Author Response - 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)

We thank the reviewers for their constructive feedback. Please find a full set of responses, including a revised manuscript and record of the corresponding changes with the changed line numbers in the attached pdf file.

Sincerely,
Akshay Kamath and authors

Reviewer Comment 1

Dear Andres Ortega Lucero & Steven Micklethwaite,

We thank you for your time and effort reviewing the submitted manuscript, and are pleased that you appreciated our results. We have incorporated your suggestions into the revised manuscript, as detailed in the following pages. Please note that to facilitate the evaluation of our revision, the page and line numbers of the reviewers' comments refer to the originally submitted manuscript while page and line numbers of our responses refer to our revised manuscript.

Kindest regards,
Akshay Kamath, Samuel Thiele, Moritz Kirsch and Richard Gloaguen

Q1) Does the paper address relevant scientific questions within the scope of SE?

Yes, it does. The exploration of new methodologies based on hyperspectral data that allow the prediction of petrophysical properties beyond the mineral-mapping traditional approach is highly relevant within the earth sciences.

We are glad the reviewer finds the work interesting, and agree completely.

Q2) Does the paper present novel concepts, ideas, tools, or data?

The paper presents a novel processing workflow to predict petrophysical properties from hyperspectral core data.

It might be worth the authors reading and possibly citing the PhD thesis from Rocio Vargas Soto, which achieves similar things using hyperspectral imaging to extract comminution properties (https://espace.library.uq.edu.au/view/UQ:c0c45d3)

The PhD thesis mentioned by the reviewer sounds very interesting but is not publicly accessible and therefore can not be cited. The author did not publish in scientific journals.

Q3) Are substantial conclusions reached?

They are substantial as part of a broader development to see if we can use hyperspectral data to extract petrophysical properties.

Agreed. Our contribution is a first-step, but hopefully a useful one.

Q4) Are the scientific methods and assumptions valid and clearly outlined?

The methodology is clearly covered, discussing the data acquisition, co-registration and processing. However:

a. In section 3.1 "Hyperspectral data acquisition", although is mentioned that the details of the dataset can be found in a previous work from “Thiele et al. (2024)” it would be worthy to mention the spectral resolution and/or number of bands in each spectral region (VNIR-SWIR, MWIR, and LWIR). It will allow to better understand the further section 3.4 “Spectral processing”, specially when it is mentioned that the first and last 10 bands of the MWIR and LWIR data were removed. Similarly, in section 3.8, reference is made to the relevance of the spectral sampling resolution for the configuration of the convolutional neural network model.

b. In section 3.3, "Data Co-registration," the spatial resolution of the slowness and hyperspectral data is mentioned. It may also be useful for the authors to include the spatial resolution of the density and gamma-ray properties to provide a clearer understanding of the downsampling method
c. The first heading in Fig 4 only references SWIR, however in the text in section 3.6 the heading in described as VNIR-SWIR. Can you please update the figure?

The reviewers are correct, and the paper could benefit from a more in-depth section on the hyperspectral dataset. In accordance, the section has been updated from L3.20 as follows:

"... with the AisaFENIX sensor capturing 450 bands in the visible-near to short-wave infrared (VNIR-SWIR; 380–2500 nm) with an average spectral sampling resolution of 3.5 nm for the VNIR and 5.5 nm for the SWIR, the SPECIM FX50 sensor capturing 308 bands in the mid-wave infrared (MWIR; 2700–5300 nm) with an average sampling resolution of 8.4 nm, and the AisaOWL sensor capturing 103 bands in the long-wave infrared (LWIR; 7700–12300 nm) with an average sampling resolution of 45 nm."

For the second point, we have added the sampling resolutions of the density and gamma-ray logs, starting from L4.23:

"... (with a sampling resolution of 2 cm), and gamma-ray (with a sampling resolution of 5 cm) logs ..."

For the third point, Figure 4 has been updated to have VNIR-SWIR as the title of the first head.

Q5) Are the results sufficient to support the interpretations and conclusions?

Yes, they are. The interpretations and conclusions are based on the model's performance in predicting each petrophysical property.

Q6) Is the description of experiments and calculations sufficiently complete and precise to allow their reproduction by fellow scientists (traceability of results)?

Yes, it’s well described, and the inclusion of the Shapley analysis provides insights into how the model is learning, which gives confidence in generalizing the model to other experiments.

Q7) Do the authors give proper credit to related work and clearly indicate their own new/original contribution?

Check

Q8) Does the title clearly reflect the contents of the paper?

Check

Q9) Does the abstract provide a concise and complete summary?

Check

Q10) Is the overall presentation well structured and clear?

It is well-structured. The paper presents a clear methodology for data acquisition, processing and analysis.

Q11) Is the language fluent and precise?

Check; Lines 254 onwards: Slightly convoluted sentence. REWRITE "This indicates that there is a sensitivity to the fundamental mechanical and petrophysical properties of the rock, and it suggests that the model could be generalised on more diverse data."

The sentence has been restructured (with minor changes to avoid repetition) at L10.22:

"The model appears to be sensitive to the fundamental mechanical and petrophysical properties of the rock, which suggests that it could be generalised on more diverse data."

Q12) Are mathematical formulae, symbols, abbreviations, and units correctly defined and used?

Check

Q13) Should any parts of the paper (text, formulae, figures, tables) be clarified, reduced, combined, or eliminated?

The paper is extremely well-written and there is no text to be eliminated, reduced or combined. The paper could benefit from adding more detail on the data used, rather than simply referencing Thiele et. Al 2024 as mentioned above.

We agree with the reviewer and have made the necessary additions (See Q4).

It may also be useful for the authors to be slightly more explicit about the challenge of generalisation in the paper. In line 250, the authors address this problem by discussing the implications of the results from the Shapley analysis. However, for readers who do not have much knowledge about CNNs it may be useful to add the following sentence at the beginning of the paragraph:

"A common challenge for deep learning models based on CNNs is whether or not they can be generalised. In this study, training and applying the model to 3 drill cores from the same geological sequence does not mean that similar results could be attained in different geological sequences. However, given the results from our Shapley value analysis, we suggest ... "

Additional clarification about the problems with generalisation has been added as suggested on L10.15:

"A common challenge for deep learning models based on CNNs is whether or not they can be generalised. In this study, training and testing the model on three drill cores from the same stratigraphic sequence does not mean that similar results could be attained in different regions. However, given..."

Q14) Are the number and quality of references appropriate?

Check

Q15) Is the amount and quality of supplementary material appropriate?

Check

Reviewer Comment 2

Dear McLean Trott,

We thank you for your time and effort reviewing the submitted manuscript, and are pleased that you appreciated our results. We have incorporated your suggestions into the revised manuscript, as detailed in the following pages. Please note that to facilitate the evaluation of our revision, the page and line numbers of the reviewers' comments refer to the originally submitted manuscript while page and line numbers of our responses refer to our revised manuscript.

Kindest regards,
Akshay Kamath, Samuel Thiele, Moritz Kirsch and Richard Gloaguen

Q1) 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.

We agree with the reviewer that direct prediction of lithology (or mineralogy) makes sense in many situations, however this is not our aim in this case. Specifically, as stated in the introduction:

"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 predict important petrophysical and mechanical properties across large drill core libraries and hyperspectral scans of outcrops"

Our aim here is to better understand the links between hyperspectral response and petrophysical properties, as a first step towards predicting more difficult to measure parameters (e.g., porosity, strength, stiffness, etc.). Hence, the derivation of a lithology log is not our end goal. Additionally, we note that our approach effectively super-resolves the borehole petrophysics data (from ~0.1 m to 0.001 m), facilitating more detailed analyses of these widely used data and, interestingly, in some ways mitigating the co-registration uncertainties addressed in the following comments.

This has been clarified by adding the following in the conclusion at L11.16, and L11.26:

"We effectively super-resolve the borehole petrophysics data (i.e. from ~0.1 m resolution of the logger to 0.001 m resolution of the hyperspectral cameras), which helps us explore the intricacies and variations of these properties that cannot be captured by running log measurements. Additionally, our workflow mitigates the co-registration uncertainties that prevent carrying out machine learning workflows over drill core data."

"Our work also serves as a stepping stone toward predicting secondary properties such as porosity and permeability, helping build a link between hyperspectral data and widely used physical rock property measurements."

We thank the reviewer for their thought provoking question!

Q2) 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.

We completely agree that depth estimates will differ between the core boxes (and hence HSI data) and downhole petrophysics logs. This is why we removed the areas (petrophysical edges, defined using the rolling standard deviation) that will be most sensitive to these co-registration uncertainties. The second paragraph of section 3.3 (L5.2) has been clarified as follows to address this:

"Due to factors like core-loss, co-registration errors are expected between the petrophysical logs and the drill-core boxes. To ensure our training dataset does not contain spectra paired with incorrect petrophysical properties, we use the rolling standard deviation to eliminate points from regions of high property variance, as these will be highly sensitive to co-registration uncertainties. Hence the underlying assumption within our preprocessing steps is that by picking points only from petrophysically homogeneous regions of the drill cores, we can partially mitigate challenges caused by co-registration errors."

Q3) 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.

We agree that the HDBSCAN algorithm is sensitive to hyper-parameter selection; so this statement has been removed (L5.31). In our case we did not optimize these hyperparameters as we're aiming to over-segment our dataset (i.e. classify into more classes than exist in the data), for the sole purpose of performing stratified random sampling. Hence, unlike classification tasks where the output classes have more meaning, our approach is less sensitive to hyperparameter selection. Instead, we tweaked the hyperparameters until we were satisfied that we had over-segmented the dataset, based on visual inspection of the clustering results (Fig. 3C). We have clarified this in the Data-Balancing section, at L6.3:

"Hyperparameters of HDBSCAN were selected manually, and iteratively assessed based on the number and separation between the clusters shown in Fig. 3. We aimed to over-segment the dataset, as the clustering was solely used to help in removing the inherent bias from a larger number of points belonging to one lithology. Any cluster distribution with over-segmented clusters (i.e., the large clusters separated from the rest) would provide similar results during training."

As for the second point, we used the HDBSCAN clusters to do the stratified sampling; this has been clarified in the document now at L6.8:

"…using the labels from the HDBSCAN clusters…"

Reviewer Comment 3

Dear Reviewer,

We thank you for your time and effort reviewing the submitted manuscript, and are pleased that you appreciated our results. We have incorporated your suggestions into the revised manuscript, as detailed in the following pages. Please note that to facilitate the evaluation of our revision, the page and line numbers of the reviewers' comments refer to the originally submitted manuscript while page and line numbers of our responses refer to our revised manuscript.

Kindest regards,
Akshay Kamath, Samuel Thiele, Moritz Kirsch and Richard Gloaguen

Q1) Does the paper address relevant scientific questions within the scope of SE?

Yes, the research aligns well with the scope of SE, especially in the areas of geophysics, spectroscopy and applications of machine learning techniques in geosciences.

Q2) Does the paper present novel concepts, ideas, tools, or data?

Yes. The idea of using deep learning to predict the petrophysical properties from high resolution hyperspectral data can be considered as a novel approach in upscaling workflows. Also, application of the Shapely method for analysing the impact of different spectral bands on the model's predictions is an innovative addition.

Q3) Are substantial conclusions reached?

Largely, yes. Conclusions are well supported by the key results, however some additional work on uncertainty quantification would be very beneficial.

Could the authors quantify the uncertainty of predictions using techniques such as Monte Carlo dropout, Bayesian inference or confidence intervals?

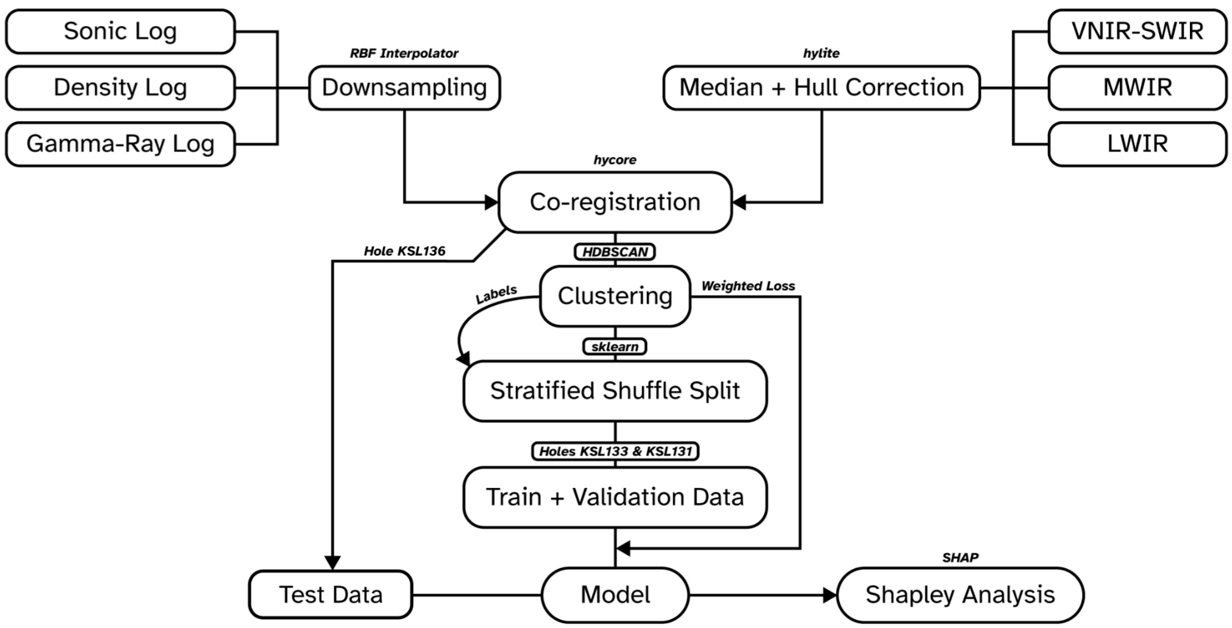
We agree that uncertainty predictions would be an interesting next step (e.g., using an ensemble model), but we consider this to be beyond the scope of the current work and suggest that this would be better addressed as a follow-up publication rather than adding extra complexity to our current manuscript. While feasible, this would require to develop an entire framework and would not add value to this submission. We argue that the XAI analysis demonstrates the robustness of the methods and that a qualitative assessment is sufficient at this stage.

Q4) Are the scientific methods and assumptions valid and clearly outlined?

Authors have outlined each step of data processing in detail; however, inclusion of a workflow flowchart would greatly enhance clarity. A visual representation of the step-by-step process, including data acquisition, data preprocessing, clustering, deep learning, evaluation, Shapely analysis, and final outputs would be very helpful for readers.

We agree with the reviewer in that a flowchart would be helpful. Due to the high number of figures already present in the manuscript, we have added the following flowchart as a supplementary figure, as referred to on L3.14:

"Hyperspectral data were acquired and coregistered with downhole petrophysical logging data, and then used to train machine learning regression models. The various steps needed to preprocess our training data and build the deep-learning models (see Supplementary Figure for an overview) are described in detail below."



Q5) Are the results sufficient to support the interpretations and conclusions?

The performance metrics used (R^2 and RMSE) support the claim that hyperspectral data can predict the petrophysical properties accurately.

Nevertheless, the model struggles with unseen lithologies, more cross validation on other datasets can help generalise the conclusions.

We agree with the reviewer on the conditions for generalisability, and have highlighted them in the Discussion, at L10.15 (after including the additions suggested by RC1):

"A common challenge for deep learning models based on CNNs is whether or not they can be generalised. In this study, training and applying the model to three drill cores from the same geological sequence does not mean that similar results could be attained in different geological sequences. However, 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. The model appears to be sensitive to the fundamental mechanical and petrophysical properties of the rock, which suggests that it could be generalised on more diverse data."

We also agree that testing on unseen lithologies would be helpful, which is why the high R^2 scores in density and gamma-ray predictions for the argillaceous basement (which has not been seen by the model) has been highlighted in L8.13:

"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 not sampled by the other training drill cores (highlighted by the grey box labelled Tonschiefer i.e., argillaceous basement). 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. The measured sonic log here shows significant fluctuations, whereas our model prediction remains steady (suggesting the lithology is spectrally quite uniform)."

As for the reviewer's comment on external validation, we agree that validation on different datasets would make the model more robust. However, we do not (currently) have access to additional datasets with hyperspectral and petrophysical data. This could thus represent an interesting follow-up work.

Q6) Is the description of experiments and calculations sufficiently complete and precise to allow their reproduction by fellow scientists (traceability of results)?

Mostly, yes. As mentioned before, all processing steps have been well documented. However, model hyperparameters could be better detailed for traceability of the results.

We have now included the jupyter notebook used to train our models, as now mentioned in the *Data and Code Availability* section (can be found here). These document the model hyperparameters we have used and ensure reproducibility.

Q7) Do the authors give proper credit to related work and clearly indicate their own new/original contribution?

Yes.

Q8) Does the title clearly reflect the contents of the paper?

Yes.

Q9) Does the abstract provide a concise and complete summary?

Partially. Authors need to clearly state the problem and motivation upfront. They should start the abstract by emphasising why upscaling petrophysical measurements using hyperspectral data is important. This would draw the attention of the readers. Since automating petrophysical property predictions could lead to major cost and efficiency benefits, this should be explicitly stated in the abstract.

We agree with the reviewer and have highlighted the possibility of speeding up petrophysical data acquisition by changing the abstract as follows:

"Hyperspectral data provides rich ~~quantitative~~ information on both the mineralogical and fine-scale textural properties of rocks which also ~~, in turn, largely~~ control their petrophysical characteristics. We propose that some physical rock properties can be predicted directly from hyperspectral data, improving petrophysical characterisation and reducing the need for often laborious measurements. In this contribution we explore correlations between hyperspectral and petrophysical data using a deep convolutional neural network. ~~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 predict ~~infer~~ slowness, density, and gamma-ray values using training and testing data from ~~. We demonstrated the performance of this approach on data acquired in the~~ Spremberg ~~region of,~~ Germany. Our results show ~~demonstrate~~ that, with careful ~~meticulous~~ pre-processing ~~steps~~ and thorough data cleaning, ~~one can overcome the~~ differences in capturing ~~resolution~~ can be overcome to ~~and~~ learn the relationship between hyperspectral data and petrophysics. Using a test dataset from a spatially independent borehole, we generated ~~generate~~ a pixel-resolution (≈ 1 mm^2) model of the petrophysical properties and resampled ~~resample~~ it to match the measured logs. This test indicated ~~indicates~~ substantial accuracy, with R^2 scores and root-mean-squared errors (RMSE) of 0.7 and 16.55 $\mu s.m^{-1}$, 0.86 and 0.06 $g.cm^{-3}$ and 0.90 and 15.29 API for the slowness, density and gamma-ray predictions readings respectively. We also analysed the Shapley values of our model to gain deeper insights into its predictions. These ~~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."

Q10) Is the overall presentation well structured and clear?

Yes.

Q11) Is the language fluent and precise?

Yes.

Q12) Are mathematical formulae, symbols, abbreviations, and units correctly defined and used?

Yes.

Q13) Should any parts of the paper (text, formulae, figures, tables) be clarified, reduced, combined, or eliminated?

The legend of Figure 7 contains German geological terms for lithologies (e.g., Salzton, Werra Anhydrit, Kupferschiefer). Authors should translate such terms into English (e.g., Salt Clay, Werra Anhydrite, Copper Shale).

Other figures are well presented.

We have translated lithological descriptions to English (e.g., Karbonat to Carbonate), as suggested by the reviewer in Figure 7. However, we have refrained from changing commonly used stratigraphic terms (e.g., Kupferschiefer). We have, however, added the translation for Kupferschiefer in Section 2, at L2.27:

> "...("Copper-Shale")..."

The corresponding occurrences for the other german terms within the text (entirely within section 4) have been changed to match the legend.

Q14) Are the number and quality of references appropriate?

Yes.

Q15) Is the amount and quality of supplementary material appropriate?

Yes.