Proceedings World Geothermal Congress 2026
Calgary, Canada, June 8 – 11, 2026

Neural Network-based Geothermal Reservoir Porosity Estimation Using Spectrally Decomposed Multiphysics Data

F. Allo*, Formento J.L. and Nustes Andrade J.
Viridien
fabien.allo@viridiengroup.com

Keywords: Machine Learning, Neural Network, Multiphysics, Porosity, Spectral Decomposition.

ABSTRACT

Finding naturally occurring porous and permeable reservoir rocks that allow sufficient fluid flow for economical geothermal applications relies on the use of geophysical measurements sensitive to the rock composition such as seismic amplitudes, seismic velocities, electromagnetics-derived resistivity or gravity-derived density. Tight budgets usually means that the quality and coverage of these data is suboptimal in addition to intrinsic resolution limitations for each technique. To partly overcome these limitations, joint inversions have been proposed to combine their strengths while alleviating their individual weaknesses. Although theoretically appealing, access and ease of use have limited their practical use so far. But recent advances in open-access user-friendly machine learning algorithms have been a game-changer. Building on existing work, we demonstrate how a series of U-net-based neural networks can be used to estimate porosity from multiphysics data. The innovative aspect of this work relies on the treatment of the input and target data provided to the networks. As the mentioned geophysical measurements have different frequency bandwidths, each input data is first spectrally decomposed in a series of frequency bands. Networks with frequency-dependent parameters are then used to estimate the porosity in each band based on the data available in each frequency range. The estimates from all networks are then recombined in the frequency domain to produce a porosity estimate in the spatial domain. The proposed workflow is first tested on a synthetic dataset to evaluate its predictive capability and then on field data where the estimated porosity is compared to wellbore-derived porosity.

1 INTRODUCTION

Geophysical measurements of the subsurface are a key source of information used to locate deep porous reservoirs that would allow sufficient fluid flow for economical geothermal applications. Surface seismic data is the most commonly used measurement for reservoir characterization in the oil and gas industry but the cost to acquire and process new state-of-the-art data is usually prohibitive in the geothermal industry. Instead, operators need to rely on existing legacy seismic data of suboptimal quality and limited coverage. Unfortunately, this type of data always lacks low and high-frequency information due to the nature of the seismic signal used in older seismic acquisition campaigns. Such bandlimited data therefore needs to be supplemented by other sources of information to cover a bandwidth broad enough to provide an accurate estimate of the subsurface porosity. A common way of building a low-frequency model in the oil and gas industry is through extrapolation of wellbore measurements following geological formations interpreted on seismic data. But this approach is not an option in geothermal exploration due to very limited well control. Instead, other geophysical measurements such as seismic velocities, electromagnetics-derived resistivity or gravity-derived density have been considered to fill the low-frequency gap and provide a more reliable estimate of porosity trends in the subsurface. Initial attempts to do so relied on the use of joint inversions to combine the strengths of the different geophysical measurements while alleviating their individual weaknesses (Tseng and Lee, 2001, Gallardo and Meju, 2004, Hoversten et al., 2006, Gao et al., 2012). But access and ease of use have limited the practical use of such joint inversions so far. Advances in open-access user-friendly machine learning algorithms are changing the picture though with existing adaptable deep neural networks being used to perform such joint inversions (Ma et al., 2024, Ren et al., 2024). Recent work by Guo et al. (2025) has shown the benefit of training a neural network with a U-net architecture on multiple geophysical data instead of performing independent estimations with each data. A study by Jo et al. (2022) has also highlighted how training a network on spectrally decomposed input data can be beneficial. The spectral decomposition tends to simplify the link between input and target data and therefore help the network extract target features at various spatial scales. Based on these two pieces of work, we propose to use an array of networks to estimate porosity from multiphysics data. Each network in the array is dedicated to a specific frequency band and trained with the geophysical measurements that contain data in that frequency range. This allows for an optimized frequency-dependent parameterization of the networks and avoids the inevitable parameterization trade-off when a single network is used to recover target features at multiple spatial scales. The predictive capability of the array of networks is first compared to that of various single networks on

a synthetic dataset. The proposed approach is then applied to field data and the estimated porosity compared to wellbore-derived porosity.

2 METHODOLOGY

Due to the limited number of wells available in geothermal projects, a large number of synthetic pseudowells are first generated following the methodology described by Allo et al. (2021). These pseudowells are used as training data for the neural networks. All tested networks have the same U-net architecture composed of 4 convolutional layers with an increasing number of filters (4, 8, 16 and 32 respectively) to progressively compress the extracted features. The only two variables in the parameterization of the networks are the image and convolutional kernel sizes. For all networks, 75% of the pseudowells are used for training and 25% are used for validation.

Table 1 below summarizes the four cases considered in this study. In case 1, a relative porosity is estimated from full-stack seismic amplitude. The relative porosity is obtained by subtracting the true low-frequency model to the target absolute porosity. The true low-frequency model is then added back to the network output to obtain the absolute porosity estimate. This is the benchmark case against which the other cases are compared.

In case 2, an absolute porosity is directly estimated from full-stack seismic amplitude. It is designed to show the effect of using band-limited seismic data as single input to derive absolute porosity and specifically to assess whether the low-frequency porosity model is recoverable.

In case 3, an absolute porosity is estimated from full-stack seismic amplitude, seismic velocity and resistivity. This case is compared to case 2 to show the potential contribution low-frequency input data such as seismic velocity and resistivity can have toward recovering the low-frequency porosity model.

In case 4, an absolute porosity is also estimated from full-stack seismic amplitude, seismic velocity and resistivity. But the input and output data are spectrally decomposed, and an array of neural networks is used to estimate porosity. The spectral decomposition has two main advantages. First, it ensures that the input and target data are defined over the same frequency bandwidth. It usually simplifies the link between them and results in a faster training process and a more accurate neural network output. Secondly, it allows tailoring the neural network parameterization for each frequency bandwidth. Large image and convolutional kernel sizes are used to extract porosity trends from low-frequency data. Both sizes are progressively decreased to extract more detailed information from higher-frequency data.

Table 1. Summary of investigated cases and their objectives.

<table>
<tr><th>Case</th><th>Inputs</th><th colspan="2">Targets</th><th>Objective</th></tr>
<tr><td rowspan="2">1</td><td>Seismic amplitude</td><td>Relative porosity</td><td rowspan="2">Absolute porosity</td><td rowspan="2">Establish best-case estimation (benchmark)</td></tr>
<tr><td colspan="2">True low-frequency porosity model</td></tr>
<tr><td>2</td><td>Seismic amplitude</td><td colspan="2">Absolute porosity</td><td>Show limitation of use of seismic amplitude alone</td></tr>
<tr><td rowspan="3">3</td><td>Seismic amplitude</td><td colspan="2" rowspan="3">Absolute porosity</td><td rowspan="3">Evaluate added value brought by seismic velocity and resistivity</td></tr>
<tr><td>Seismic velocity</td></tr>
<tr><td>Resistivity</td></tr>
<tr><td rowspan="3">4</td><td>Spectrally decomposed seismic amplitude</td><td rowspan="3">Spectrally decomposed porosity</td><td rowspan="3">Absolute porosity</td><td rowspan="3">Evaluate added value of spectral decomposition and use of independent network per frequency band</td></tr>
<tr><td>Spectrally decomposed seismic velocity</td></tr>
<tr><td>Spectrally decomposed resistivity</td></tr>
</table>

2.1 Synthetic dataset

A synthetic dataset has been used to quantitatively assess the predictive ability of the proposed methodology. The dataset is made of a single line extracted from a real dataset from the Paris Basin, France (Figure 1). The

porosity model is the result of a previous neural network estimation presented in Allo et al. (2021). It is considered the true porosity and the target to be recovered. Calibrated rock physics models for carbonates (in the reservoir interval) and marls (outside the reservoir interval) were applied to compute elastic attributes (P-wave and S-wave velocities, density). A full-stack seismic amplitude section was then derived from the elastic attributes using Zoeppritz's equation for reflectivity computation followed by a convolution with a 1D wavelet statistically extracted from the seismic data. A seismic velocity section was produced by applying a low-pass frequency filter (to remove frequencies above 30Hz) followed by a lateral smoothing filter (based on 30 consecutive traces) to the modeled P-wave velocity section. Similarly, a resistivity section was produced by applying the same filters to the resistivity modeled with Archie's law.

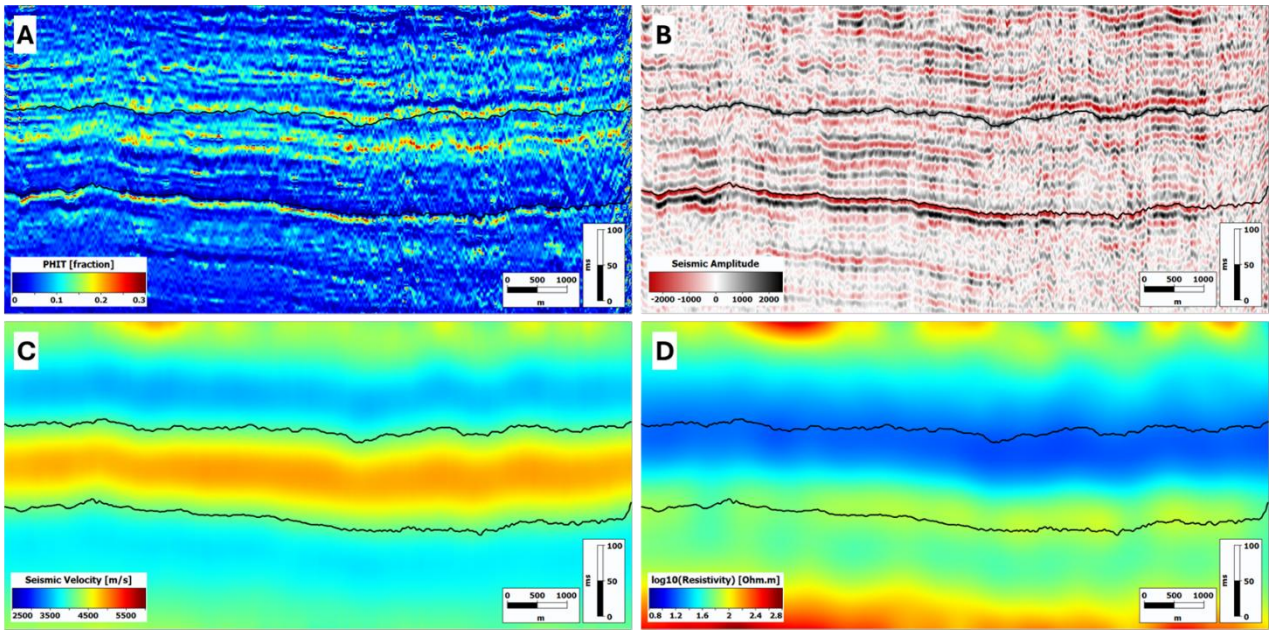


Figure 1. Synthetic model used to assess the predictive ability of the proposed methodology. The data includes a target porosity model (A), full-stack seismic amplitude (B), seismic velocity (C) and resistivity (D). Black lines indicate the interpreted positions of the top and bottom of the reservoir.

2.2 Porosity estimates

Figure 2 shows the porosity estimates for the different cases along with their associated mean error values. Case 1 (Figure 2A-2B) produces a porosity section very similar to the target one with all high-porosity layers successfully recovered along with the correct low-frequency trend due to the use of the true low-frequency model. The mean error of 4.22% is mainly caused by a mismatch of very fine details that are not recoverable from the bandlimited seismic data.

Case 2 (Figure 2C-2D) illustrates the drop in predictive ability incurred when attempting to derive the low-frequency trend from the bandlimited seismic data. While some of the high-porosity layers are still successfully recovered, the low-frequency trend is clearly missing which explains the larger mean error of 10.82%. This result was expected as the seismic data does not contain information below 15Hz so the neural network struggles to find a way to recover the low-frequency porosity trend.

Unlike the seismic data, the seismic velocity and resistivity sections contain lower frequency information so case 3 (Figure 2E-2F) was designed to quantify the uplift in predictive ability when the neural network is provided all three geophysical inputs. In this case, part of the low-frequency trend is recovered while some high-porosity layers are still preserved. This is most obvious in the reservoir interval where the decrease of porosity with depth is properly recovered. The resulting mean error of 8.67% shows that a network provided with the three geophysical inputs outperforms a network only trained on seismic data. The mean error is still twice the error of case 1 though. This can partly be explained by the use of a single network to extract a link between porosity and three input data with different resolutions and frequency contents. In this case, a trade-off in the network parameterization (image size and convolutional kernel size) is necessary to extract both low-frequency and high-frequency information from the various input data.

We propose using an array of networks to overcome this limitation. In case 4 (Figure 2G-2H), the three geophysical input data and the target porosity are spectrally decomposed in 13 frequency bands by applying a

Fourier transform. A separate neural network is used for each frequency band to find a relationship between the porosity data and the three input data in this specific frequency range. The various outputs are then recombined into a full spectrum porosity data by applying an inverse Fourier transform. The porosity section obtained with this approach looks qualitatively very close to the one obtained in case 1. This is quantitatively verified with a mean error of 5.52%, only 1.3% higher than what is obtained with the true low-frequency model. The spectral decomposition of the data and the use of an array of networks appears to be an effective way to recover the most information from the input data. As each network only deals with a given frequency band, its parameterization can be tailored to extract information at that specific spatial resolution. In practice, large image and convolutional kernel sizes are used for low-frequency data and these sizes are progressively decreased to recover more detailed features from higher-frequency data.

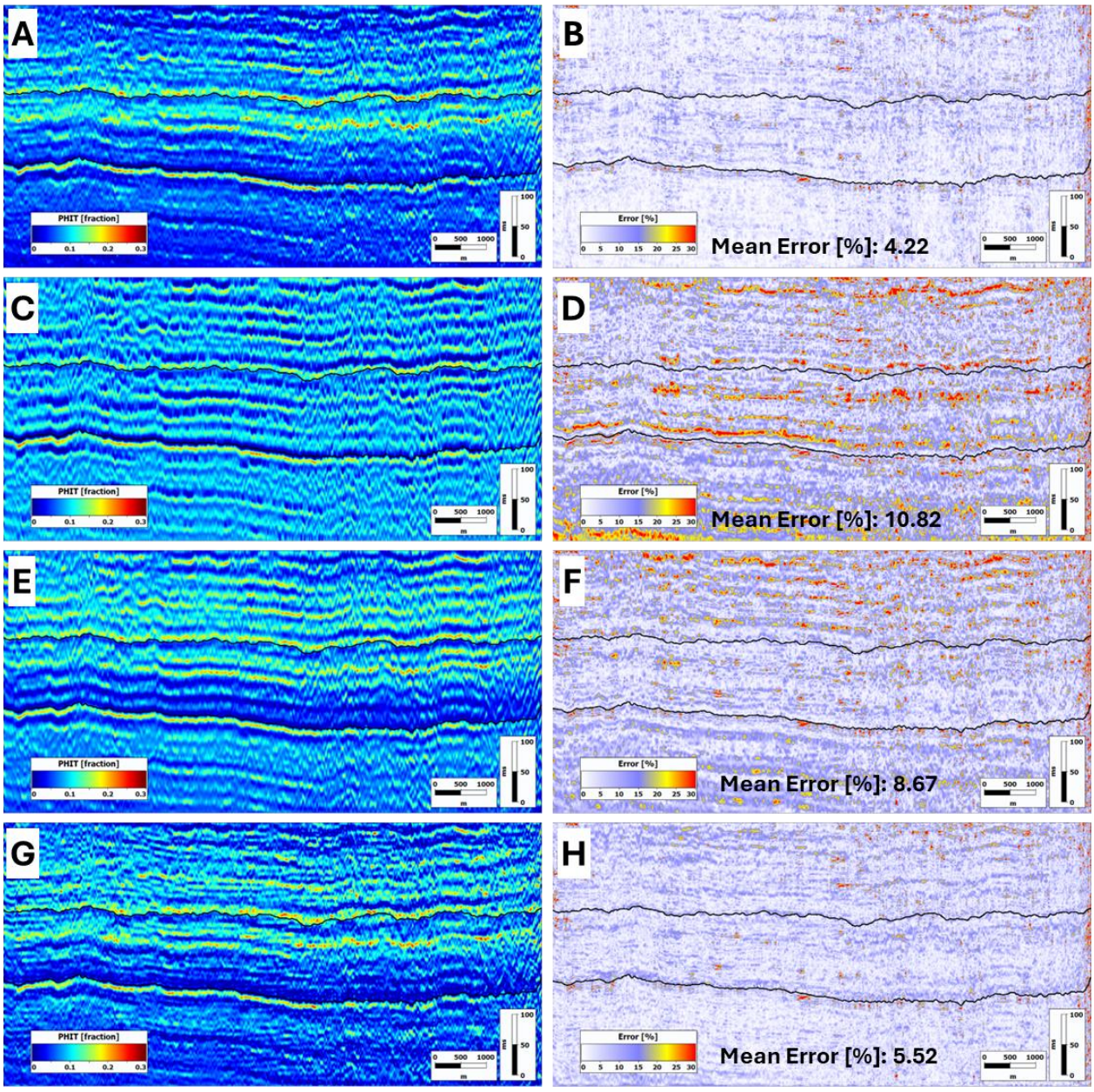


Figure 2. Porosity estimates (A, C, E, G) and associated estimation errors (B, D, F, H) for the four investigated cases. Black lines indicate the interpreted positions of the top and bottom of the reservoir.

Figure 3 further illustrates the result of the spectral approach (case 4) at a single trace location in the middle of the section. It is a simple and effective way to control the quality of the estimation and assess the contribution of each frequency band toward the overall prediction error. In this case, the quality of the estimation is similar in the first 12 frequency bands although the individual contribution of each input data varies across frequency bands with the seismic velocity and resistivity compensating for the lack of seismic data in the first two bands. The estimation in the last frequency band is clearly failing. This was expected as none of the input data contains information beyond 120Hz. The small variations present in the target porosity in that high-frequency band are

therefore not recoverable and account for the majority of the prediction error as was observed in case 1. The rest of the difference mainly originates from small errors in the first two frequency bands that highlight limitations on what can be achieved with the available geophysical data.

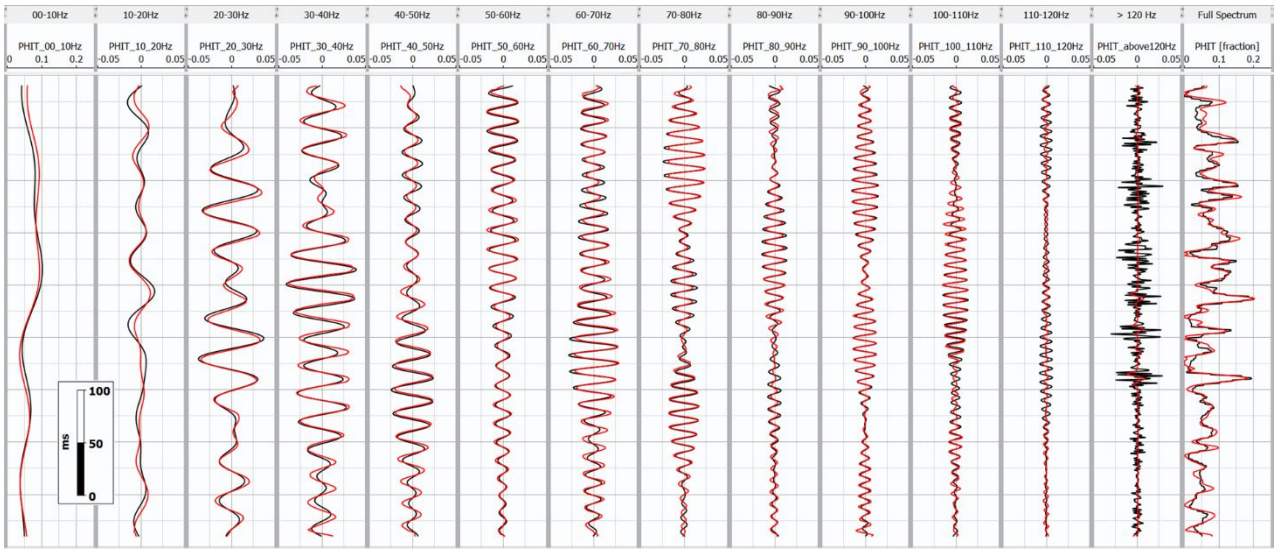


Figure 3. Estimates from an array of networks trained on spectrally decomposed data (red curves) compared to the spectrally decomposed target values (black curves). The last track shows the porosity curves obtained after recombination of all frequency bands.

3 APPLICATION TO FIELD DATA

To further validate the methodology based on an array of networks, it has been applied to field data from the Upper Rhine Graben in the northeast of France. The geothermal reservoir is made of sandstone overlying a granitic basement. As in the synthetic example, the data includes full-stack seismic amplitude, seismic velocity and resistivity (Figure 4). The resistivity is the result of an inversion of magnetotelluric and controlled source electromagnetic data constrained by the top of the granitic basement and the deep resistivity logs measured at three nearby well locations. Measurements of density, compressional and shear sonic, deep resistivity are available at these well locations. Porosity derived from density, sonic and resistivity logs are used to evaluate the quality of the networks estimates. As these wells did not contain sufficient data to train the array of networks, they were used as a base for the simulation of 162 pseudowells. Simulations included variations of reservoir thickness, reservoir porosity and reservoir clay volume. All data have been spectrally decomposed in 8 frequency bands with a width of one octave (0-2Hz, 2-4Hz, 4-8Hz, 8-16Hz, 16-32Hz, 32-64Hz, 64-128Hz and 128-256Hz). The decomposition in octaves puts more emphasis on the low-frequency components and therefore maximizes the amount of information extracted from the low-frequency seismic velocity and resistivity input data. In practice, it also simplifies the recombination of the output as the inverse Fourier transform in this case is the same as the addition of the output of all frequency bands.

Figure 5 shows the porosity resulting from the estimation with an array of networks based on spectrally decomposed data. A complex porosity distribution is visible in the reservoir interval with relatively discontinuous high-porosity layers. In contrast, low porosity values are predicted in the tight carbonate formation above the reservoir and in the granitic basement below the reservoir.

This porosity estimate was compared to porosity logs available at the three well locations. A mean error of 12.7% is observed over the logged intervals and the mean error is 12.6% within the reservoir interval. Figure 6 shows that there is a large amount of variability from one well location to the other though. This variability is due to a combination of factors that include variations in input data quality along the composite line and variable distances between the line and the well locations.

Like any other neural network-based approach, the quality of the predicted target is directly linked to the quality of the input data. In this example, an unconstrained resistivity section was also available and, when used as input instead of the constrained resistivity section, it led to a less accurate result at the well locations (mean error of 17.9%). This shows how, beyond the choice of network architecture and parameterization, acquisition and processing of the input data is critical in getting as accurate a target estimate as possible.

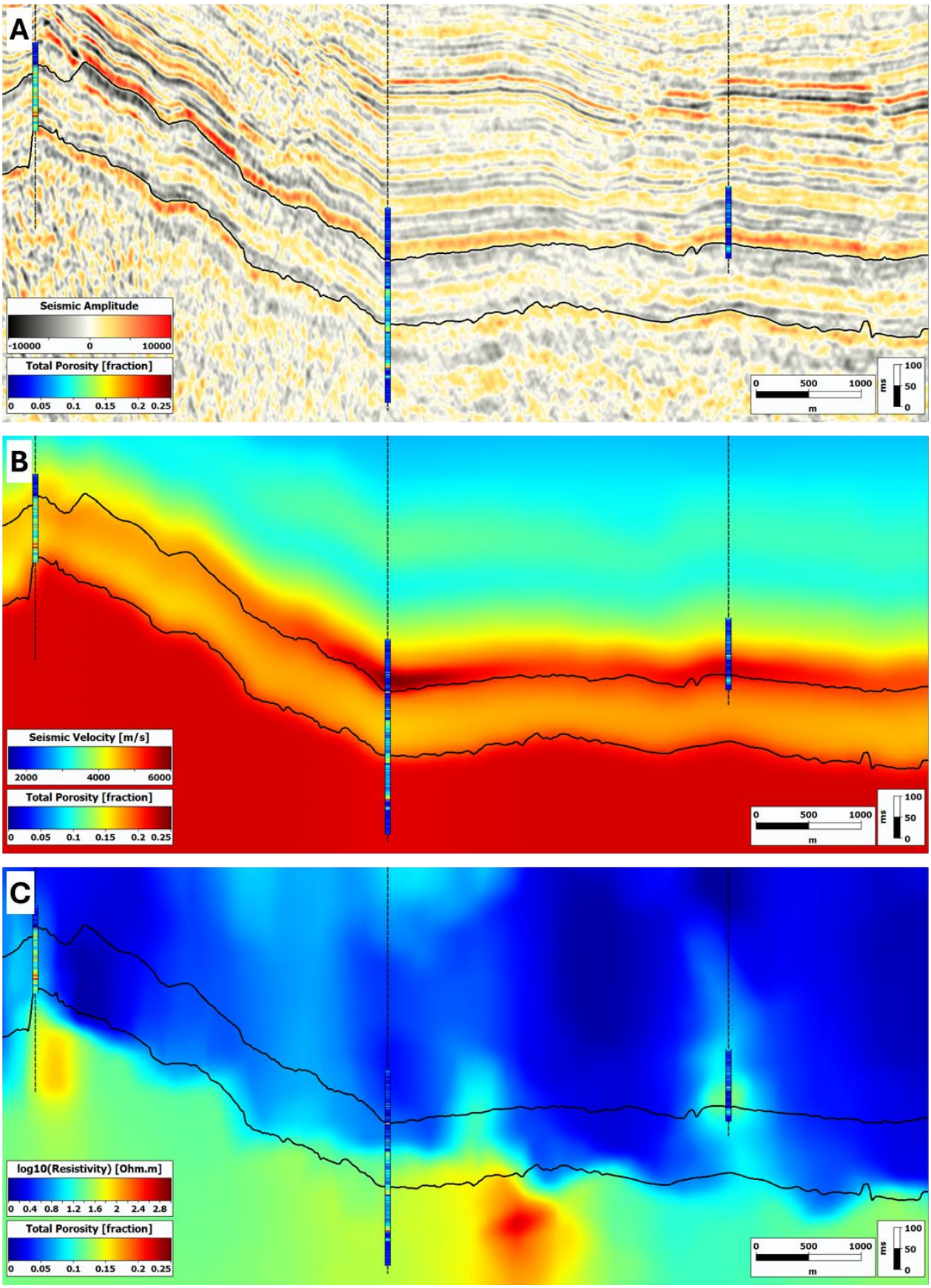


Figure 4. Field data from the Upper Rhine Graben (France) used to further test the proposed methodology. The data includes full-stack seismic amplitude (A), seismic velocity (B) and resistivity (C). Total porosity logs from the three available wells are superimposed on the sections. Black lines indicate the interpreted positions of the top and bottom of the reservoir. The bottom of the reservoir coincides with the top of the granitic basement.

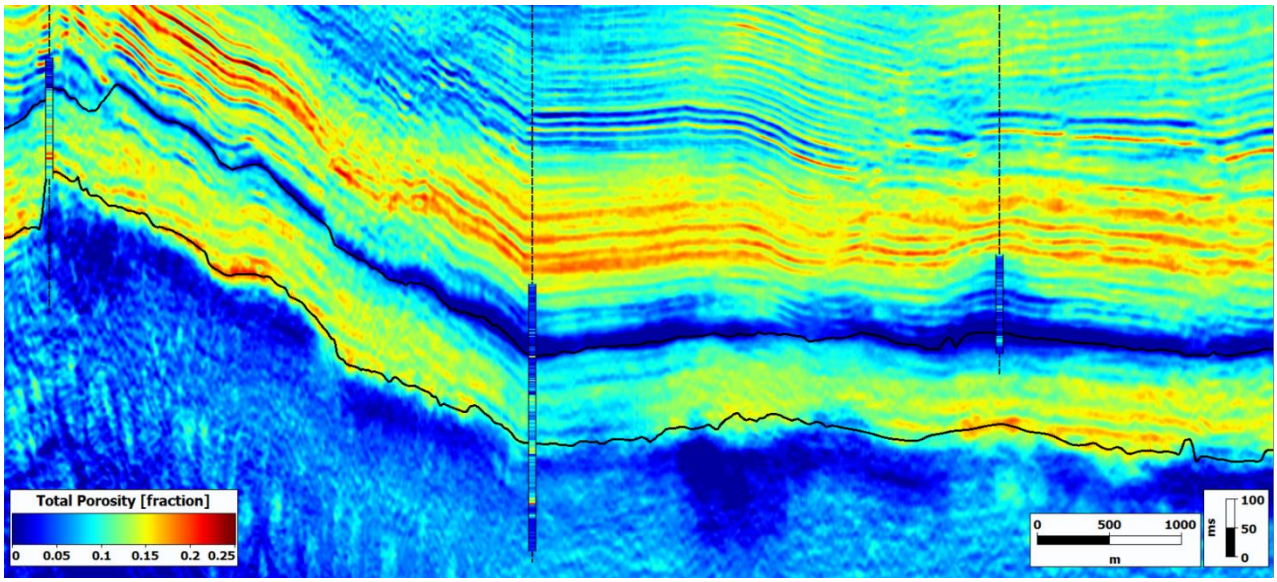


Figure 5. Porosity estimated from spectrally decomposed multiphysics data using an array of deep neural networks.

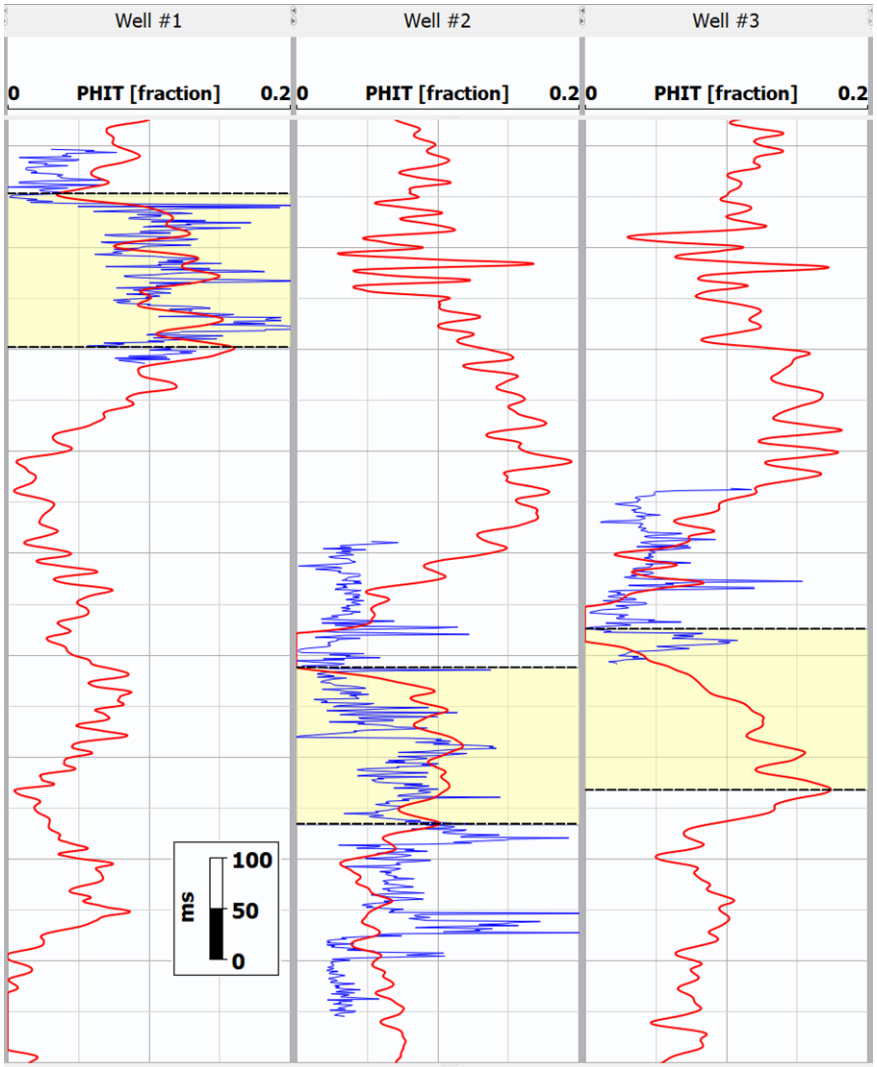


Figure 6. Comparison of estimated porosity (red curves) and measured porosity logs (blue curves) at three of the four well locations. The reservoir interval is highlighted in yellow.

4 CONCLUSION

A technique to estimate porosity from multiphysics data using an array of U-net-based neural networks is proposed. The input and target data are spectrally decomposed and each network in the array estimates the contribution to porosity within a specific frequency band based on the information from the multiphysics data within that band. The use of an independent network per frequency band allows tailoring its parametrization to extract the porosity features that exist at the corresponding spatial resolution. This technique has been compared to approaches using a single network and delivered superior results on synthetic data. The application to field data resulted in porosity values that correlate reasonably well with well-derived porosity and could help locate the most favorable locations for future geothermal wells.

Acknowledgements

The authors would like to thank Viridien for permission to share this work. We are also grateful to BRGM (French Geological Survey) and ES Geothermie for providing access to the field data presented in this study.

REFERENCES

Allo, F., Coulon, J.P., Formento, J.L., Reboul, R., Capar, L., Darnet, M., Issautier, B., Marc, S., and Stopin, A.: Characterization of a carbonate geothermal reservoir using rock-physics-guided deep neural networks, *The Leading Edge*, **40**, (2021), 751-758.

Gallardo, L.A. and Meju, M.A.: Joint two-dimensional DC resistivity and seismic travel time inversion with cross-gradients constraints, *J. Geophys. Res., Solid Earth*, **109**, (2004), Art. no. B03311.

Gao, G., Abubakar, A. and Habashy, T.M.: Joint petrophysical inversion of electromagnetic and full-waveform seismic data, *Geophysics*, **77**, (2012), WA3–WA18.

Guo, R., Zhou, H., Wei, X., Lin, Z., Li, M., Eldar, Y.C., Yang, F., Xu, S and Abubakar, A.: Deep joint inversion of multiple geophysical data with U-net reparameterization, *Geophysics*, **90**, (2025), WA61-WA75.

Hoversten, G.M., Cassassuce, F., Gasperikova, E., Newman, G.A., Chen, J., Rubin, Y., Hou, Z. and Vasco D.: Direct reservoir parameter estimation using joint inversion of marine seismic AVA and CSEM data, *Geophysics*, **71**, (2006), C1-C13.

Jo, H., Cho, Y., Pyrcz, M., Tang, H. and Fu, P. (2022): Machine-learning-based porosity estimation from multifrequency poststack seismic data, *Geophysics*, **87**, (2022), M217-M233.

Ma, J., Deng, Y., Li, X., Guo, R., Zhou, H. and Li, M.: Recent advances in Machine Learning - Enhanced joint inversion of seismic and electromagnetic data, *Surveys in Geophysics*, **46**, (2024), 197-225.

Ren, Y., Liu, B., Liu, B., Liu, Z. and Jiang, P.: Joint inversion of seismic and resistivity data powered by deep learning, *IEEE Transactions on Geoscience and Remote Sensing*, **62**, (2024), 5929614.

Tseng, H., and K. Lee: Joint inversion for mapping subsurface hydrological parameters, *Expanded Abstracts*, 71st Annual International Meeting, SEG, San Antonio, USA (2001), 1341–1344.