differentiate between formation and mud effects (since the the travel time through the mud stays constant)(Sheriff and Geldart, 1995). This travel time, recorded in microseconds per meter (a measure of slowness), is given by:

70 $$\Delta t = t_{far} - t_{near} \tag{1}$$

where t_{far} is the travel time from the farther transducer and t_{near} is travel time from the nearer transducer.

For density logging, a Century-9036 density logging tool was used. The radioactive source, typically a directional Cs^{137} emitter, emits medium-energy gamma rays into the formation. These gamma rays undergo Compton scattering upon interacting with electrons in the formation, and the scattered gamma rays reaching the detector provide a measure related to the formation's
75 electron density. This electron density (ρ_e) is then associated with the formation's bulk density (ρ_b) using the relationship:

$$\rho_e = 2\rho_b \frac{Z}{A}$$

where Z is the atomic number, and A is the molecular wei⋯ the response of the density tool (Baker, 1957).

Gamma-ray logging quantifies gamma radiation levels⋯ and potassium (K) concentrations. The gamma-ray intens⋯ concentrations as follows:

> Downhole geophysical tools typically start measuring distance from surface at 0 and measure in a linear fashion downhole based on how much line has been unspooled. Really they have the best depth registration of virtually all the drillhole analysis methods, including core scanning. Core scanning hardware typically registers depth between driller blocks or on a per-box basis. Not as accurate, and depending on the circumstances may be significantly different from the depths provided by wireline geophysical tools. Section 3.3 does not address this issue of co-registration. Or perhaps there is an underlying assumption that the scanned data depths are accurate and correspond to the geophysical depths? Either way this should be addressed or at least acknowledged. It's one of the greatest barriers to performing ML workflows on drillhole data.

$$GR\ (\mathrm{API}) = 8 \times \mathrm{U}(ppm) + 4 \times \mathrm{Th}(ppm) + 16 \times \mathrm{K}(wt.\%) \tag{3}$$

This reading allows for comprehensive characterization of radioactive content in the formation, providing insights into its mineralogical composition (Asquith et al., 2004).

3.3 Data co-registration

Given the varying resolutions and sensor positions inherent in the different geophysical logging tools used to measure slowness, density and gamma-ray values, it was necessary to resample them to derive measurements over comparable depth ranges and ensure comparability. Given that the lowest resolution data was the slowness log, with a sampling of 10 cm, we employed *scipy*'s (Virtanen et al., 2020) `RBFInterpolator` to downsample the logs. Specifically, a thin-plate spline kernel and a different number of nearest neighbours for the slowness, density, and gamma-ray logs—10, 50, and 20 nearest neighbours respectively (to make sure that the truncation distance for the interpolator was the same for each property) were used, to downsample the logs and match the resolutions while minimising potential interpolation bias and ensuring consistent alignment across each sensor.

Finally, a rolling standard deviation of the measured petrophysical properties was computed using a 1-metre window (Fig. 1). This was used to exclude high variance regions from the training data, to limit potential co-registration errors associated with uncertainties in the depths recorded by both the petrophysical and hyperspectral logs. The hyperspectral data also needed to be downsampled by two orders of magnitude to match the spatial resolution of the petrophysical measurements, given its $\approx 1\ \mathrm{mm}^2$ spatial resolution (two orders of magnitude higher than the downhole geophysical logs). We computed the median spectra within a defined window of $10 \times 5\ \mathrm{cm}^2$ (equivalent to $100 \times 50\ \mathrm{pixels}^2$) centred along the core. The results represent the (linearly mixed) spectral response of the rock volume sampled by the petrophysical loggers, with some assumptions on e.g., the representativeness of the scanned core surface.

3.4 Spectral processing

In addition to the standard white-reference normalisation, lens correction and sensor co-registration described by Thiele et al. (2024), the hyperspectral data were further preprocessed to (1) remove deleterious noisy bands, (2) reduce illumination effects

105 that result from the 3D geometry of the scanned drill cores (e.g., shadowing), and (3) enhance mineralogically significant

absorption features (Fig. 2).

Specifically, the VNIR-SWIR data w[...]he

MWIR and LWIR data were also remov[...]g

hylite (Thiele et al., 2021), which simul[...]y

110 associated with drill core shape and ass[...]

You've used HBSCAN to cluster the data and further identify noise (class -1) which you've removed from the dataset. You've mentioned resiliency to hyperparameter selection--- it is actually very well established that HDBSCAN outcomes are highly sensitive to hyperparameter selection, particularly the distance metric, min_samples, and min_clusters parameters. sklearn can automate hyperparameter selection using Randomized Search Cross Validation, which seeks to optimize the validity index for iterated hyperparameters, otherwise hyperparameter tuning is highly manual and hugely impacts the number of clusters and cluster distributions returned. I'd strongly suggest addressing this.

3.5 Data balancing

The petrophysical and associated hyperspectral data form distinct clusters, as clearly shown when they are plotted in the

petrophysical property space (Fig. 3A). After using the rolling standard deviation to remove high-variance points and associated

mixed measurements (Fig. 3B), these clusters are even more distinct. Importantly, the cluster sizes are unbalanced due to

115 different thicknesses of the lithologies in the boreholes, leading to an overrepresentation of abundant units in the training

dataset.

The Hierarchical Density-Based Spatial Clustering of Applications with Noise (HDBSCAN) algorithm was used to identify

these clusters and mitigate their translation (as representation bias) into the deep learning models. HDBSCAN is well adapted

for its ability to identify clusters of varying densities and its resilience to parameter selection (McInnes et al., 2017). We

120 used HDBSCAN to automatically identify clusters in a dimensionality reduced feature space containing the depths of each

measurement (to include a spatial component) and the 30 normalised principal components of the median spectra for each data

point. This clustering resulted in a total of 24 distinct groups, with one additional noisy cluster excluded from the dataset (Fig.

3C) to leave 2946 remaining spectrum-property pairs.

A stratified sampling strategy was then employed to select well distributed training and validation subsets from drill cores

125 KSL131 and KSL133 (KSL136 was kept separate for independent testing), using a six-fold stratified train-validation split

implemented by the `StratifiedShuffleSplit` utility in *scikit-learn* (Pedregosa et al., 2011). This stratification ensures

the data diversity is captured in both the train and validation datasets, while a weighted loss function was developed (see Sec.

3.6) to ensure the differing cluster sizes do not bias our model.

you describe using stratified sampling; what was the category that you stratified? The drillhole ids? The HDBSCAN clusters? Did you stratify to bins of a petrophysical parameter? Please clarify.

3.6 Model architecture and loss function

130 Our model was developed using PyTorch, an open-source machine learning framework known for its flexibility and strong sup-

port for deep learning applications (Paszke et al., 2019). We implemented a multi-headed variant of a 1D Convolutional Neural

Network (CNN) architecture (Albawi et al., 2017). The 1D CNN enhances feature extraction by incorporating convolutional

layers within a Multi-Layer Perceptron (MLP) framework, using kernels with learnable weights that are convolved with the

input data (LeCun et al., 2015).

135 Each spectral range— VNIR-SWIR, MWIR, and LWIR—was assigned a separate head within the CNN (Fig. 4), to allow

sensor-specific feature extraction. This configuration was chosen to accommodate the different spectral sampling resolutions

provided by each sensor and the distinct spectral characteristics associated with each of the respective spectral ranges. By

tailoring the kernel sizes to each spectral range, the model was better adapted to identify and capture relevant features across the diverse spectral data. The extracted features were then concatenated and fed through additional fully connected layers

140 before outputting the petrophysical predictions. To ensure balanced contributions from each of the three outputs during loss computation, the slowness, density, and gamma-ray counts were normalised by factors of 10^{-3}, 10^{-1}, 10^{-3} and respectively, to ensure that their values fall between 0 and 1.

A Mean Squared Error (MSE) loss function was used for training, with the following key adjustments. First, to address the imbalance in the number of data points per cluster, we multiplied each output loss by a factor inversely proportional to the

145 number of points in its assigned cluster (see Sec. 3.5). This modification ensured that the model treated each cluster equally, avoiding the disproportionate influence of more populous clusters. Each data point in the kth cluster was associated with a weight w_k given by:

$$w_k = \frac{N_k^{-1}}{\sum_{j=1}^{N_c} N_j^{-1}} \tag{4}$$

where N_k is the number of points in the kth cluster in the training set, and N_c is the total number of clusters.

150 Secondly, we introduced a penalty term into the loss function that discouraged negative outputs, as none of the predicted properties can be negative. This ensures that the model predictions stay within valid (positive) bounds by increasing the loss in proportion to the square of the sum of all negative outputs. A penalty term was used instead of e.g., a ReLU activation function to avoid issues with "dying neurons" in the final layer Lu et al. (2019). The final loss function utilised for training was thus:

$$Loss = \sum_{i=1}^{n} \left[\frac{1}{n} \left(w_i \left(\mathbf{m}(x_i) - y_i \right) \right) + \lambda \left(\max\left(-\mathbf{m}(x_i),\, 0 \right) \right) \right] \tag{5}$$

155 where $\mathbf{m}$ is the model, x_i is the input spectrum, y_i are the corresponding petrophysical properties, λ is the penalty factor for negative outputs, n is the batch size, and w_i is the cluster-defined weight (Equation 4) associated with the ith datum.

The training dataset was shuffled and split into six parts, with the model trained iteratively on each split. Every 100 epochs, a new split was introduced for training and validation, effectively reshuffling the data to help prevent overfitting. Over 600 total epochs, the model iteration with the lowest validation loss was chosen as the best-fitting model. The use of Stratified Shuffle

160 Split preserved data distribution across splits, reducing overfitting without needing dropout layers or batch normalisation, common in CNNs but sometimes performance-limiting. As mentioned previously, a spatially separated borehole (KSL136) was then used to test the model and evaluate its accuracy, after the training phase was completed.

3.7 Shapley analysis

Finally, to help understand the spectral properties that influenced the model predictions, we performed a Shapley value anal-

165 ysis. Shapley values are derived from cooperative game theory, and have been widely employed in the machine learning context to quantify each input feature's contribution to model predictions (Lipovetsky and Conklin, 2001). We apply Shapley

values using the Python package *shap* (**SH**apley **A**dditive ex**P**lanations) (Lundberg and Lee, 2017), specifically leveraging the `DeepExplainer` utility optimised for deep learning models. This allowed us to efficiently compute feature impacts, enhancing interpretability and transparency in our neural network.

170   To reduce computational complexity, we subsampled the data before computing Shapley values, selecting 10 random samples from each of the HDBSCAN clusters. This ensures that the samples extracted for Shapley analysis are representative of the entire dataset. We then visualised the results in the form of a raster image to help qualitatively assess the spectral regions used by the model for each of its predictions. Lastly, to better understand the impact of certain bands on the outputs, we computed the mean Shapley values for the top ten percent of each predicted output, to explore which factors resulted in a sample being

175   predicted to be fast (i.e., bottom 10 percent of the slowness), dense or exhibiting high gamma ray counts.

4   Results

4.1   Model accuracy

The model predictions for the training set showed scores of 0.990, 0.996, and 0.964 for the slowness, density and gamma-ray values respectively (Fig. 5), suggesting the model explains more than 96% of the petrophysical variance in the training set. The

180   corresponding RMSE for these properties was 2.544 μs.m^{-1}, 0.023 g.cm^{-3}, and 4.993 API, respectively.

To better assess the model's accuracy for upscaling applications, we applied it to every 1 mm^2 pixel in the hyperspectral dataset (including the test drill core), to create a high-resolution (upscaled) map of the petrophysical properties. The results identify fine-scale and cross-core variations that could not be captured by traditional running log measurements (Fig. 6).

Our workflow for hyperspectral downsampling (see Sec. 3.3) was then run on these petrophysical property prediction maps,

185   to derive an estimate of the petrophysical logging data (noting that the predicted properties slowness, density and gamma-ray readings are all additive, so can be arithmetically averaged).

The resulting predicted logging data showed a close match to the measurements (Fig. 7), both for the two training boreholes and for the independent test hole (KSL136). R^2 scores for this test hole were 0.86 and 0.9 for the density and gamma-ray logs, respectively, indicating very reasonable accuracy on unseen data, which even included a basement lithology that was

190   not sampled by the other training drill cores (highlighted by the grey box labelled Tonschiefer i.e., argillaceous shale). The slowness prediction in KSL136 showed a relatively lower R^2 score of 0.7, with most of the erroneous predictions lying within the unseen lithology (Tonschiefer basement). The measured sonic log here shows significant fluctuations, whereas our model prediction remains steady (suggesting the lithology is spectrally quite uniform).

4.2   Shapley analysis

195   The calculated Shapley values (Fig. 8), represent the quantitative importance of every input feature in the form of a push away from the mean prediction. Since the theoretical foundation of Shapley analysis looks at every band as a separate feature and not necessarily the overall shape of the absorption feature itself, the overall contribution from a particular spectral range cannot be

ascertained from looking at the raw Shapley values. However, a single band in the LWIR, on average, contributes more towards the output than a single band in the VNIR-SWIR-MWIR ranges.

200   To look at the total contribution each spectral range makes towards the output, we use the additive nature of Shapley values to create a stackplot (Fig. 9). The plot shows that the contributions from the VNIR-SWIR range are minimal in all predictions. The net contribution from the MWIR bands is the largest, followed by the LWIR bands for slowness predictions, except for the fastest samples, where the LWIR contribution exceeds that of the MWIR bands. For the density predictions, the MWIR dominates the predictions across all the samples, closely followed by LWIR. For the gamma ray values, the SWIR bands

205   contribute more than in the other two properties, but the LWIR appears to be the dominant steerer of the predicted value.

The median Shapley values of the samples in the bottom 10 percentile of the slowness predictions (i.e. fastest predicted P-wave velocities), the top 10 percentile of density predictions, and the top 10 percentile of gamma-ray predictions were calculated as well (Fig. 10). The hull-corrected spectra for these samples (Fig. 11) were then inspected to identify possible spectral features associated with fast, dense or high gamma-ray predictions. From Figure 10, it is clear that the model is

210   focusing on known mineral absorption features (and their surrounding bands) in each of the spectral ranges. These include, but are not limited to, the carbonate absorption feature at 11200 nm, as well as the fundamental overtones of the Silicon Oxide (νSiO) absorption feature at 8575 nm and 9200 nm. The 1400 and 1900 nm water and 2200 AlOH SWIR absorption features are also clearly being used, as are the bands around 2300 nm (which are likely influenced by the relatively broad carbonate absorption at 2345 nm).

215   The results highlight some spectral variability within these top 10 percentile groups, as would be expected, but also some commonalities. For instance, in the slowness plots (Fig. 10A), the samples show a significant push below the mean due to the bands around 11200 nm and 4000 nm, corresponding to primary and secondary overtones of carbonate absorption features. This observation aligns with our geological understanding, as the fastest samples are from the Stassfurt-Hauptdolomit formation (see Legend, Fig. 7). The spectra in Figure 11A further support this inference, as they exhibit a clear carbonate absorption feature.

220   Conversely, the densest samples in the region are associated with the Werra-Anhydrit formation (see Legend, Fig. 7). They do not exhibit the carbonate absorption feature as seen in the spectra (Fig. 11B), and the low Shapley values for those bands (Fig. 10B) confirm that the model is no longer looking at any of the carbonate absorption features.

The samples with the top 10 percent of gamma ray readings come from the Salzton (salt-clay) formation, the Stassfurt-Salz (salt) formation and the Rotliegend (Sandstone) formation. The model considers the bands corresponding to the water

225   absorption feature, as well as the clay feature in the SWIR favourably (Fig. 10C). Furthermore, there is a considerable peak in the 9200 - 9600 nm range in the LWIR, which might correspond to the tecto- and sheet-silicates absorption features (Laukamp et al., 2021) containing potassium salts from the Salzton and Stassfurt-Salz formations, or the K-Feldspar absorption feature from the Rotliegend sandstones, both of which would give rise to high gamma ray values.

> Great use of Shapley values!

5 Discussion

230 Our deep learning approach has managed to accurately predict the variability in petrophysical properties from hyperspectral data. This prompts several important questions: what is the model using to make these predictions? Are the results potentially generalisable, or likely only valid for cores from the Spremberg area? And how could this novel approach help improve our understanding of, and ability to predict, subsurface behaviour?

The Shapley values (Fig. 8, 9, 10) clearly indicate that the model is extracting features associated with the depth and shape

235 of known mineral absorption features, especially those relating to quartz, feldspars and carbonates in the LWIR range. This suggests a primary sensitivity to mineralogy, which we suggest is reasonable given the mineralogical control on all the predicted petrophysical properties. The model also clearly places a strong emphasis on the MWIR and the LWIR bands, even though much of the mineralogical variation could also be identified using these other spectral ranges. Based on this, we speculate that the model could be learning to identify relatively subtle changes in LWIR absorption feature shape that are known (Zaini

240 et al., 2012; Rost et al., 2018; Salisbury et al., 1987) to vary with grain size and surface-roughness (which we suggest is likely correlated to porosity).

Importantly, the Shapley values for predicted gamma emission differ substantially, presumably as gamma emission is primarily controlled by mineralogy. Gamma-ray emissions are triggered by natural radioactive decay of elements like Potassium, Thorium, and Uranium (Asquith et al., 2004). These are typically hosted in clay (and feldspar) minerals so the SWIR bands,

245 especially those around 1400 nm (associated with hydrated minerals) and 2340 nm (related to clay minerals) are expected to play a more prominent role for gamma prediction. The increased Shapley values in these regions indicate that the model is correctly associating higher absorption at these wavelengths with elevated gamma-ray values, as expected in shaly rocks or clays. Similarly, the predicted gamma emissions do not appear sensitive to many of the spectral regions used to predict density and slowness, such as the carbonate feature at 11200 nm.

250 Given the results from our Shapley value analysis, we suggest that it is unlikely that the model is "just" learning to distinguish different lithologies and returning appropriate (average) predictions. Instead, it appears to generate predictions based on the mineralogical and textural information captured by the spectra. This is key to its demonstrated ability to identify intra-lithology variations in each of the petrophysical properties (Fig. 7), and possibly explains why it produced broadly reasonable predictions for the unseen basement lithology. This indicated sensitivity to the fundamental rock properties linked to mechanical and

255 petrophysical behaviour is promising, and suggest that models trained on more diverse data could have generalisation potential. Hence, while we would not apply our model outside of the Spremberg region (given its very limited training), it appears that our approach could be applied, given appropriate training information, to derive more widely applicable predictions.

While a generalised model would require significantly more training data (and rigorous evaluation), we suggest that locally trained models such as the one presented here can be very useful. The translation of ≈ metre scale petrophysical logger

260 measurements to millimetre resolution maps provides significantly more detail on the physical and mechanical structure of the sampled rock mass. This combination of large extent and high spatial resolution help to address the scaling challenges

commonly encountered when extrapolating properties measured at small scale (using ≈ 1-10 cm sized laboratory tests) to larger, reservoir-scale models (Christie, 1996).

We suggest that the information on mesoscale heterogeneity captured by these hyperspectral predictions could be used to
265 bridge the scale gap between laboratory and reservoir properties. Further tests in different settings are required to confirm this assumption. Extensive literature describes various solutions to this scaling problem. Traditional approaches, including statistical averaging, renormalisation techniques, and geostatistical methods, have been employed to derive effective medium parameters suitable for reservoir simulations (Renard and de Marsily, 1997; King, 1989). Our work introduces a novel approach that leverages hyperspectral data as an intermediary between different measurement resolutions. By integrating hyperspectral
270 data with traditional petrophysical data, we can capture fine-scale heterogeneity (at ≈ 1 mm^2 resolution; Fig. 6) while retaining coverage applicable at the reservoir scale (extent of several hundred metres). A similar approach might also be possible using outcrop hyperspectral data, if data quality issues associated with hyperspectral outcrop imaging can be resolved, further improving our ability to quantify mesoscale mechanical and petrophysical variability and so link laboratory and reservoir scale properties.

275 6   Conclusion

To conclude, our results demonstrate the link between hyperspectral and petrophysical properties. We have developed a preprocessing workflow that overcomes challenges associated with differing data resolutions, and successfully trained a deep-learning model to predict slowness, density and gamma-ray count from VNIR, SWIR, MWIR and LWIR hyperspectral data. The model was tested on an independent drill core, which included an unseen lithology, and gave accurate predictions (test R^2 score > 0.7).
280 A Shapley analysis suggests the model is leveraging the sensitivity of the MWIR-LWIR range to mineralogy, grain-size and surface roughness to inform predictions of slowness and density, indicating some degree of generalisation might be possible. The gamma predictions are much more closely linked to mineralogy, so could be quite broadly applicable.

We suggest that this novel approach enables the quantification of petrophysical properties at high spatial resolution over large areas, providing the information on mesoscale heterogeneity needed to help bridge laboratory and reservoir scale mechanical
285 and physical behaviour.

Code and data availability. The hyperspectral data presented in this study can be found here: Spremberg Data. Our code is available here: Vector-Geology Repository.

Acknowledgements. The author(s) declare that financial support was received for the research, authorship, and/or publication of this article. This research has received funding from the European Union's Horizon Europe research and innovation programme under grant agreement
290 nᵒ 101058483 (VECTOR) and from UK Research and Innovation.

Author contributions. AK: Formal analysis, Investigation, Methodology, Visualization, Writing – original draft; ST: Conceptualization, Methodology, Writing – original draft, review & editing; MK: Writing – review & editing; RG: Writing – review & editing,

Competing interests. The authors declare that they have no conflict of interest.

References

Albawi, S., Mohammed, T. A., and Al-Zawi, S.: Understanding of a convolutional neural network, pp. 1–6, IEEE, ISBN 978-1-5386-1949-0, https://doi.org/10.1109/ICEngTechnol.2017.8308186, 2017.

Asquith, G., Krygowski, D., Henderson, S., and Hurley, N.: Basic well log analysis, American Association of Petroleum Geologists, 2004.

Baker, P.: Density Logging with Gamma Rays, Transactions of the AIME, 210, 289–294, https://doi.org/10.2118/940-g, 1957.

Borg, G., Piestrzyński, A., Bachmann, G. H., Puttman, W., Walther, S., and Fiedler, M.: An overview of the European Kupferschiefer deposits, Economic Geology Special Publication, Special Publication, 455–486, 2012.

Bourbié, T., Coussy, O., and Zinszner, B.: Acoustics of Porous Media, Editions TECHNIP, 1987.

Christie, M.: Upscaling for Reservoir Simulation, Journal of Petroleum Technology, 48, 1004–1010, https://doi.org/10.2118/37324-JPT, 1996.

Clark, R. N.: Spectroscopy of Rocks and Minerals, and Principles of Spectroscopy, in: Remote Sensing for the Earth Sciences: Manual of Remote Sensing, edited by Rencz, A. N., vol. 3, pp. 3–58, John Wiley & Sons, 1999.

Géring, L., Kirsch, M., Thiele, S., Ribeiro, A. D. L., Gloaguen, R., and Gutzmer, J.: Spectral characterisation of hydrothermal alteration associated with sediment-hosted Cu–Ag mineralisation in the central European Kupferschiefer, Solid Earth, 14, 463–484, https://doi.org/10.5194/se-14-463-2023, 2023.

Kereszturi, G., Heap, M., Schaefer, L. N., Darmawan, H., Deegan, F. M., Kennedy, B., Komorowski, J.-C., Mead, S., Rosas-Carbajal, M., Ryan, A., Troll, V. R., Villeneuve, M., and Walter, T. R.: Porosity, strength, and alteration – Towards a new volcano stability assessment tool using VNIR-SWIR reflectance spectroscopy, Earth and Planetary Science Letters, 602, 117 929, https://doi.org/10.1016/j.epsl.2022.117929, 2023.

King, P. R.: The use of renormalization for calculating effective permeability, Transport in Porous Media, 4, 37–58, 1989.

Kopp, J. C., Spieth, V., and Bernhardt, H.-J.: Precious metals and selenides mineralisation in the copper-silver deposit Spremberg-Graustein, Niederlausitz, SE-Germany, Zeitschrift der Deutschen Gesellschaft für Geowissenschaften, 163, 361–384, https://doi.org/10.1127/1860-1804/2012/0163-0361, 2012.

Laukamp, C., Rodger, A., LeGras, M., Lampinen, H., Lau, I. C., Pejcic, B., Stromberg, J., Francis, N., and Ramanaidou, E.: Mineral Physicochemistry Underlying Feature-Based Extraction of Mineral Abundance and Composition from Shortwave, Mid and Thermal Infrared Reflectance Spectra, Minerals, 11, https://doi.org/10.3390/min11040347, 2021.

LeCun, Y., Bengio, Y., and Hinton, G.: Deep learning, Nature, 521, 436–444, https://doi.org/10.1038/nature14539, 2015.

Lee, J., Cook, O. J., Argüelles, A. P., and Mehmani, Y.: Imaging geomechanical properties of shales with infrared light, Fuel, 334, https://doi.org/10.1016/j.fuel.2022.126467, 2023.

Lipovetsky, S. and Conklin, M.: Analysis of regression in game theory approach, Applied Stochastic Models in Business and Industry, 17, 319–330, https://doi.org/10.1002/asmb.446, 2001.

Lu, L., Shin, Y., Su, Y., and Karniadakis, G. E.: Dying ReLU and Initialization: Theory and Numerical Examples, https://doi.org/10.4208/cicp.OA-2020-0165, 2019.

Lundberg, S. M. and Lee, S.-I.: A Unified Approach to Interpreting Model Predictions, in: Advances in Neural Information Processing Systems 30, edited by Guyon, I., Luxburg, U. V., Bengio, S., Wallach, H., Fergus, R., Vishwanathan, S., and Garnett, R., pp. 4765–4774, Curran Associates, Inc., http://papers.nips.cc/paper/7062-a-unified-approach-to-interpreting-model-predictions.pdf, 2017.

330 McInnes, L., Healy, J., and Astels, S.: hdbscan: Hierarchical density based clustering, Journal of Open Source Software, 2, 205, https://doi.org/10.21105/joss.00205, 2017.

Paszke, A., Gross, S., Massa, F., Lerer, A., Bradbury, J., Chanan, G., Killeen, T., Lin, Z., Gimelshein, N., Antiga, L., Desmaison, A., Köpf, A., Yang, E., DeVito, Z., Raison, M., Tejani, A., Chilamkurthy, S., Steiner, B., Fang, L., Bai, J., and Chintala, S.: PyTorch: An Imperative Style, High-Performance Deep Learning Library, 2019.

335 Pedregosa, F., Varoquaux, G., Gramfort, A., Michel, V., Thirion, B., Grisel, O., Blondel, M., Prettenhofer, P., Weiss, R., Dubourg, V., Vanderplas, J., Passos, A., Cournapeau, D., Brucher, M., Perrot, M., and Duchesnay, E.: Scikit-learn: Machine Learning in Python, Journal of Machine Learning Research, 12, 2825–2830, 2011.

Renard, P. and de Marsily, G.: Calculating equivalent permeability: A review, Advances in Water Resources, 20, 253–278, 1997.

Rost, E., Hecker, C., Schodlok, M. C., and Van der Meer, F. D.: Rock Sample Surface Preparation Influences Thermal Infrared Spectra,
340 Minerals, 8, 475, https://doi.org/10.3390/min8110475, 2018.

Salisbury, J. W., Hapke, B., and Eastes, J. W.: Usefulness of weak bands in midinfrared remote sensing of particulate planetary surfaces, Journal of Geophysical Research: Solid Earth, 92, 702–710, https://doi.org/10.1029/jb092ib01p00702, 1987.

Schlumberger: Log Interpretation Principles/Applications, Schlumberger Limited, 1972.

Serra, O.: Fundamentals of Well-Log Interpretation: The Acquisition of Logging Data, Elsevier, 1984.

345 Sheriff, R. E. and Geldart, L. P.: Exploration Seismology, Cambridge University Press, ISBN 9781139168359, https://doi.org/10.1017/cbo9781139168359, 1995.

Spieth, V.: Zechstein Kupferschiefer at Spremberg and related sites : hot hydrothermal origin of the polymetallic Cu-Ag-Au deposit, https://doi.org/10.18419/OPUS-10530, 2019.

Thiele, S. T., Lorenz, S., Kirsch, M., Acosta, I. C. C., Tusa, L., Herrmann, E., Möckel, R., and Gloaguen, R.: Multi-scale, multi-sensor data
350 integration for automated 3-D geological mapping, Ore Geology Reviews, 136, 104 252, https://doi.org/10.1016/j.oregeorev.2021.104252, 2021.

Thiele, S. T., Kirsch, M., Lorenz, S., Saffi, H., Alami, S. E., Acosta, I. C. C., Madriz, Y., and Gloaguen, R.: Maximising the value of hyperspectral drill core scanning through real-time processing and analysis, Frontiers in Earth Science, 12, https://doi.org/10.3389/feart.2024.1433662, 2024.

355 Virtanen, P., Gommers, R., Oliphant, T. E., Haberland, M., Reddy, T., Cournapeau, D., Burovski, E., Peterson, P., Weckesser, W., Bright, J., van der Walt, S. J., Brett, M., Wilson, J., Millman, K. J., Mayorov, N., Nelson, A. R. J., Jones, E., Kern, R., Larson, E., Carey, C. J., İlhan Polat, Feng, Y., Moore, E. W., VanderPlas, J., Laxalde, D., Perktold, J., Cimrman, R., Henriksen, I., Quintero, E. A., Harris, C. R., Archibald, A. M., Ribeiro, A. H., Pedregosa, F., van Mulbregt, P., and Contributors, S. .: SciPy 1.0: Fundamental Algorithms for Scientific Computing in Python, Nature Methods, 17, 261–272, https://doi.org/10.1038/s41592-019-0686-2, 2020.

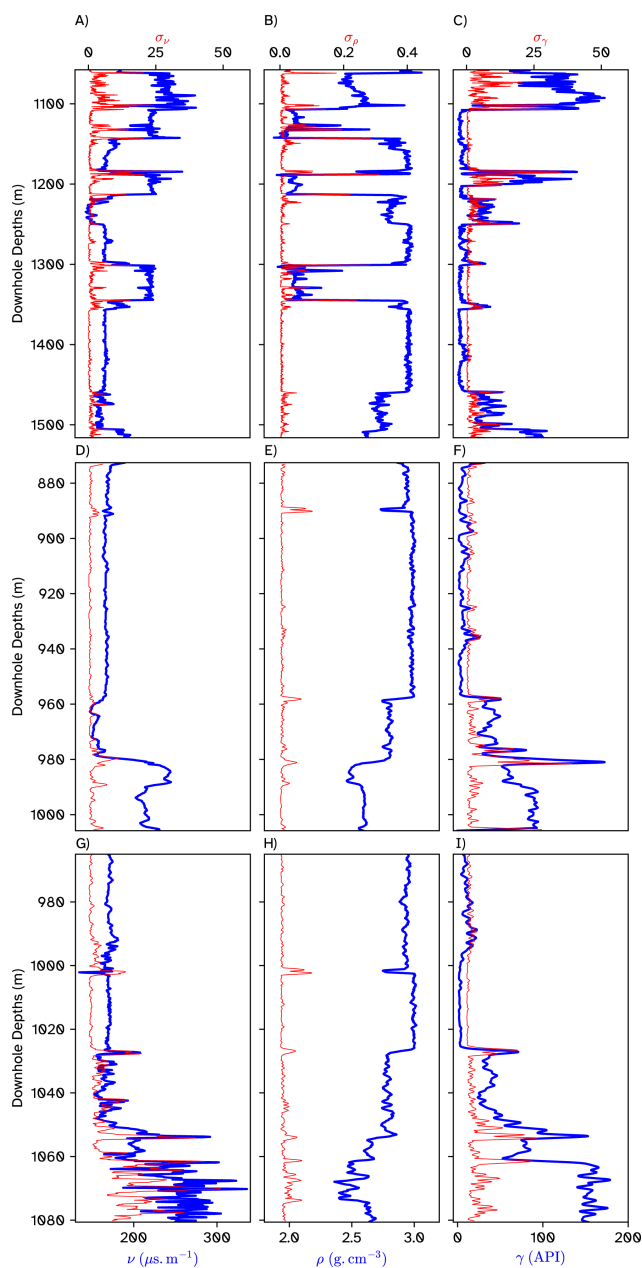
360 Zaini, N., Van der Meer, F., and Van der Werff, H.: Effect of Grain Size and Mineral Mixing on Carbonate Absorption Features in the SWIR and TIR Wavelength Regions, Remote Sensing, 4, 987–1003, https://doi.org/10.3390/rs4040987, 2012.

Figures

Figure 1. Measured logs for KSL133 (A-C), KSL131 (D-F) and KSL136 (G-I). Figures A, D and G show the sonic log, B, E and H show the gamma-gamma density log, and C, F, and I show the gamma-ray log; measured values are shown in blue along with the rolling standard deviation for a 1m window in red.

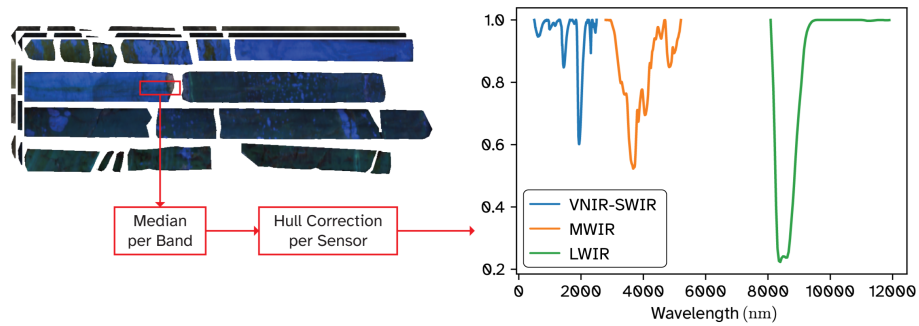



Figure 2. Preprocessing of the hyperspectral data. The images on the left show an example set of core scans, with the long-wave infrared being shown at the top as a false colour composite. The red rectangle is an example window measuring 100 pixels wide by 50 pixels tall, centred along the core. The spectra within this window are averaged (using a median), and hull-corrected to give the spectrum on the right.

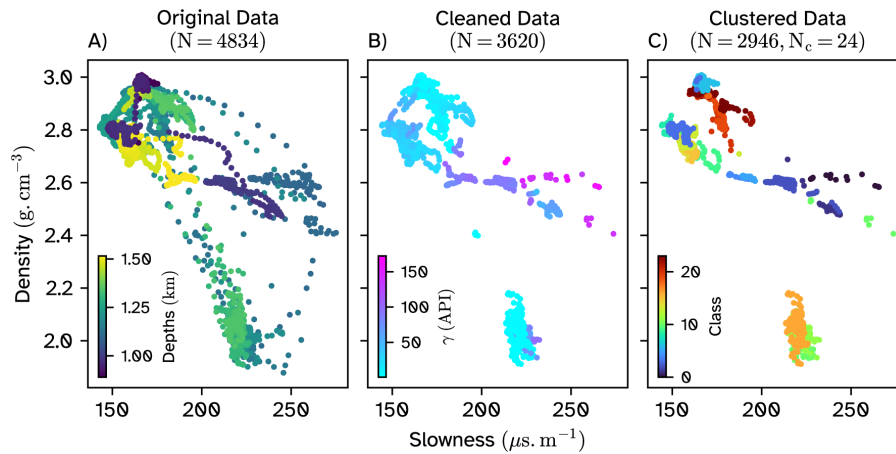



Figure 3. Preprocessing of the training dataset. Slowness-density scatter plots of A) original dataset, B) filtered dataset with high standard deviation points removed, and C) the clustered dataset with HDBSCAN class labels. N refers to the number of points in the dataset at the current stage of processing, and N_c refers to the number of clusters defined by the HDBSCAN algorithm (excluding the noisy cluster).

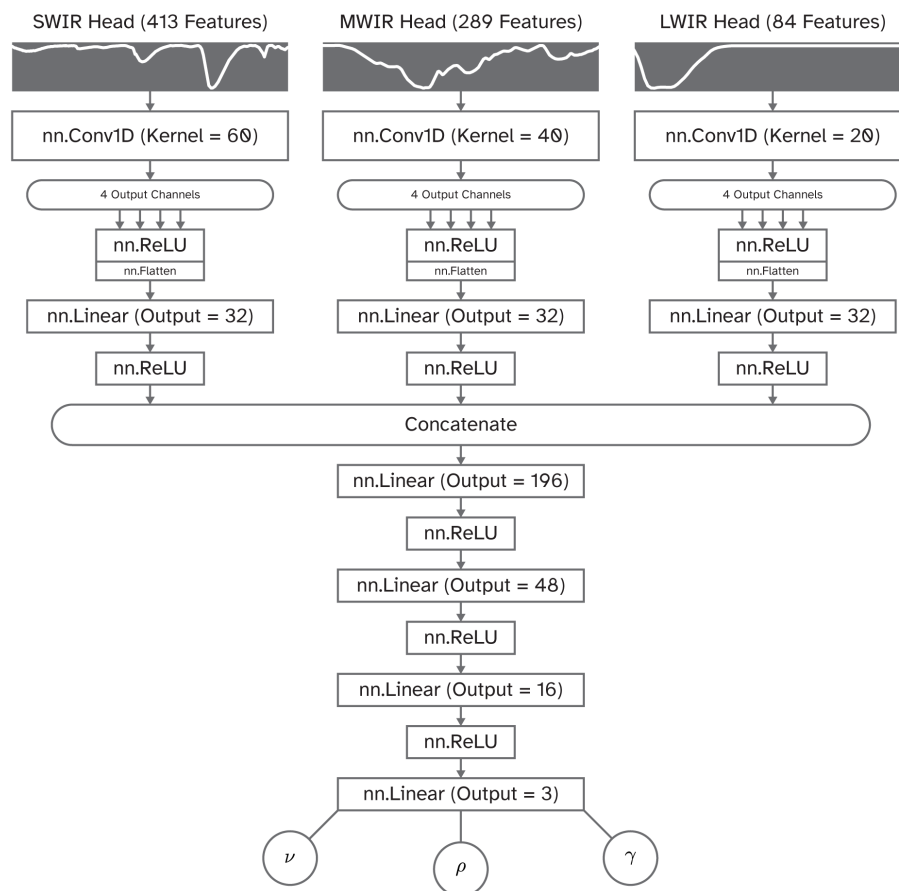



Figure 4. The architecture of the best-performing Multi-headed CNN Model. Note the difference in convolutional kernel sizes for the different heads. As the number of bands and the shape and size of the absorption features are different across different sensors, the convolutional kernel sizes of 60, 40 and 20 were chosen to better facilitate feature extraction.

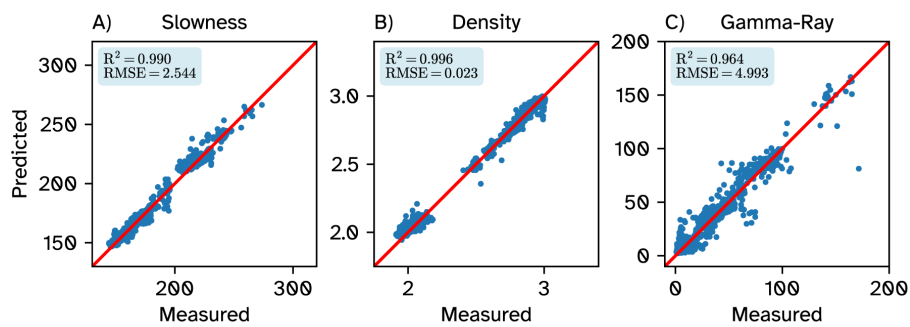



Figure 5. Scatter plots of the measured v/s. predicted properties for the training dataset. A) slowness (in μs.m^{-1}), B) density (in g.cm^{-3}) and C) gamma-ray (in API).

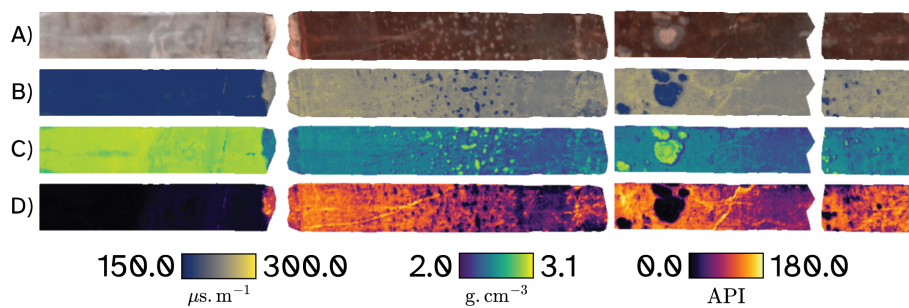



Figure 6. Pixel-scale petrophysical property predictions for an example core-scan from the KSL133 borehole. A) RGB image (derived from the RGB bands of the VNIR sensor), B) slowness, C) density and D) gamma-ray.

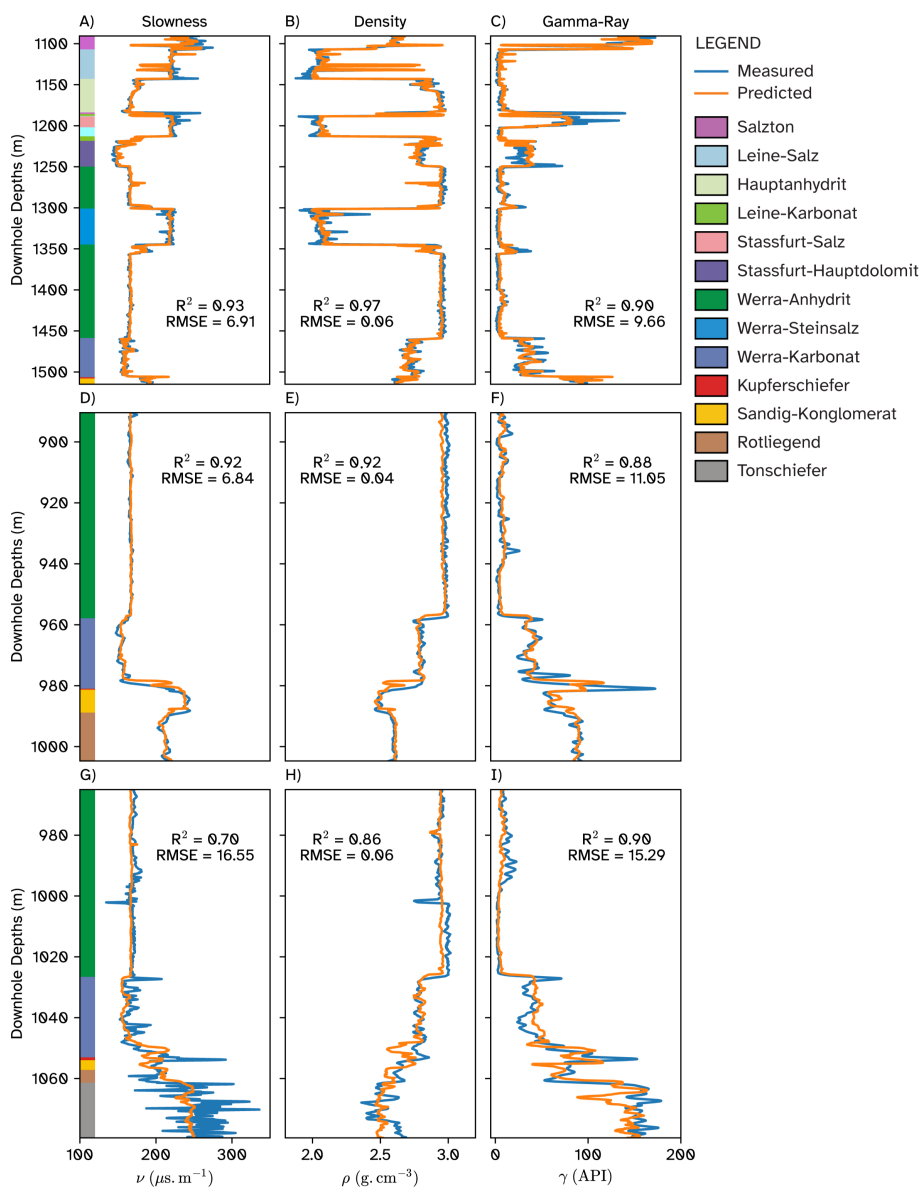



Figure 7. Reconstructed downhole logs for KSL133, KSL131 and KSL136. Figures A-C show the slowness, density and gamma-ray for the KSL133 borehole, D-F for the KSL131 borehole and G-I for the KSL136 borehole. The coloured rectangles in plots A, D, and G are the lithologies acquired from the geological logs of the three boreholes.

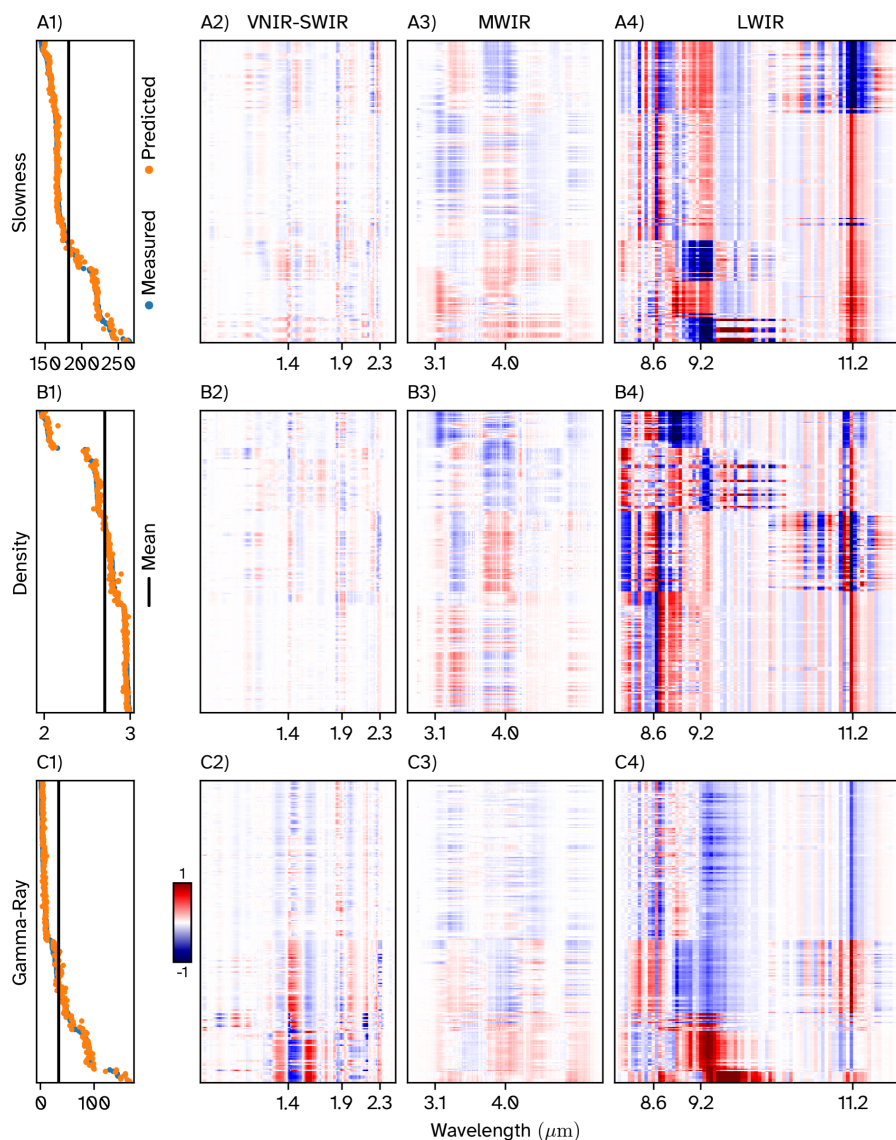



Figure 8. Shapley values for our model. Every row within a sub-figure corresponds to one input spectrum (i.e., a 'Sample'), and every column corresponds to a particular band within the input spectrum. Figures A1, B1 and C1 show the measured and predicted slowness (in μs.m^{-1}), density (in g.cm^{-3}) and gamma-ray values (in API) respectively (sorted in decreasing order of predicted value). The black vertical line corresponds to the mean predicted output. Figures A2-A4 show the Shapley values for the VNIR-SWIR bands; B2-B4 for the MWIR bands; and C2-C4 for the LWIR bands. Wavelengths at which key absorption features are expected have been annotated along the x-axis of these plots.

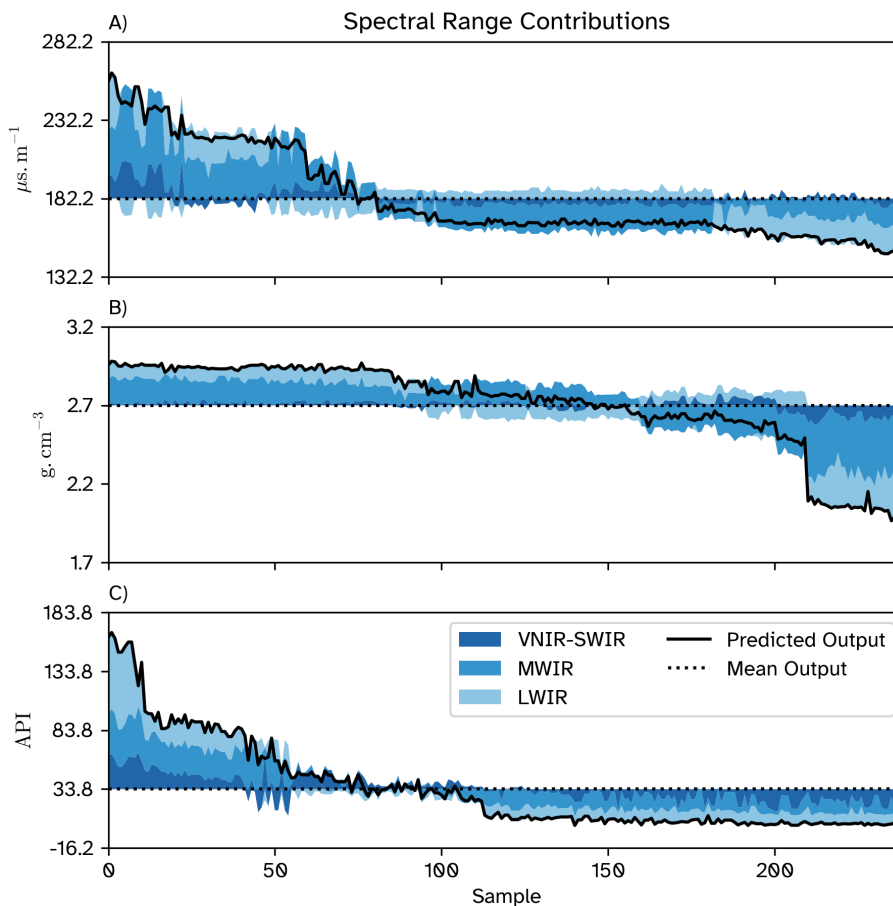



Figure 9. Contribution of each spectral range towards the output. The contribution from each spectral range is calculated by taking the signed sum of all the Shapley values corresponding to the bands in that range for all the input spectra (referred to as 'Sample'). The dotted line shows the mean predicted output, and the solid line shows the predicted output value, which is the signed sum of the mean and the individual spectral range contributions.

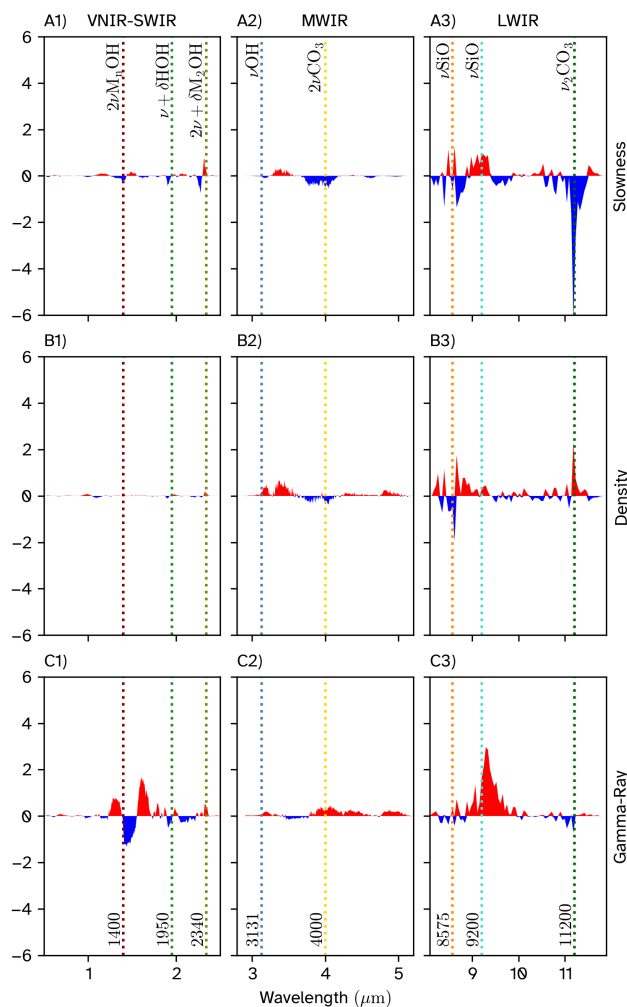



Figure 10. Median of the Shapley values for the selected samples. The columns are split according to the sensors, while the rows are split according to the properties, with Figure A showing the median Shapley value for slowness, B for the density, and C for the gamma-ray. The red regions correspond to the bands that push the output above the mean, whereas the blue regions correspond to the bands that push the output below the mean. The dotted vertical lines mark the key absorption features of interest within an IR spectrum (Laukamp et al., 2021) with the annotations specifying the wavelength (in nm) and the corresponding molecular vibrational features.

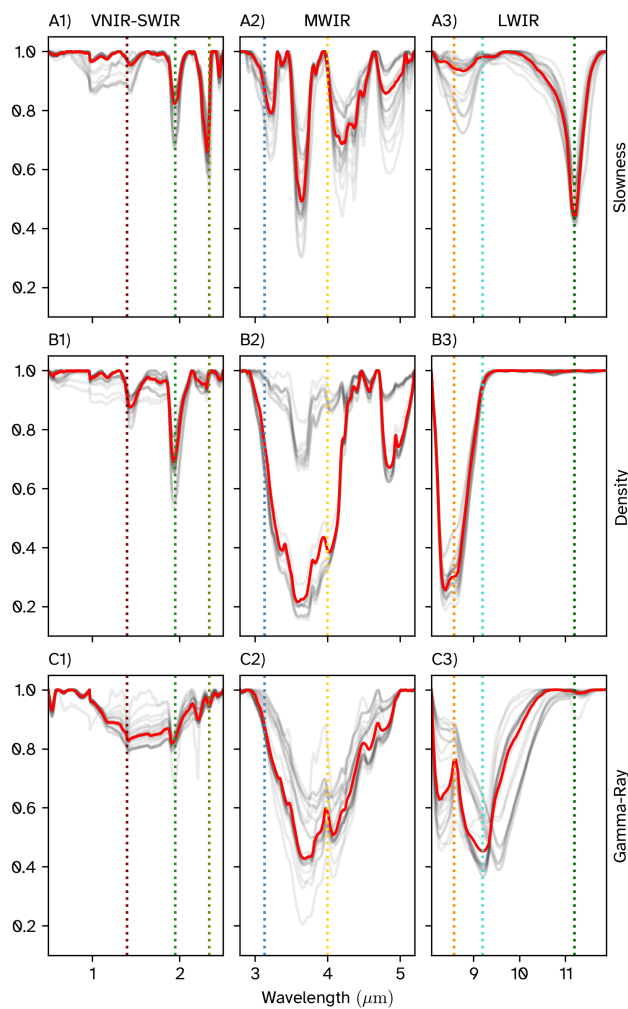



Figure 11. Reflectance spectra of selected samples. The columns are split according to the sensors, while the rows are split according to the properties, with Figure A showing the spectra of the samples in the bottom 10 percentile of slowness predictions, B showing the spectra of the samples in the top 10 percentile of density predictions, and C showing the spectra for the samples in the top 10 percentile of gamma-ray reading predictions. The low opacity grey lines plot the individual spectra, whereas the red line corresponds to the median spectrum.