DOI 10.4430/bgo00409

# Polarity assessment of reflection seismic data: a Deep Learning approach

G. Roncoroni[1], E. Forte[1], L. Bortolussi[1], L. Gasperini[2] and M. Pipan[1]

[1] *University of Trieste, Department of Mathematics and Geosciences, Trieste, Italy*
[2] *ISMAR, CNR, Bologna, Italy*

(Received: 20 June 2022; accepted: 31 October 2022; published online: 30 November 2022)

***ABSTRACT*** We propose a procedure for the polarity assessment in reflection seismic data based on a Neural Network approach. The algorithm is based on a fully 1D approach, which does not require any input besides the seismic data since the necessary parameters are all automatically estimated. An added benefit is that the prediction has an associated probability, which automatically quantifies the reliability of the results. We tested the proposed procedure on synthetic and real reflection seismic data sets. The algorithm is able to correctly extract the seismic horizons also in case of complex conditions, such as along the flanks of salt domes, and is able to track polarity inversions.

**Key words:** polarity assessment, seismic phase, Deep Learning.

## 1. Introduction

The time delay of a reflected event between adjacent traces is a crucial parameter in the interpretation of seismic reflection data. Several seismic attributes help to analyse the lateral continuity of the reflected events, for instance coherency and phase-based attributes. Polarity is a further crucial element in this analysis and attributes, such as apparent polarity, can help in tracking reflections and variations related to fluid substitutions. Polarity is related to the reflection coefficient values, which are in turn related to the contrasts of acoustic impedance and also to the petrophysical parameters of the subsurface materials. Therefore, polarity changes with depth and polarity reversals along a reflector are essential indicators in the quantitative analysis of seismic data.

Besides the well-known problems related to phase distortions due to both seismic data acquisition, analysis and processing (Brown, 2009), the polarity assessment (a.k.a. phase detection) is often far from a trivial step during reflection seismic data interpretation. In fact, phase assessment is a common problem affecting autopicking methods (Forte *et al.*, 2016). In the ideal case, the final interpretation objective is to pick (and extract) the first phase of each reflection and the related peak amplitude. In this way, the subsurface geometries and the seismic impedance contrasts can be extracted and inversion procedures can be properly performed. Picking of reflected events can be somehow related to techniques applied in seismology, in particular first-break picking (Sabbione and Velis, 2010). The phase assessment is a generalisation of such an issue, with the main difference related to the extremely high number of reflections and records (i.e. traces) typical of reflection seismic data.

While in recent years several phase picking techniques have been developed and implemented that exploit machine learning techniques for both seismological (e.g. Cano *et al.*, 2021) and reflection seismic applications (e.g. Tschannen *et al.*, 2020), less effort has been specifically addressed to the polarity assessment.

In this paper, we describe and test a procedure for the polarity assessment of reflection seismic data based on a fully 1D Neural Network (NN) approach, requiring in input only the seismic data, thus minimising the subjectivity level and the intervention of the interpreter. The procedure is at first tested on 1D and 2D synthetic data, with various noise levels. We successively apply it on real cases previously used for testing other published methods, for an objective assessment of the effectiveness of the proposed procedure. In particular, we thoroughly analyse the benefits of using a 1D strategy, which can be applied to any type of seismic data set, including large 3D volumes thanks to its moderate computational load.

## 2. Methods

The algorithm is based on a Long Short-Term Memory (LSTM) architecture because we wish to ensure the causality of the data and the long-term memory fits the physics better behind the wave propagation (Hughes *et al.*, 2019). In fact, the bi-directional LSTM is a strategy able to improve the accuracy of NN classification (Guo *et al.*, 2019) and in the present case it can help the NN to find the correct shape of the wavelet by working on both sides of it (Fig. 1).

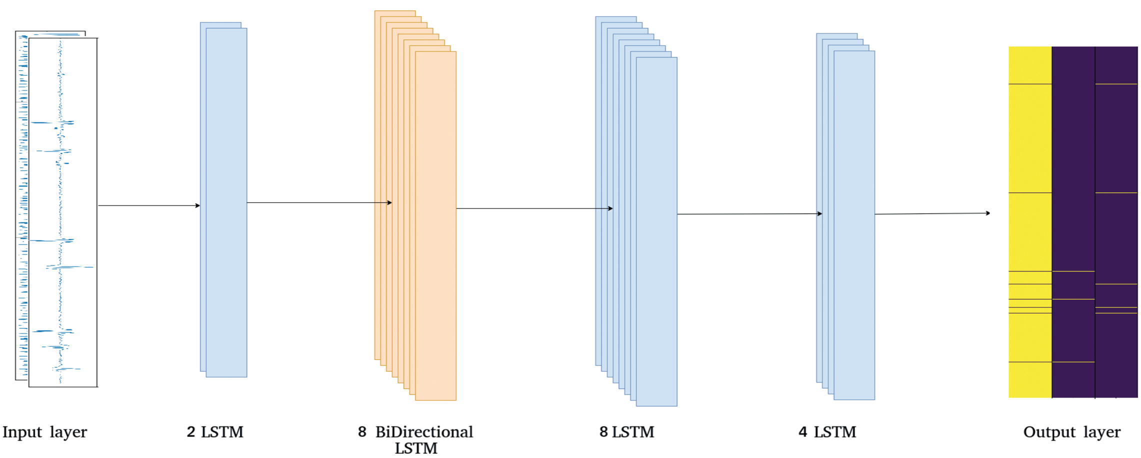


Fig. 1 - Example of NN architecture of the proposed algorithm. From left to right: input, hidden layers, and output with three possible categories: no reflection, positive reflection, negative reflection.

The output is driven by a dense layer with 3 neurons and a SoftMax activation function that outputs a probability value equal to 1 on the maximum phase of a reflection in time: the three classes represent negative polarity, no reflection, and positive polarity.

We use the CuDNNLSTM, a fast approximation of LSTM (Hochreiter and Schmidhuber, 1997) that works on Nvidia CUDA (Chetlur *et al.*, 2014) and AdaMax as optimiser (Kingma and Ba, 2014), a modified version of Adam with infinity norm and categorical cross-entropy as loss function (Mannor *et al.*, 2005).

We adopt a double input, namely synthetic data and its cosine of the instantaneous phase, which should give more information to the algorithm and exploits the possibility given by the NNs to deal with layered information.

The training is fully performed on synthetic data obtained from a convolutional model-based scheme, while the subsequent horizon extraction step can be applied to any type of field seismic data set.

The training data set is often a crucial issue for the performances of the algorithm on field data sets: we train the NN on synthetic data to avoid any link to a specific field data set and to have a complete control over the NN performance through the knowledge of the subsurface model that generates the training data.

After tests, we found that the best way to simulate field data for NN is to add pure random noise to the convolved trace and also to the reflection coefficient series before the convolution:

$$f(t) = \left\{ w(t) \times \left[ \mathrm{n}_1(t) + r(t) \right] \right\} + n_2(t) \tag{1}$$

where $r(t)$ is a randomly uniform distributed coefficient series with random values, while $n_1(t)$ and $n_2(t)$ are noise series randomly generated for each data. $n_1(t)$ has been added to simulate realistic cases. While $n_2(t)$ could represent an instrumental or environmental noise, $n_1(t)$ is a noise linked to the nature of the material and can be seen in the seismic trace as random noise convoluted to the seismic wavelet. The other term is the wavelet: we use different frequency ranges to generate the wavelet in order to simulate the stretching, spectral variations, and variability that occur under natural conditions.

We do not need to train on a previously defined wavelet since by feeding a NN with a recorded signal, the link between its temporal discretisation and the actual recording time is unknown for the NN, unless specified, as the NN takes as input only amplitude information. This allows us to define the temporal discretisation of the desired wavelet and just resample our data to the desired discretisation value, equal to 36 in this training.

In order to reduce the prediction uncertainty, we use the ensemble learning technique that exploits multiple learning algorithms to obtain better predictive inferences. In particular, we tested two solutions, namely: prediction with different NNs trained on a data set with the same characteristics, and prediction with the same NN on a single trace and on its time-inverted version. The two approaches produced similar results; we thus decided to use a single NN to reduce the required training effort.

With the application of ensemble learning, we obtain two different predictions: one on the single trace and the other on its inverted version in time. We successively combine them with the geometric mean, as it gives better results than, for instance, the arithmetic mean.

The prediction is given as a probability set that associates a probability value to each point: the value indicates the probability of the point to be a reflector, i.e. to belong to a reflecting surface. A minimum threshold above which a point is labelled as a reflector can be set.

The optimum threshold is estimated by evaluating the number of points classified as reflectors vs. the threshold. We perform this task by using the algorithm described in Satopaa *et al*. (2011). The threshold is set at the sharp inflection point clearly visible in the resulting curve, thus limiting the subjectivity of the choice. We apply this methodology both for positive and negative values (Fig. 2).

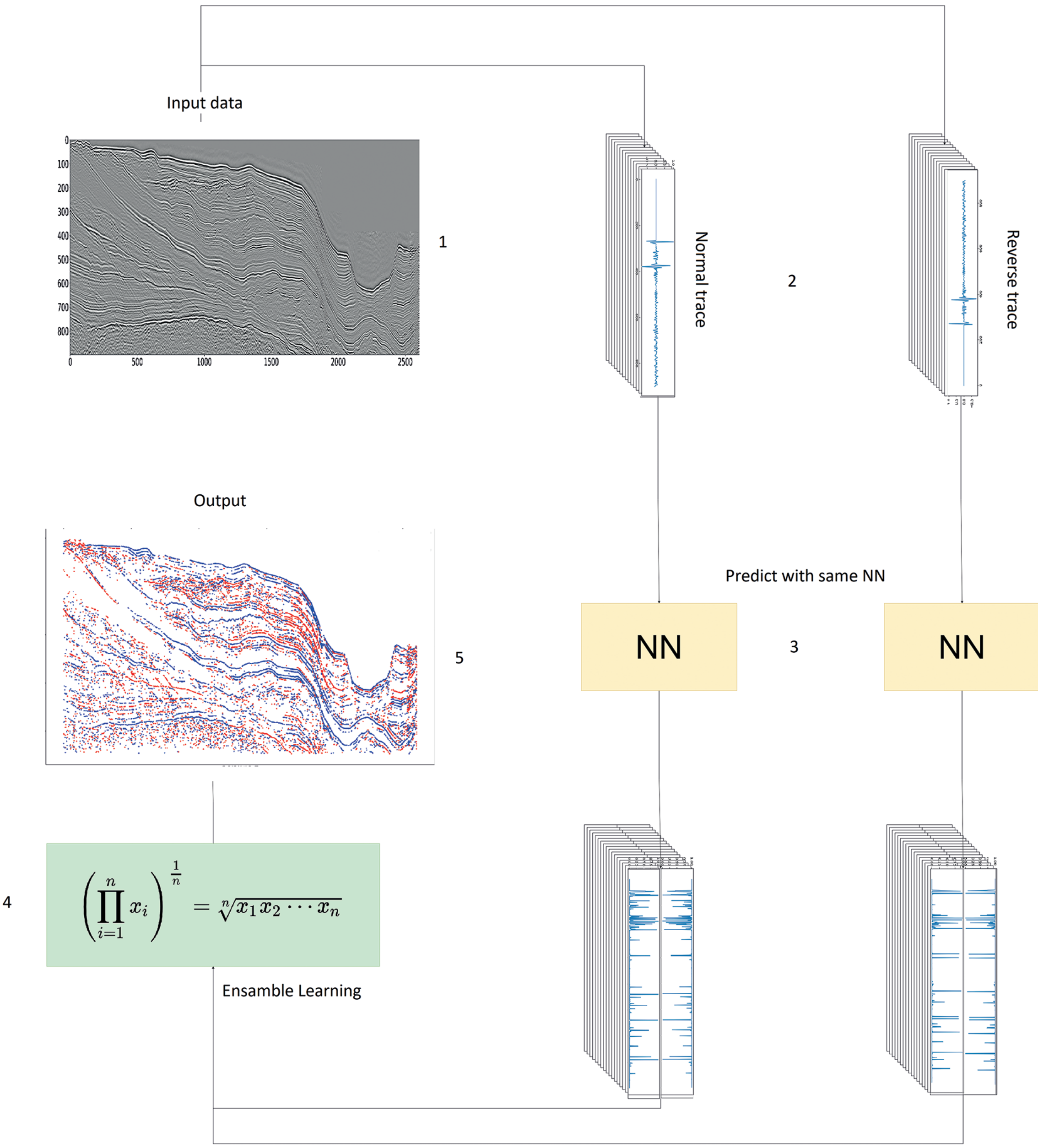


Fig. 2 - Example of the proposed workflow: 1) input data; 2) normal and reversed-in-time version of the input, reshaped as trace-by-trace matrix; 3) NN prediction; 4) ensemble learning, i.e. a geometrical mean of the input; 5) final results.

## 3. Results and discussion

In this section, we present the application of the method to both synthetic and real data.

In Fig. 3 we can see a synthetic 1D test. It shows the performances of the NN on a data set, which is similar to training data. As we can see, the NN performs very well and the output, expressed as a probability in range [0, 1] for positive and negative classes, shows very high values

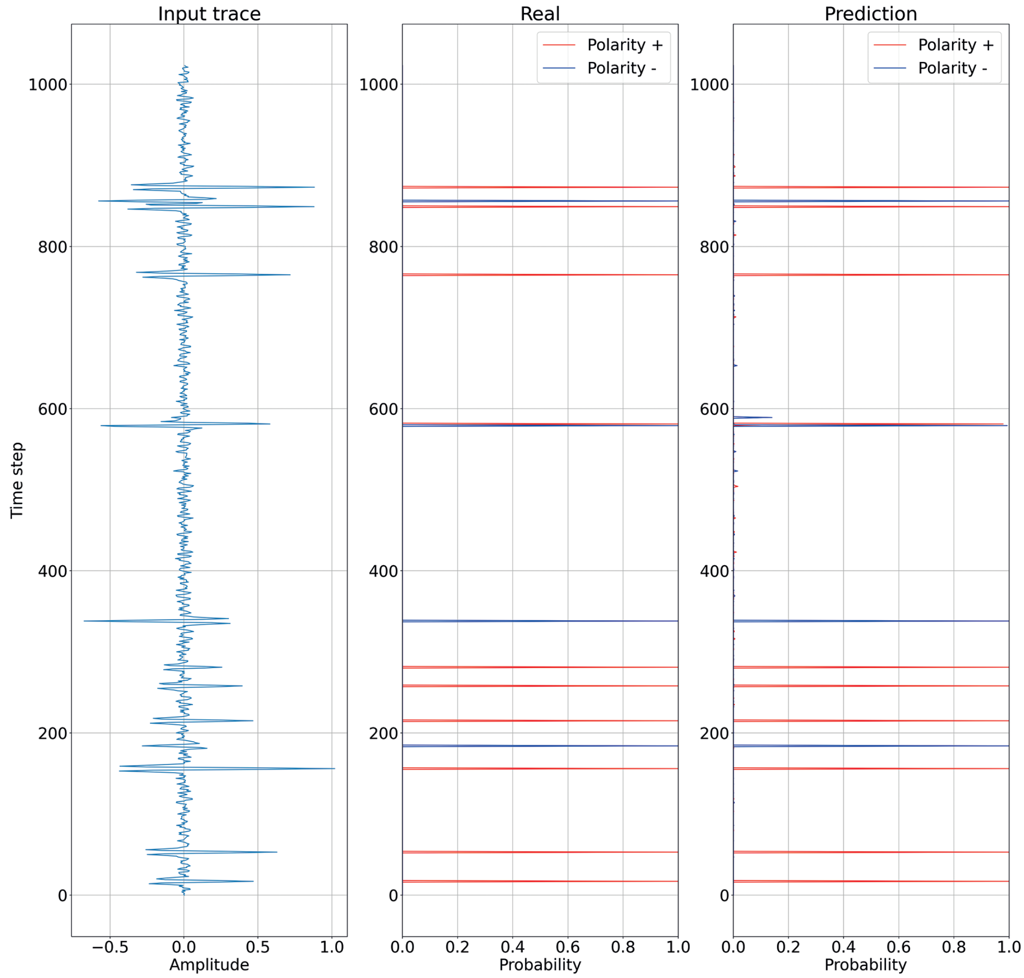


Fig. 3 - Synthetic example of prediction: from left to right we have the input trace, the real output and the NN prediction.

in almost all cases. At time step 590, the approach is able to predict a very close overlapping between positive and negative reflections and it correctly predicts a very low probability value for a wrong negative polarity signal at 596 time step.

In order to test the 2D potentiality of our trace-by-trace approach we generated some random 2D models. Fig. 4 shows the target data (left) and the NN prediction (right). As can be seen, here we applied a threshold to the data to decide if one point should be considered a reflection or not: in this case the minimum threshold was set to 0.9 probability.

The algorithm performances are quite good since it is able to spot reflection events with the right polarity and it just misses some points on low amplitude horizons, in which continuity is not complete, e.g. along the event marked with A in Fig. 4. In the prediction we can also spot some erroneous randomly distributed predictions, which could easily be deleted in a post processing step, since they do not show any lateral coherency.

In order to test the proposed methodology, we use a 2D marine seismic profile of the WS10 exploration project, acquired in autumn 2010 in the west Mediterranean Sea by the Istituto Nazionale di Oceanografia e Geofisica Sperimentale (OGS), which also performed the data processing (Geletti *et al.*, 2014). The selected portion of the seismic profile crosses a rifted margin

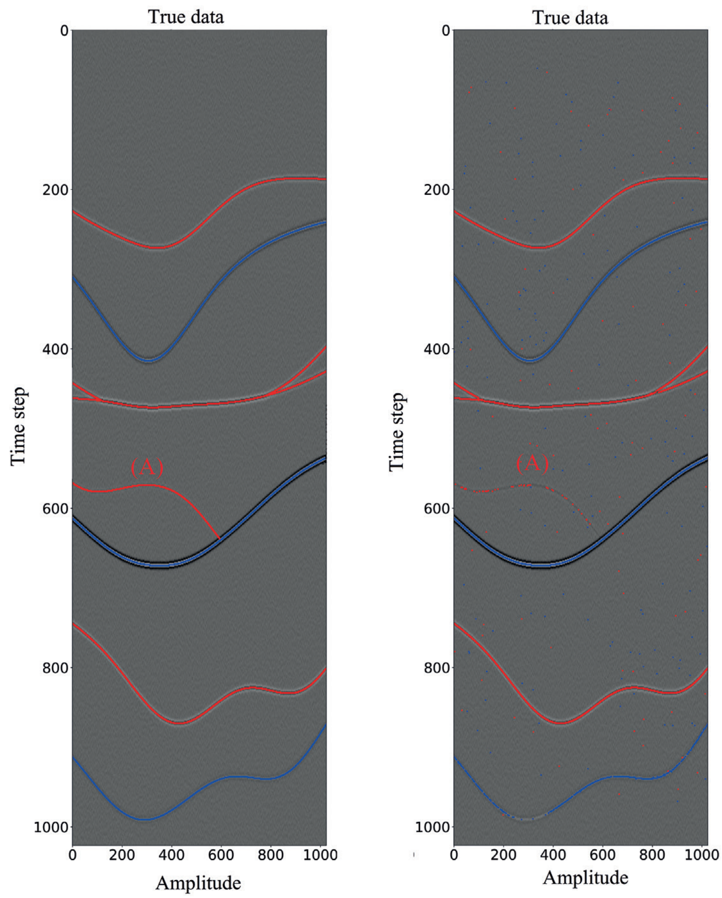


Fig. 4 - Prediction on a 2D synthetic seismic profile generated with a convolutive approach. Real polarity, on the left, and related prediction, on the right. With A we marked horizon with a low amplitude that was not continuously reconstructed by the methodology.

of the eastern Sardo-Provençal Basin characterised by a faulted salt dome and by a portion of an almost undisturbed sedimentary sequence (Fig. 5). For this reason, the analysed data represent an

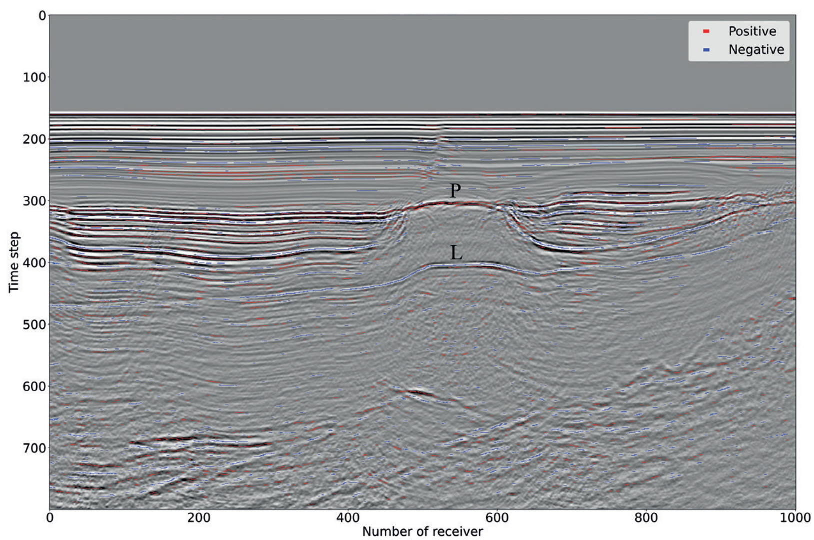


Fig. 5 - Prediction on a 2D marine seismic profile of the WS10 exploration project in the west Mediterranean Sea acquired by OGS, a salt dome with top at time step 300 is located between positions 450 and 650. P and L mark the positive and negative polarity, respectively, at the top and at the bottom of the salt body.

interesting and complex test for the proposed procedure. We focused on this portion also because it is exactly the same used in Forte *et al*. (2016) (Fig. 5a) to test an automated picking and phase assessment approach based on phase seismic attributes. The NN is able to properly extract all the main horizons, both where they are sub-horizontal (i.e. in the shallow part) and where they exhibit a significant dip (i.e. along the flanks of the salt dome), and is able to keep track of the inversion from positive to negative polarity under the salt dome (marked with P and L in Fig. 5, respectively).

The second test considers marine seismic data acquired on the Ionian continental platform (seismic source boomer). Over a total length of 25 km, the profile samples an extremely irregular sea bottom topography along with complex sub-bottom structures that include steep and conflicting dips, faulted horizons, thin layers, and sedimentary wedges. The algorithm correctly follows horizons and polarity reversals even in the steeply sloping parts (see A and B in Fig. 6).

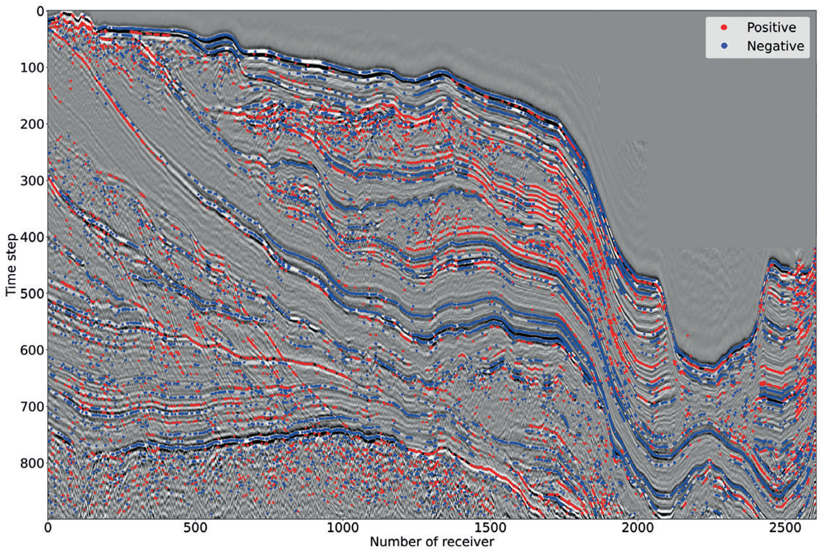


Fig. 6 - Example of prediction performed by the proposed NN on marine seismic data from the Ionian continental platform.

## 4. Conclusions

Bi-directional LSTM architecture shows good prediction performance and is an effective solution for polarity assessment as demonstrated by synthetic and field data set testing. The double input strategy (amplitude data and cosine of instantaneous phase) apparently provides a further improvement in performance, probably related to the NN characteristic of exploiting the layering of information and benefitting from the relative insensitivity to variations in the amplitude of the instantaneous attribute. Training on synthetic data allows great flexibility as well as to prepare NNs for real data applications of any complexity. The training phase for the NN presented in this work required a total time of 6 hours on a machine with GPUs[1] , which can be considered an acceptable computational cost when compared with the subsequent performance of the NN in the analysis of the field data. The training effort is further reduced by the solution of a single NN working on direct and reverse time data. The threshold for labelling the reflectors is set through a user-independent procedure, which greatly reduces the subjectivity of the whole process. The application of the algorithm to the seismic data of the eastern margin of the Sardinian-Provençal Basin and of

[1] On Cineca Marconi 100: 2×16 cores IBM POWER9 AC922 at 3.1 GHz, 4 x NVIDIA Volta V100 GPUs, 256 GB RAM.

the Ionian continental platform shows that the performance of the NN is not influenced by the complexity of the structural conditions or by the topography of the seabed.

Further analysis is needed to test the proposed strategy on different types of data (land seismic, GPR), but the tests performed to date show that the 1D approach and the complete adaptability to different wavelets and time scales make the algorithm robust and able to cope with a virtually unlimited range of applications.

**Acknowledgments.** This This research was partially supported by PNRA projects IPECA (PNRA18_00186) and CRIOVEG (PNRA18_00288) and by 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 Halliburton Landmark through their academic grants. We further thank the anonymous reviewers for their fruitful comments and suggestions.

REFERENCES

Brown A.; 2009: *Phase and polarity issues in modern seismic interpretation*. Search and Discovery, #40397, 22 pp.

Cano E.V., Akram J. and Peter D.B.; 2021: *Automatic seismic phase picking based on unsupervised machine-learning classification and content information analysis*. Geophys., 86, V299-V315, doi: 10.1190/geo2020-0308.1.

Chetlur S., Woolley C., Vandermersch P., Cohen J., Tran J., Catanzaro B. and Shelhamer E.; 2014: *cuDNN: efficient primitives for deep learning*. Arxiv, 9 pp., doi: 10.48550/arXiv.1410.0759, < arxiv.org/abs/1410.0759 >.

Forte E., Dossi M., Pipan M. and Del Ben A.; 2016: *Automated phase attribute-based picking applied to reflection seismics*. Geophys., 81, V141-V150, doi: 10.1190/geo2015-0333.1.

Geletti R., Zgur F., Del Ben A., Buriola F., Fais S., Fedi M., Forte E., Mocnik A., Paoletti V., Pipan M., Ramella R., Romeo R. and Romi A.; 2014: *The Messinian salinity crisis: new seismic evidence in the west-Sardinian Margin and eastern Sardo-Provençal Basin (west Mediterranean Sea)*. Mar. Geol., 351, 76-90, doi: 10.1016/j.margeo.2014.03.019.

Guo Z., Tian Y., Zhang D., Wang T. and Wu M.; 2019: *A novel stick-slip based linear actuator using bi-directional motion.* Mech. Syst. Sig. Process., 128, 37-49, doi: 10.1016/j.ymssp.2019.03.025.

Hochreiter S. and Schmidhuber J.; 1997: *Long short-term memory.* Neural Comput., 9, 1735-1780, doi: 10.1162/neco.1997.9.8.1735.

Hughes T.W., Williamson I.A.D., Minkov M. and Fan S.; 2019: *Wave physics as an analog recurrent neural network*. Sci. Adv., 5, eaay6946, doi: 10.1126/sciadv.aay6946.

Kingma D.P. and Ba J.; 2015: *Adam: a method for stochastic optimization.* In: Proc. 3rd International Conference for Learning Representations, San Diego, CA, USA, 15 pp., doi: 10.48550/arXiv.1412.6980, < arxiv.org/abs/1412.6980 >.

Mannor S., Peleg D. and Rubinstein R.; 2005: *The cross entropy method for classification*. In: Proc. 22nd International Conference on Machine Learning - ICML '05, Bonn, Germany, pp. 561-568, doi: 10.1145/1102351.1102422.

Sabbione J.I. and Velis D.; 2010: *Automatic first-breaks picking: new strategies and algorithms*. Geophys., 75, V67-V76, doi: 10.1190/1.3463703.

Satopaa V., Albrecht J., Irwin D. and Raghavan B.; 2011: *Finding a "Kneedle" in a haystack: detecting knee points in system behavior.* In: Proc. 31st International Conference on Distributed Computing Systems Workshops, Minneapolis, MN, USA, pp. 166-171, doi: 10.1109/ICDCSW.2011.20.

Tschannen V., Delescluse M., Ettrich N. and Keuper J.; 2020: *Extracting horizon surfaces from 3D seismic data using deep learning.* Geophysics, 85 (3), N17-N26, doi: 10.1190/geo2019-0569.1.

*Corresponding author:* Giacomo Roncoroni
University of Trieste, Department of Mathematics and Geosciences
Via Weiss 1, Trieste, Italy
Phone: +39 040 5582287; e-mail: groncoroni@units.it