Geophysical Journal International

Geophys. J. Int. (2024) **237,** 378–388 https://doi.org/10.1093/gji/ggae053
Advance Access publication 2024 February 07
GJI Applied and Marine Geophysics

Deep attributes: innovative LSTM-based seismic attributes

G. Roncoroni, E. Forte and M. Pipan

Department of Mathematics, Informatics and Geosciences, University of Trieste, Trieste 34128*, Italy. Email:* groncoroni@units.it

Accepted 2024 February 6. Received 2023 December 22; in original form 2023 July 28

SUMMARY

Seismic attributes are derived measures from seismic data that help characterize subsurface geological features and enhance the interpretation of subsurface structures: we propose to exploit the hidden layers of Long–Short Time Memory neural network predictions as possible new reflection seismic attributes. The idea is based on the inference process of a neural network, which in its hidden layers stores information related to different features embedded in the input data and which usually are not considered. Neural network applications typically ignore such intermediate steps because the main interest lies in the final output, which is considered as the exclusive exploitable feature of the process. On the contrary, here we analyse the possibility to exploit the intermediate prediction steps, hereafter referred as 'deep attributes' because they are produced by a deep learning algorithm, to highlight features and emphasize characteristics embedded in the data but neither recognizable by traditional interpretation, nor evident with classical attributes or multi-attribute approaches. Nowadays, classical signal attributes are numerous and used for different purposes; we here propose an original strategy to calculate attributes previously never exploited, which are potentially complementary or a good alternative to the classical ones.

We tested the proposed procedure on synthetic and field 2-D and 3-D reflection seismic data sets to test and demonstrate the stability, affordability and versatility of the entire approach. Furthermore, we evaluated the performance of deep attributes on a 4-D seismic data set to assess the applicability and effectiveness for time-monitoring purposes and comparing them with the sweetness attribute.

Key words: Image processing; Machine learning; Neural networks, fuzzy logic.

1 INTRODUCTION

Seismic attributes have been developed since the 1970s to help reflection seismic interpretation exploiting additional quantities to classical reflection amplitude and multitrace horizons correlation. In fact, the basic principle of seismic data interpretation is the detection of reflections that are related to subsurface impedance contrasts (e.g. Anstey 2013). Therefore, seismic horizons are linked not only to stratigraphic or structural contacts, but also to porosity or fluid content variations within the same geological unit.

An incredibly large number of seismic attributes have been developed since the first applications. While most of the attributes are well defined from the mathematical point of view, their correlation with specific physical parameters or geological elements is still not apparent (Li & Zhao 2014). Since the 1990s, the further development of new seismic attributes or the more sophisticated calculation of already available ones benefited from 3-D seismic surveys, characterized by a spatial data coverage that was previously unconceivable (Chopra & Marfurt 2005). In addition, 3-D based seismic attributes improved (semi) automated data interpretation and 3-D (auto) picking and surface/volume extraction, as well as more accurate estimates of reservoir parameters and their spatial variations.

Following the definition of Chopra & Marfurt (2007), seismic attributes are 'any measure of seismic data that helps us visually enhance or quantify features of interpretation interest'; we here propose a new approach to automatically extract seismic attributes from reflection seismic data set based on neural network (NN) calculations.

In recent years, NN, and in a broader sense, deep learning (DL) techniques have been extensively applied in different branches of geophysics (a comprehensive review is provided by Yu & Ma 2021), and particularly in different steps of reflection seismic surveys with several different objectives:

(i) Optimize data acquisition maximizing the information while avoiding redundant data (e.g. Lu *et al*. 2019; Deighton & Olsen 2021).

(ii) Perform data simulations creating realistic synthetic data sets (e.g. Moseley *et al*. 2020; Roncoroni *et al*. 2021).

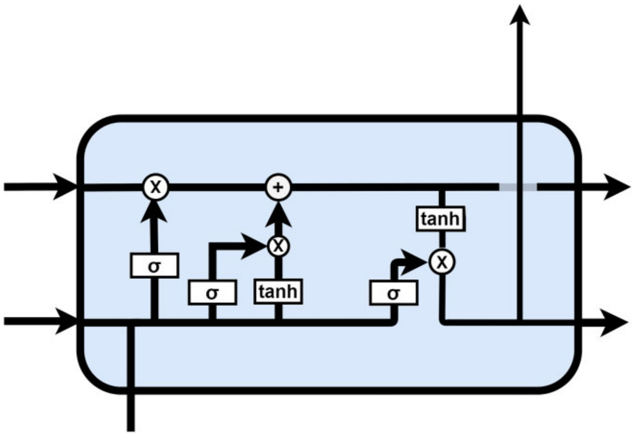

Figure 1. Conceptual scheme of a neuron in an LSTM, which contains four interacting layers. In the white rectangles, the activation functions, while in the white circles the mathematical operations are depicted.

(iii) Enhance data quality by means of interpolation, processing and imaging algorithms (e.g. Jia & Ma 2017; Wang *et al*. 2019; Hou & Hoeber 2020; Klochikhina *et al*. 2020).

(iv) Improve interpretation, data classification, information extraction and seismic horizon/surface picking (e.g. Di *et al*. 2018; Duan *et al*. 2019; Das & Mukerji 2020; Wrona *et al*. 2021; Roncoroni *et al*. 2022a; Roncoroni *et al*. 2022b; Sain & Kumar 2022).

(v) Perform data inversion and parameters extraction (e.g. Wang *et al*. 2019, 2020; Ruiz *et al*. 2021).

Alongside all these applications, we propose a totally new way of using hidden layer (HL) predictions, which are usually the 'transparent' steps of any NN lying in between data input and the expected output. The idea is based on the inference process of a Long–Short Time Memory (LSTM)-based NN. In its HL, an NN extracts information related to some features (i.e. attributes) embedded in the input data and uses them to get the required inference (i.e. the output). DL applications typically ignore the information from the intermediate steps because the main interest lies in the final output, which is considered the manageable result of the entire process. In other words, we analyse the possibility to exploit the intermediate prediction steps to highlight features and emphasize characteristics embedded in the data but recognizable neither by traditional interpretation, nor by classical attributes or multi-attribute approaches.

Since the obtained features are produced by a DL procedure, we decided to refer to them as deep attributes (DA). In addition, such a definition follows similar ones already used, even for slightly different issues, in other fields of research and technology like face recognition (Jadhav *et al*. 2016) and, more in general, to image analysis and pattern recognition (Kim *et al*. 2022).

To avoid any confusion, we remark that our procedure does not combine, mix, classify, fuse, extrapolate previous attributes. Therefore, it is different from the ones recently proposed by Qian *et al*. (2018) and further modified by Li *et al*. (2022) in which DL is used to 'extract deep features from seismic waveforms and combine multiple attributes for seismic attribute fusion at the same time' so typically exploiting the NN output and not considering its HL.

We applied our analysis on both synthetic and field 2-D and 3-D reflection seismics thus demonstrating the affordability and versatility of the whole procedure. Furthermore, we evaluated the performance of DA on a 4-D seismic data set to assess the applicability for time-monitoring purposes critically evaluating the performance as respect to a standard attribute.

2 METHODS

In this paper, we focus on LSTM-based NN, however, the proposed approach could be applied, in principle, to any NN geometry. The choice of LSTM is motivated by the causal nature of the recurrent neural network (RNN), which can provide a more reliable representation of the different signal components embedded in the seismic data. Other typical NNs like for instance U-Net or convolutive neural networks (CNN) work more on the geometrical features, while LSTM is more signal-related and combines both the dynamic and spectral information of the data. Although the peculiarities of each different NNs architecture are more suitable for different specific tasks, HL information (i.e. DA) can in principle always be exploited.

The simplest RNN is made up of a single neuron that receives an input, produces an output, and sends that output to itself and to the output vector. At each time step t, the recurrent neuron receives the inputs x(t), as well as its own output from the previous time step $y(t-1)$. Special neurons have been introduced to handle long-term dependencies: one of the most used is the LSTM, at first introduced by Hochreiter and Schmidhuber (1997). LSTMs are in fact explicitly designed to avoid the long-term dependency problems: they have a chain-like structure, as for the RNN, but the repeating module has a different architecture (see Fig. 1), with the peculiar ability to keep on information even from $x(t) < x(t-1)$. The capability to spot long time dependencies within a series (in our case a time series) is the reason for the choice of this type of layers for the proposed extraction method.

Another crucial point of the methodology is the NN geometry: the use of an encoder–decoder structure is directly linked to the typical encoder–decoder convolutive NN (ED-CNN). A classical geometry of ED-CNN is made of a chain of couples of CNN layers, linked with pooling layers in the encoder and with up-sampling layer in the decoder. A pooling layer takes values in an interval (defined as kernel) and gives out a single value, that is, the maximum values in this case; doing this we can perform a reduction of the length of the trace equal to the kernel size. An up-sampling layer makes exactly the opposite: it takes a single value at and replicates it kernel-size times.

The mix between this kind of neurons and the encoder–decoder geometry grants us that each values of the HL can be directly linked to the correct position in time or, in some way, to what happened at previous times. This assure the causality typical of seismic waves propagation.

In Fig. 2, we explain the main NN geometry used. If we try to focus on the change of a single time step, assuming the kernel size = 2 both for the pooling and for the up-sampling layers, we have before the first pooling layer point D1, which refers to time t. After the first pooling layer we get point D2, referring to time [t, t+1] and at the third layer, we have point D3 referring to time [t, t+3]. In the encoded version of the signal, we have point E1 and E2 referring both to time [t, t+5] and the information for their creation comes just from time $<t+5$, since the recurrent neuron does not go backwards. In the decoder part, we have exactly the inverse of what described above.

Since the base of the methodology is the down sampling of the seismic wavelet, we have to take care of the input sampling rate. In order to estimate the optimal sampling rate for each signal (beside the Nyquist–Shannon theorems and related issues, see e.g. Dossi *et al*. 2018) we compute the amplitude spectrum in the Fourier domain with a unit sampling interval. We then resample it in order to get a desired frequency with this unitary value. This approach allows to work without changes on every input frequency and without

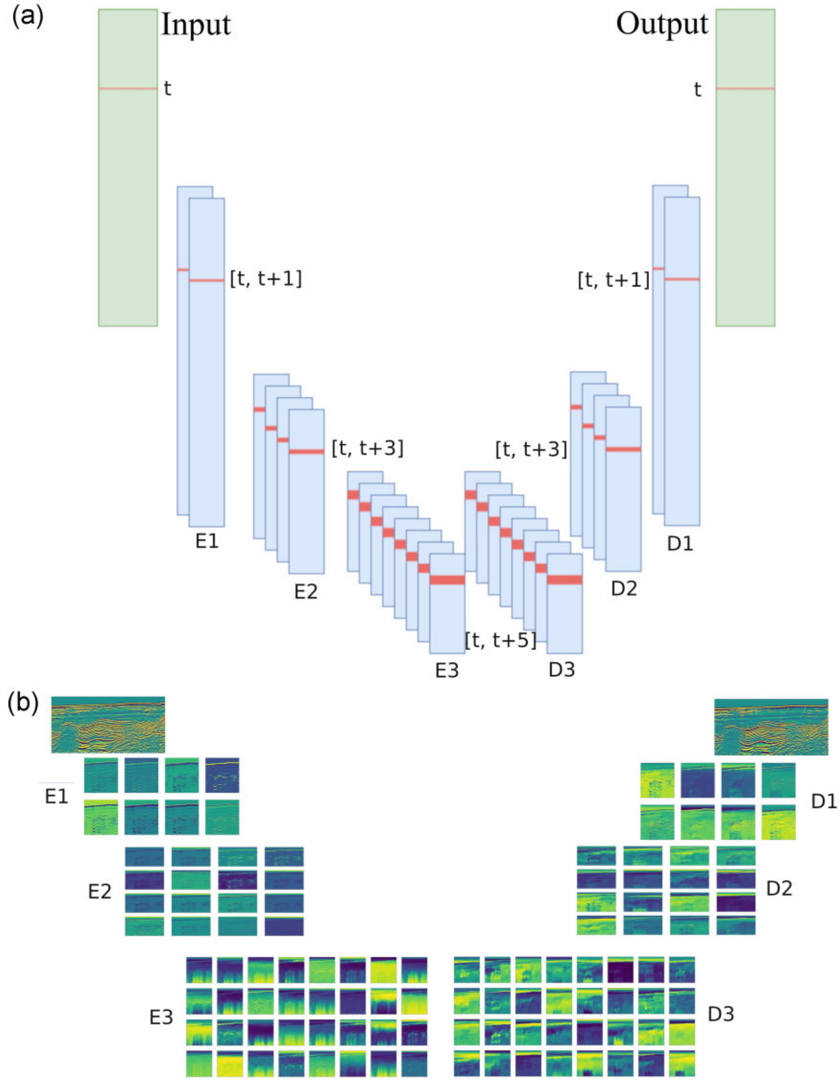

Figure 2. (a) Base structure of the proposed method. The geometry is based on an encoder–decoder LSTM NN with input equal to output. Recurrent layers are depicted in blue, while a sample at time t is highlighted to show the loss of resolution of DA. (b) Example of applications of the LSTM ED to the real data shown in Fig. 6. We can see the loss of resolution and the implementation of DA (lower part). We can also evaluate the good reconstruction of the input, comparing the two top right pictures (for a better visual comparison see also Fig. 7). E1, E2, E3 and D1, D2, D3 are the HL of the encoder and decoder parts, respectively.

considering the sample rate. This allows to generalize the values we will state later in this paper to any source frequency.

The best value was estimated taking into account the minimum depth of the NN needed to lose the wave nature during down sampling: if we oversample a trace, for example, a wavelet is discretized in 64 points, after three pooling layers we will still have 8 sample per wavelet and in turn the task of reconstructing the signal will be easy and no DA will be needed. On the other hand, if we use a too low sampled trace, we will have to create a very deep NN and we will lose a large amount of information.

By taking into account the maximum discretization and the loss value, since we still have to get a good reconstruction of our input data, we found as an optimal value 16 samples per wavelet. We tested this value on different data with various complexities and we always got good results.

As we can see in Fig. 3, the loss function has a plateau between 4 and 8 samples, while it starts decreasing with higher values. From 32 samples, the number of pooling layers is not enough to make the NN able to lose the waveform during the training: this means that we get very good results in terms of prediction (which is not the main objective of DA extraction), but poor results in term of features extraction.

Once we have defined the procedure, the main objective of the NN is to obtain an exact copy of the input by processing it into the encoder part so that it can retrieve the original trace in the decoder part.

The proposed workflow is:

(i) Create an LSTM-based NN that fits a specific problem.
(ii) Train the NN on the data.
(iii) Use the HL predictions as a set of additional information (i.e. DA) for improved seismic analysis and interpretation purposes.

Once we have a trained NN we can start working on the prediction of each single neuron from each HL. In Fig. 4, we show all the DA produced by the trained NN during the horizons extraction step.

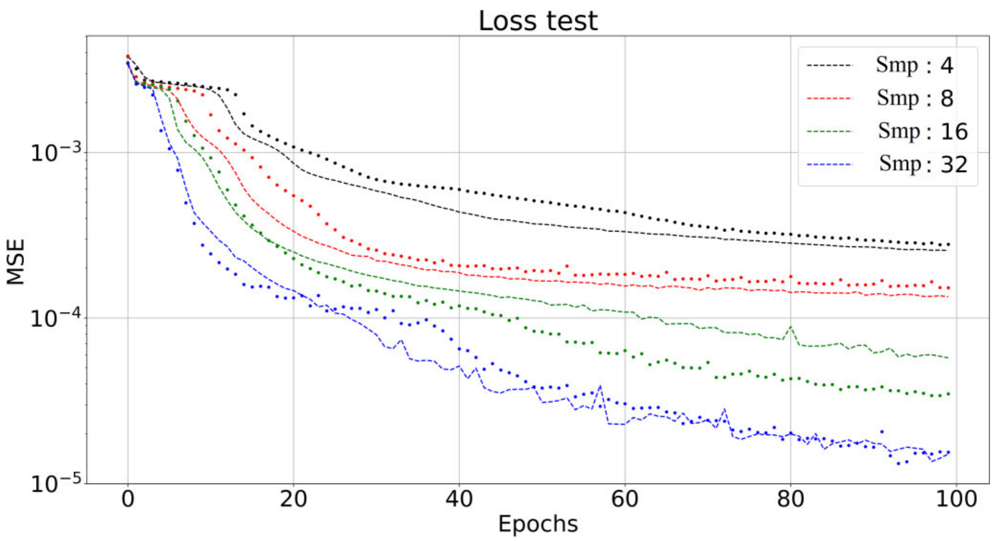

Figure 3. Loss function values with varying discretization steps of the input: with 4 and 8 samples we got to a plateau, while increasing the sampling rate we get better results but to avoid overfitting we chose 16 as a good value since it can get acceptable reconstructions without keeping the waveform memory. With 32 (or more) samples, the problem is overfitted. Dotted and dashed lines represent the validation and training losses, respectively.

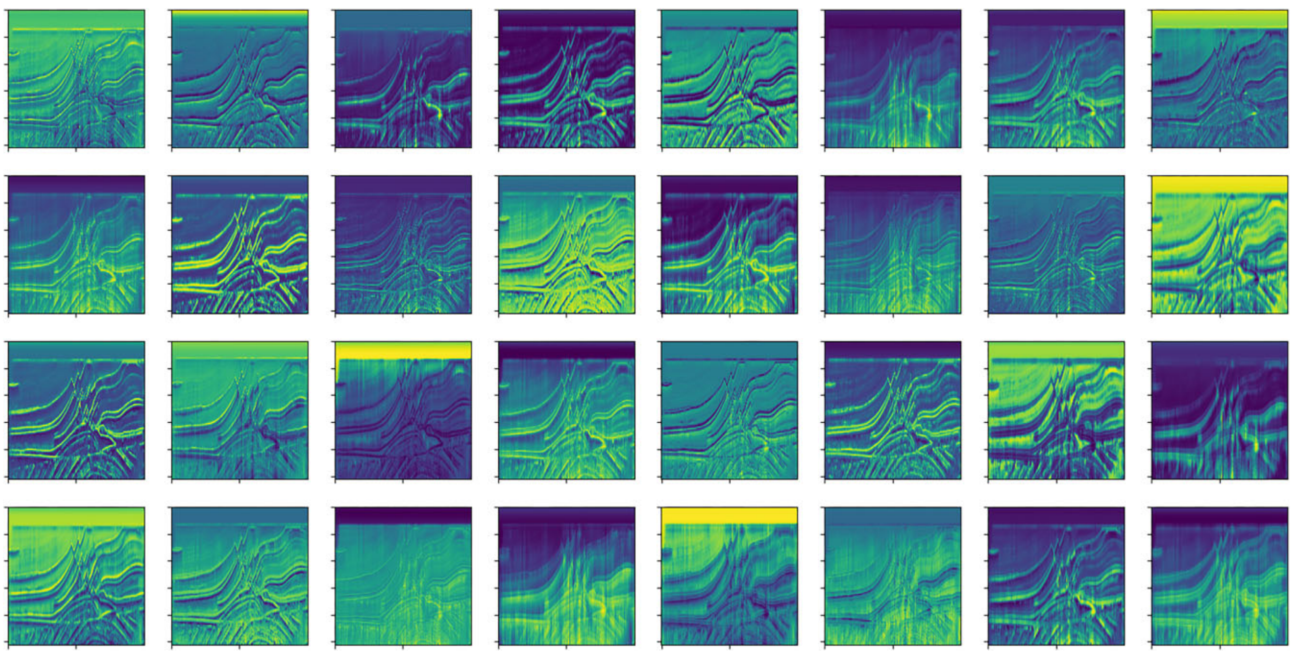

Figure 4. 32 DA of HL 3 of the encoder phase applied to a seismic profile of the Marmousi data set.

The proposed methodology can be applied to several different data sets and with slightly different purposes. In this paper, we propose two different approaches to evaluate its potential.

At first, we used the method to infer complex geometries in 2-D seismic profiles. We tested the effectiveness of the proposed DA approach on synthetic and field data considering the obtained DA then computing a principal component analysis (PCA) in order to reduce the number of attributes, while summarizing the whole information content.

We further exploit a single DA applying the same NN to a 3-D seismic survey repeated in different years (i.e. a 4-D data set) to monitor a CO_2 storage field. This allows to evaluate the stability and the repetitiveness of the methodology and its use for monitoring purposes.

3 RESULTS

3.1 Application on 2-D data

3.1.1 Synthetic example

In order to test the methodology, we performed at first an analysis on the elastic Marmousi model (Martin, 2004) which is a classical data

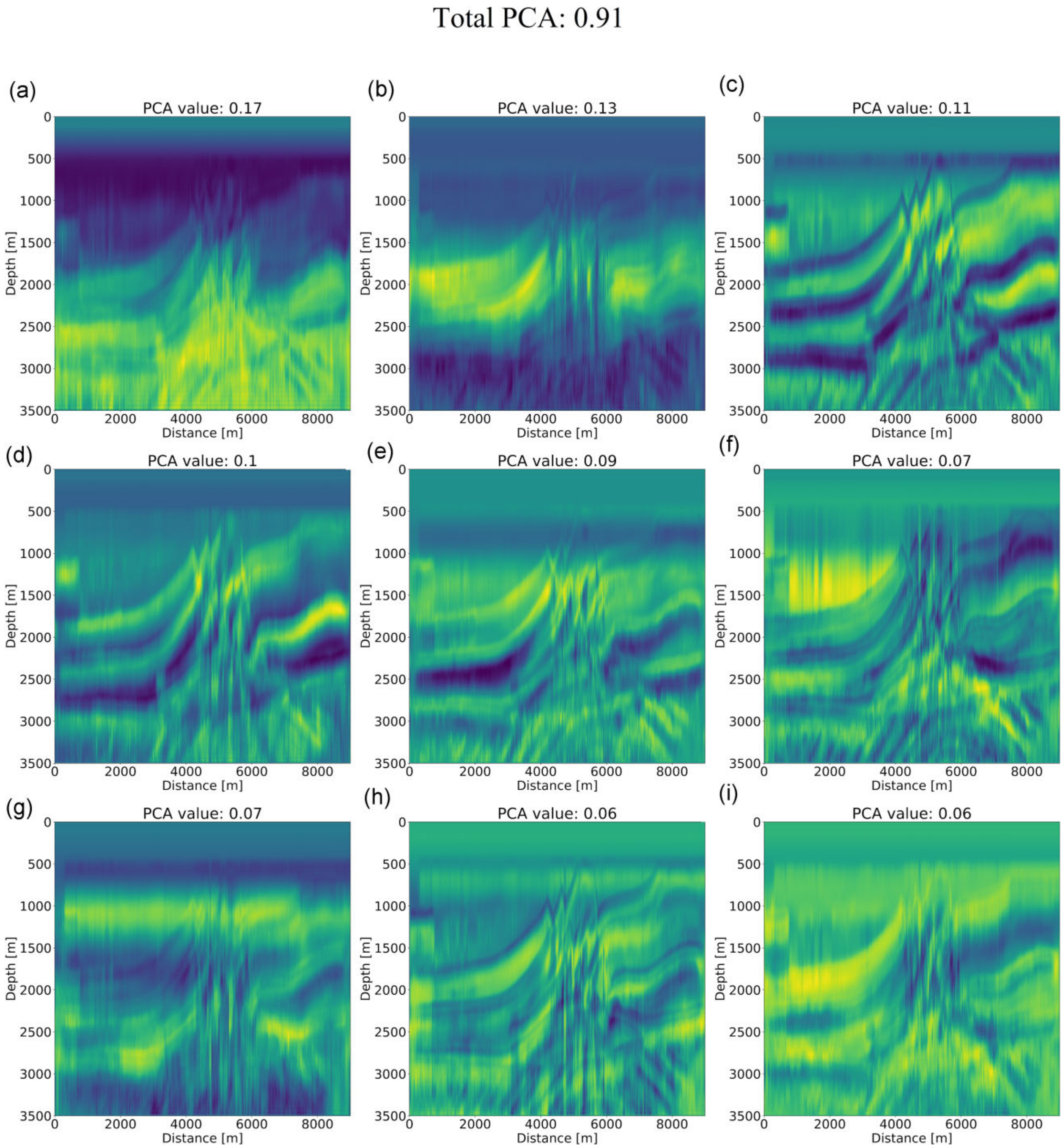

Figure 5. Test of the application of PCA on the HL depicted in Fig. 4. The total explained variance ratio is plotted on the top of each picture. As we can see, nine PCA results can explain 0.91 of the total variance.

set used for evaluation purposes. An examplary HL of the Encoder part of the NN applied to a profile of this data set is shown in Fig. 4.

From Fig. 4, it is apparent that some of the DA provide information about the single reflectors geometry, while other are more related to the different seismic signatures of the various zones. Moreover, it is interesting to notice that the resolution level of the DA panels is variable and spans over a wide range. Since it is difficult to get and summarize the whole information contained in many DA, we decided to apply a PCA to each HL to get a lower number of attributes that could condense all the extracted features. In order to better understand how much information we are losing with the PCA analysis we compute the explained variance ratio, which is the percentage of variance related to each DA. As we can see in Fig. 5, with just nine components we can explain 0.91 (i.e. 91 per cent) of the total variance of the first HL of the encoder.

From the PCA panels, it is apparent that they encompass not only the information related to the geometry and location of the reflections, but even more important, there are additional details related to materials in-between reflectors (i.e. transmission attributes) or to changes in the frequency content or main geological domains. In order to better understand the prediction performances of the proposed methodology, we directly compare some of the velocity

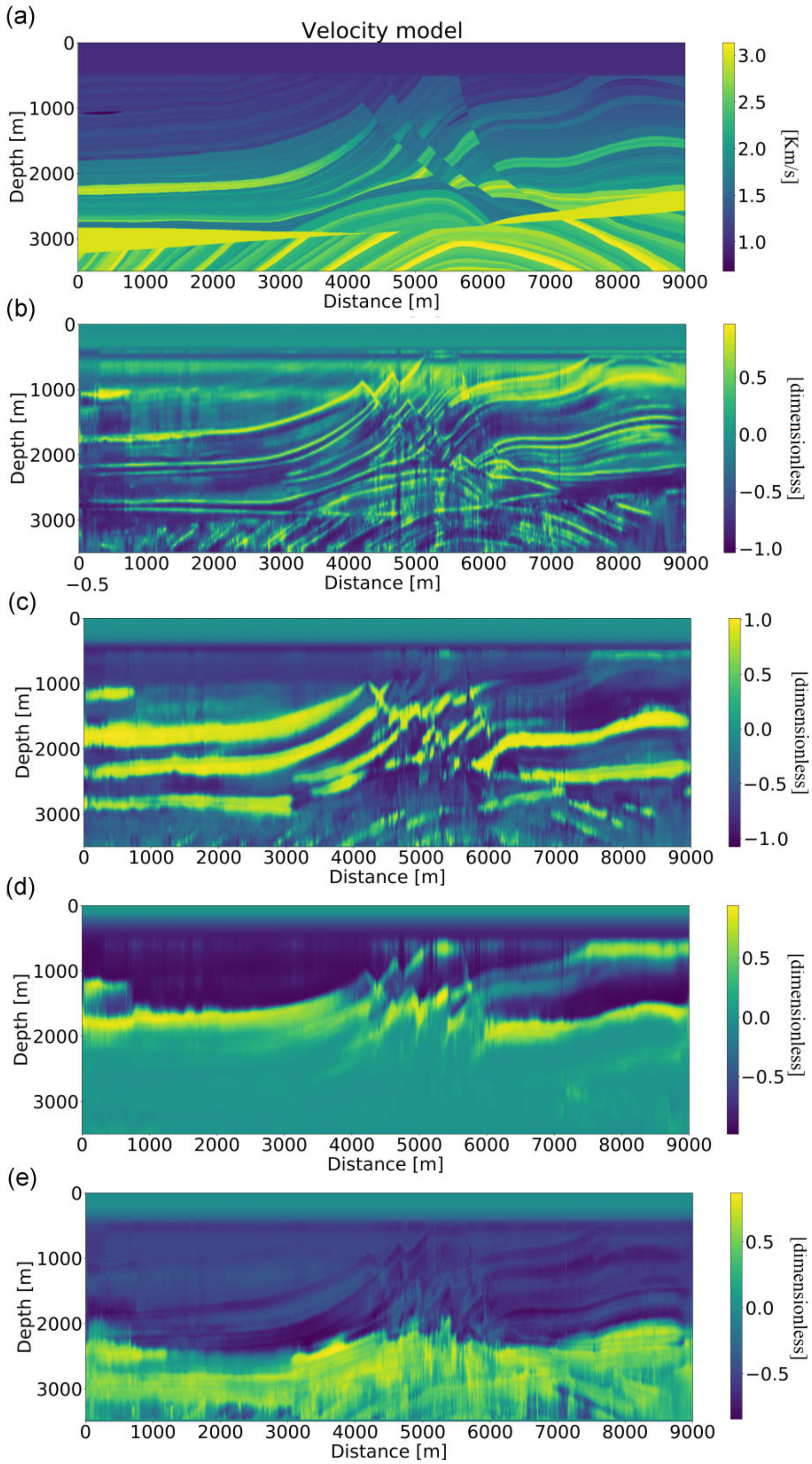

Figure 6. (a) Velocity model and (b)–(e) DA for an example of a seismic profile from the Marmousi model.

models used to generate the Marmousi data set with some randomly selected DA (Fig. 6).

It is interesting to note that DA can highlight elements at different resolution levels, and with different geological meaning. For instance, while in Fig. 6(b) several horizons related to the main velocity variations are apparent, in Fig. 6(c) a general overview of the stratigraphy is provided. Moreover, Figs 6(d) and (e) provide a clear distinction between the shallow lower velocity zone and the high velocity one in depth, with some slightly different details.

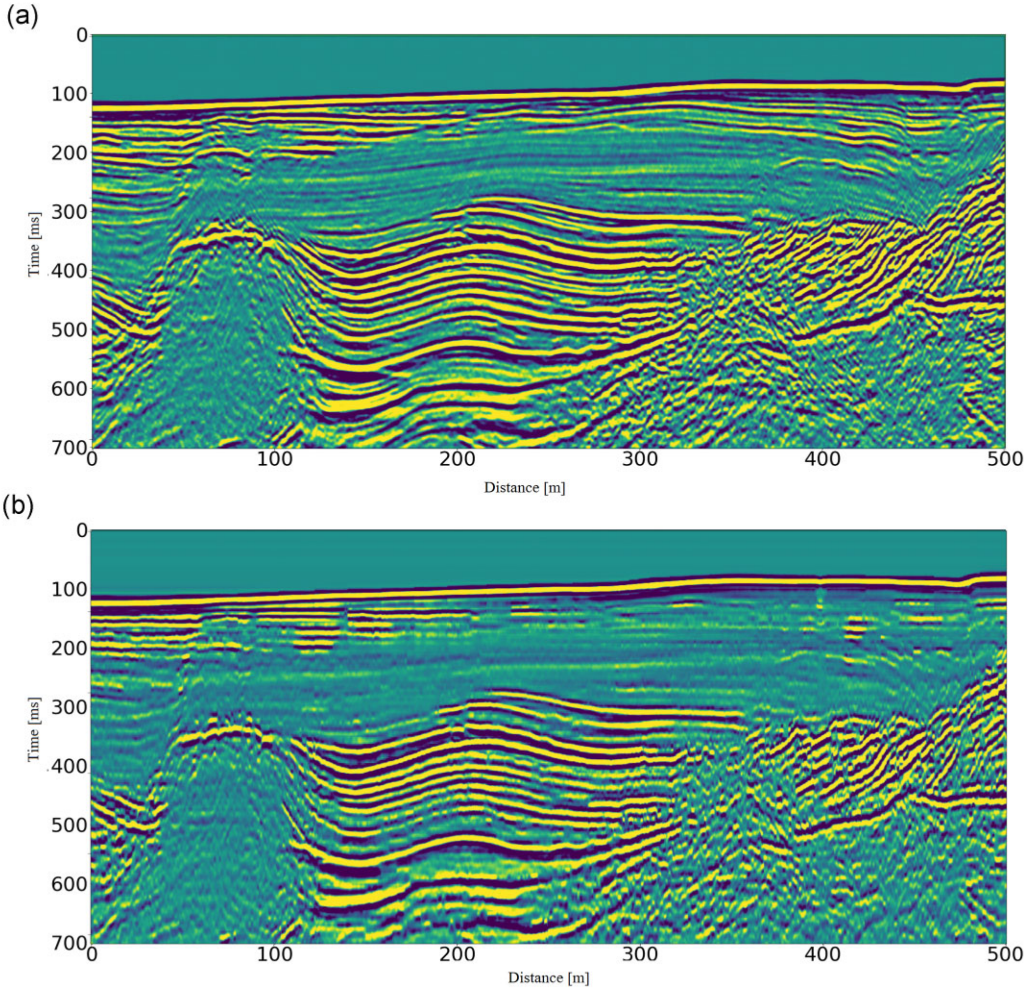

Figure 7. Example of a seismic profile of the WS10 exploration project: (a) field processed data and (b) NN reconstruction (output).

3.1.2 Field data example

We further tested the proposed methodology on field data of the 2-D marine seismic profile of the WS10 exploration project, obtained in autumn 2010 in the west Mediterranean Sea by the Istituto Nazionale di Oceanografia e Geofisica Sperimentale, which also performed the data processing (Geletti *et al*., 2014).). The selected portion of the seismic profile images a rifted margin of the eastern Sardo-Provençal Basin characterized by a faulted salt dome and by a portion of an almost undisturbed sedimentary sequence (Fig. 7).

The PCA analysis performed allow to condense the main part of DA information (equal to 92 per cent) in just four panels (Fig. 8). Beside the effectiveness in reducing the number of DA panels PCA is quite efficient in keeping both high- and low-frequency features. In fact, while panels (a) and (b) provide information about the general geological (and velocity) trends, and panels (c) and (d) give evidence of local structures like the salt dome and the most coherent horizons.

3.1.3 Application on a 4-D data set

The Sleipner 4-D seismic data set is a reference data set from the Sleipner CO_2 storage site. Several seismic surveys have been performed at the storage site since 1994, when the first survey was completed before the commencement of fluid injection. This data set contains seismic cubes from 1994 to 2010. In 2007, Petroleum Geo-Services (PGS) performed a full reprocessing of 1994, 2001, 2004 and 2006 data sets collected with the same geometry, acquisition parameters, with an identical processing flow (Equinor 2020). An example of crossline for the four data sets is presented in Fig. 9.

We applied the technique to the processed data cube from 2001 acquisition and we selected just a DA panel that enhances the strong reflection of the injection area. We applied the same trained NN on the data set of all the years from 1994 to 2006 to test the stability of the approach. A time slice showing some of the results is provided in Fig. 10.

As we can see, the methodology is stable over different and independent data sets collected in different years and the single DA plotted is able to highlight the CO_2 saturated area and its limits. DA (Fig. 10a) gives results comparable to the ones provided by Sweetness attribute (Fig. 10b) which is a composite signal attribute calculated by dividing the instantaneous amplitude by the square root of the instantaneous frequency thus condensing both dynamics and spectral informations. DA results can be further compared with the analysis performed by Chadwick *et al*. (2010) on the same data set.

Values plotted in Fig. 10(a) are scaled to the maximum values of more recent 2006 time slice: results evidence that also the magnitude of the prediction is independent from the training data and that the

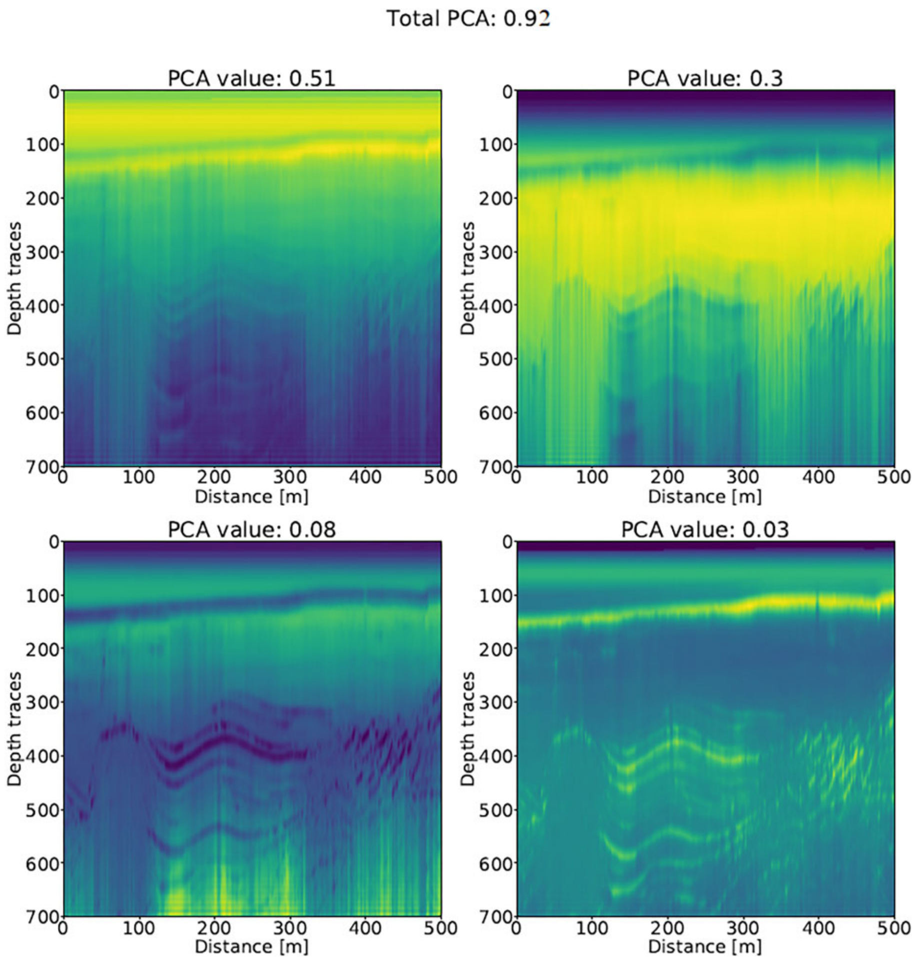

Figure 8. Results of the PCA analysis on the last layer before the encoding part of the geometry applied on the same data set as in Fig. 7a.

This is a crucial and not obvious result that points out the high stability and affordability of the whole approach, which is helpful not only for qualitative analyses but also for detailed monitoring purposes.

Moreover, a simple threshold (values > 0) on the DA on the 4-D cube is effective to enhance the evolution of the injected area through time (Fig. 11). While plots in Fig. 10 are very helpful in estimating the real boundaries of the CO_2 saturated volume, as already discussed, the purpose of the DA threshold is rather aimed at testing the stability of the approach.

4 DISCUSSION

We remark that the proposed approach is quite different from all the existing seismic attribute analyses, including the ones exploiting different strategies to combine and condense attributes (see e.g. Sain & Kumar 2022). In fact, PCA, clustering, cross plot, NN methods are effective in decreasing the total number of attributes, but all have to calculate the attributes beforehand and are somehow subjective in terms of the selection of attributes to be combined (Meldahl *et al*., 2001). The proposed technique, on the contrary, exploits LSTM networks using their intermediate prediction steps, which are usually disregarded, as a way to recover additional information through quantities here referred as DA. Meta-attributes, which are the aggregation of seismic attributes combined with the interpreter's insight through 'artificial intelligence' techniques (Sain & Kumar 2022) are indeed something different, but future research could be directed on the one hand to combine DA with the experience of the interpreter, and on the other to infer a one-to-one correlation between DA and specific physical parameters and/or geological features.

From the computational point of view the methodology is profitable on GPU machines, since LSTM gets a great improvement from cuda implementation. For a single 3-D Sleipner data cube, training time on a laptop with a RTX-1080 with 8Gb of memory takes only few hours for a 116 532 traces training. Prediction time on the same number of traces is just of 2.4 s for the whole data set.

5 CONCLUSIONS

We exploit the HLs of LSTM networks as DA of reflection seismic data sets. The proposed methodology is definitely different from the ones in which NN are applied to combine/condense previously calculated attributes.

The results obtained from synthetic and field data show that the new method is able to manage even complex geometries highlighting not only single seismic reflectors (i.e. features with high spatial frequency) but, even more important, the main geological and geophysical features related for instance to the low spatial frequency seismic velocity trend. PCA can be successfully applied on DA in

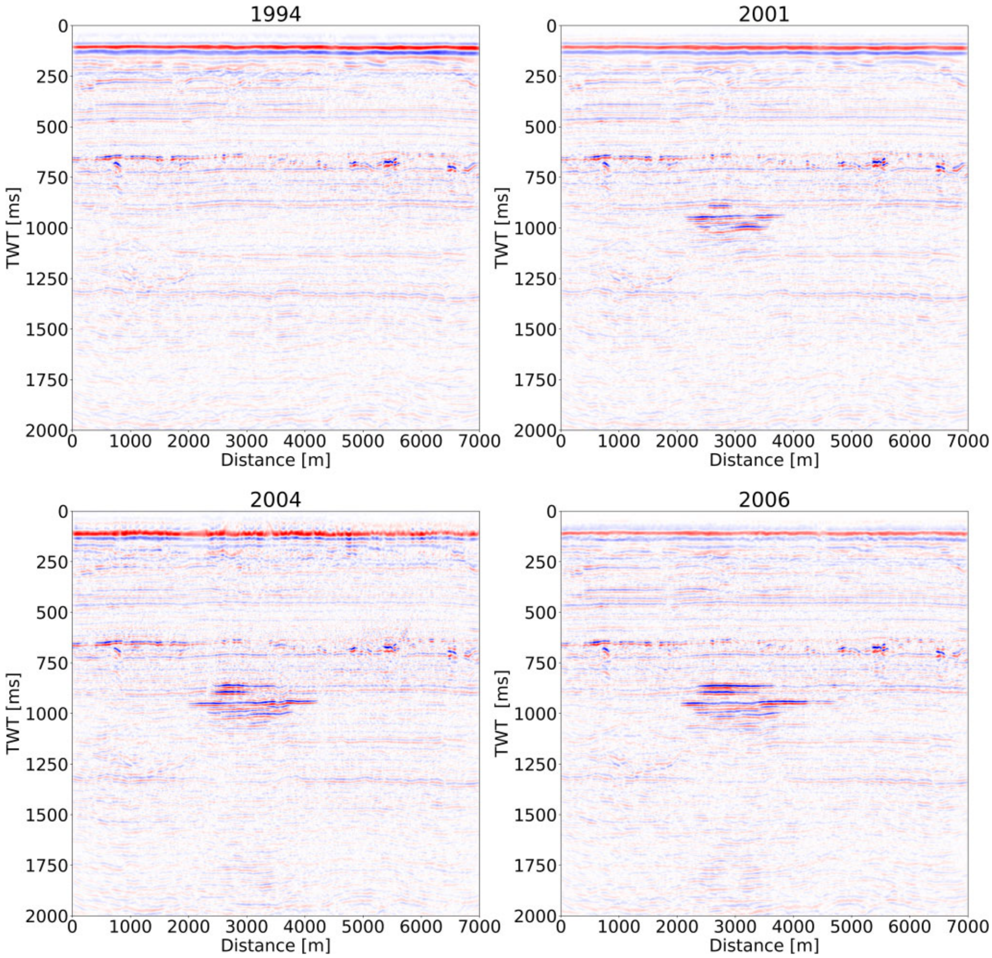

Figure 9. Example of a crossline of 1994, 2001, 2004 and 2006 data sets reprocessed by PGS in 2007 (data from Equinor 2020).

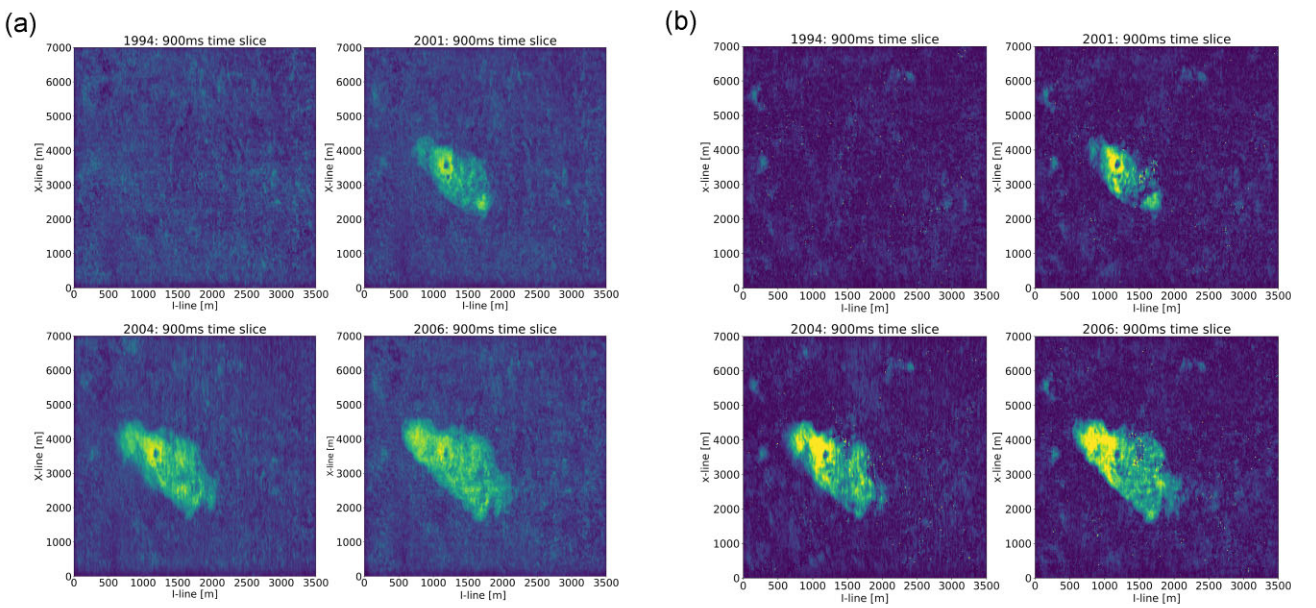

Figure 10. 900 ms time slices extracted from the 1994, 2001, 2004 and 2006 3-D Sleipner data sets representing the same (a) single DA and the (b) corresponding values of the Sweetness attribute. The NN training was performed on the 2001 data and then applied to all the other data as a stability test of the whole DA calculation procedure.

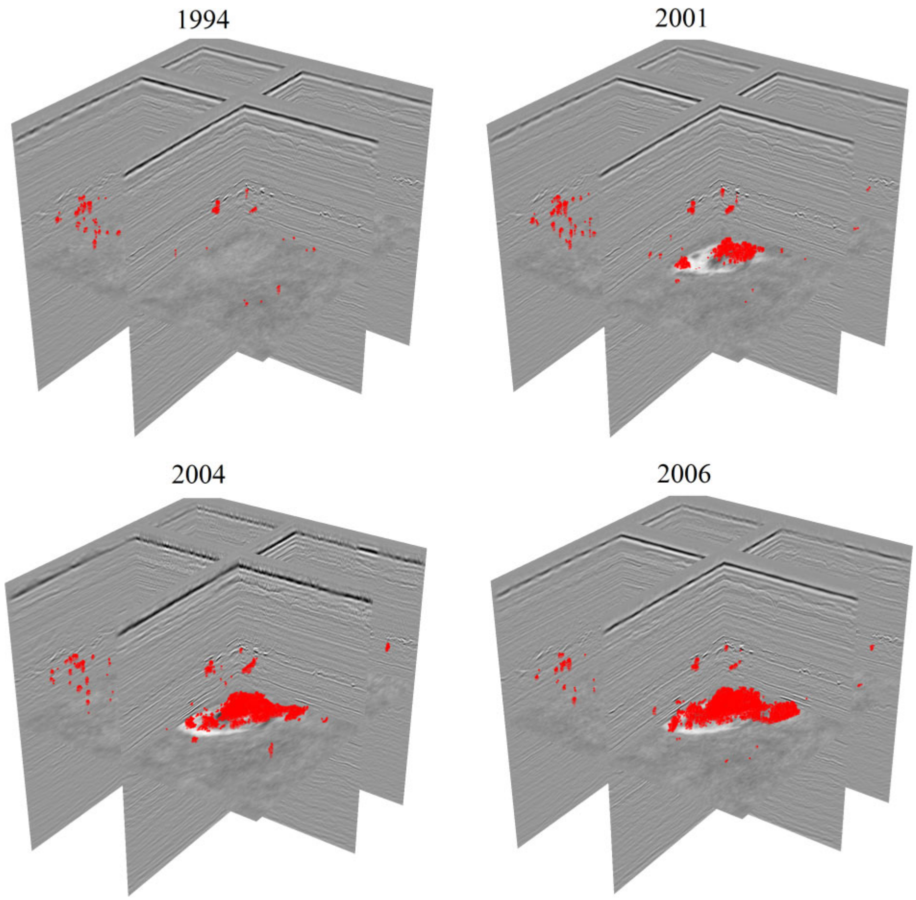

Figure 11. 3-D display of the same DA as shown in Fig. 10, where all values greater than 0 are plotted highlighted. See the text for more details.

order to reduce the total number of attributes and is effective in retaining the information content.

The stability tests performed demonstrate the high affordability of the procedure that can be trained on a single data set (or just on a portion) and then applied to larger seismic volumes. The application of DA to a 4-D data set collected to monitor a controlled CO_2 injection in an underground gas storage further demonstrate the stability and the repetitiveness of the methodology and in turn its full applicability for monitoring purposes.

Future research will be directed to infer specific correlations between DA and single or integrated physical parameters which would make the DA a new possible strategy for the quantitative subsurface petrophysical characterization at different scale and resolution levels.

ACKNOWLEDGMENTS

This research was partially supported the project 'Dipartimento di Eccellenza' of the Department of Mathematics and Geosciences of The University of Trieste. We gratefully acknowledge the support of Shearwater and Hallliburton Landmark through their academic grants. We furthermore thanks the PNRA projects IPECA (PNRA18_00186) and CRYOVEG (PNRA18_00288).

We would like to thank the editor prof. B. Rouet-Leduc and the two reviewers, one anonymous and prof. D. Verschuur for their advisable comments and suggestions.

DATA AVAILABILITY

The full implementation of the DA extraction flow can be found at https://github.com/Giacomo-Roncoroni/DeepAttributes.

REFERENCES

Anstey, N.A., 2013, *Seismic Interpretation,* Dordrecht: Springer Netherlands. https://public.ebookcentral.proquest.com/choice/publicfullrecord.aspx?p=5577515.

Chadwick, A., *et al.*, 2010. Quantitative analysis of time-lapse seismic monitoring data at the Sleipner CO2 storage operation. *Leading Edge,* **29,** 170–177.

Chopra, S. & Marfurt, K.J., 2005. Seismic attributes — a historical perspective. *Geophysics,* **70,** 3SO–28SO.

Chopra, S. & Marfurt, K.J., 2007. Seismic attributes for prospect identification and reservoir characterization, *SEG,* pp.481. doi: 10.1190/1.9781560801900.

Das, V. & Mukerji, T., 2020. Petrophysical properties prediction from prestack seismic data using convolutional neural networks. *Geophysics,* **85**(5), N41–N55,

Deighton, M. & Olsen, S., 2021. AI: a game changer in seismic acquisition and processing. *Geoexpro,* **18,** 4. source: https://geoexpro.com/ai-a-game-changer-in-seismic-acquisition-and-processing/

Di, H., Wang, Z. & AlRegib, G., 2018. Deep convolutional neural networks for seismic salt-body delineation, in Annual, AAPG, doi: 10.1306/70630Di2018.

Dossi, M., Forte, E. & Pipan, M., 2018. Quantitative analysis of GPR signals: transmitted wavelet, amplitude decay, and sampling-related amplitude distortions. *Pure appl. Geophys.,* **175**(3), 1103–1122.

Duan, Y., Zheng, X., Hu, L. & Sun, L., 2019. Seismic facies analysis based on deep convolutional embedded clustering. *Geophysics,* **84**(6), IM87–IM97.

Equinor. 2020. Sleipner 4D seismic dataset [Data set]. Archive2014. doi: 10.11582/2020.00005.

Geletti, R. *et al.*, 2014. The Messinian Salinity Crisis: New seismic evidence in the West-Sardinian Margin and Eastern Sardo-Proven¸cal basin (West Mediterranean Sea), *Marine Geology,* **351,** 76–90.

Hochreiter, S. & Schmidhuber, J. 1997. Long Short-Term Memory, *Neural Computation,* **9, 8 (Nov. 1997),** 1735–1780. doi: 10.1162/neco.1997.9.8.1735.

Hou, S. & Hoeber, H., 2020. Seismic processing with deep convolutional neural networks: opportunities and challenges, in *82nd EAGE Annual Conference & Exhibition,* vol. **2020,** pp. 1–5. doi: 10.3997/2214-4609.202010647.

Jadhav, A., Namboodiri, V.P. & Venkatesh, K.S., 2016. Deep attributes for one-shot face recognition. In: Hua, G. & Jégou, H.(eds) *Computer Vision – ECCV 2016 Workshops. ECCV 2016. Lecture Notes in Computer Science,* vol. **9915**. Springer, Cham. doi: 10.1007/978-3-319-49409-8_44.

Jia, Y. & Ma, J. 2017. What can machine learning do for seismic processing? An interpolation application. *Geophysics,* **82**(3), V163–V177. doi: 10.1190/geo2016-0300.1.

Kim, H., Lee, J. & Byun, H., 2022. Discriminative deep attributes for generalized zero-shot learning. *Pattern Recog.,* **124,** 108435. doi: 10.1190/geo2016-0300.1.

Klochikhina, E., Crawley, S., Frolov, S., Chemingui, N. & Martin, T., 2020. Leveraging deep learning for seismic image denoising. *First Break,* **38,** 41–48.

Li, K., Zong, J., Fei, Y., Liang, J. & Hu, G., 2022. Simultaneous seismic deep attribute extraction and attribute fusion. *IEEE Trans. Geosci. Remote Sens.,* **60**(5906410), 1–10.

Li, M. & Zhao, Y., 2014. Seismic attribute analysis, in *Geophysical Exploration Technology,* Li, M. & Zhao, Y.eds., chapter 5, Elsevier, pp. 103–131.

Lu, P., Xiao, Y., Zhang, Y.Y. & Mitsakos, N., 2019. Deep learning for 3D seismic compressive sensing technique: a novel approach. *Leading Edge,* **38,** 698. doi: 10.1190/tle38090698.1.

Martin, G.S., 2004. *The Marmousi2 Model, Elastic Synthetic Data, and an Analysis of Imaging and Avo in a Structurally Complex Environment*. PhD's thesis. University of Houston.

Meldahl, P., Heggland, R., Bril, B. & de de Groot, P., 2001. Identifying targets like faults and chimneys using multi-attributes and neural networks, *The Leading Edge, SEG,* **20**(5), 474–482. doi: 10.1190/1.1438976.

Moseley, B., Nissen-Meyer, T. & Markham, A., 2020. Deep learning for fast simulation of seismic waves in complex media. *Solid Earth,* **11,** 1527–1549.

Qian, F., Yin, M., Liu, X.Y., Wang, Y.J., Lu, C. & Hu, G.M., 2018. Unsupervised seismic facies analysis via deep convolutional autoencoders. *Geophysics,* **83,** 3, A39–A43.

Roncoroni, G., Forte, E., Bortolussi, L. & Pipan, M., 2022a. Efficient extraction of seismic reflection with Deep Learning, *Computers & Geosciences,* **166,** 105190. ISSN 0098-3004. doi: 10.1016/j.cageo.2022.105190.

Roncoroni, G., Forte, E., Bortolussi, L., Gasperini, L. & Pipan, M., 2022b. Polarity assessment of reflection seismic data: a deep learning approach, BGTA, **63–4,** 693–700.

Roncoroni, G., Fortini, C., Bortolussi, L., Bienati, N. & Pipan, M., 2021. Synthetic seismic data generation with deep learning. *J. appl. Geophys.,* **190,** 104347. doi: 10.1016/j.jappgeo.2021.104347.

Ruiz, R., Roubickova, A., Reiser, C. & Banglawala, N., 2021. Data mining and machine learning for porosity, saturation, and shear velocity prediction: recent experience and results. *First Break,* **39**(7), 71–76.

Sain, K. & Kumar, P.C., 2022. *Meta-Attributes and Artificial Networking: a New Tool for Seismic Interpretation,* American Geophysical Union, ISBN: 978-1-119-48176-8, 288pp.

Wang, B., Zhang, N., Lu, W. & Wang, J., 2019. Deep-learning-based seismic data interpolation: a preliminary result. *Geophysics,* **84,** V11–V20.

Wang, W., McMechan, G.A. & Ma, J., 2020. Elastic full-waveform inversion with recurrent neural networks. SEG Technical Program Expanded Abstracts, 860–864. doi: 10.1190/segam2020-3425921.1.

Wrona, T., Pan, I., Bell, R.E., Gawthorpe, R.L., Fossen, H. & Brune, S., 2021. 3D seismic interpretation with deep learning: a brief introduction. *Leading Edge,* **40,** 524–532.

Yu, S. & Ma, J., 2021. Deep learning for geophysics: current and future trends. *Rev. Geophys.,* **59,** e2021RG000742. doi: 10.1029/2021RG000742.