



# Addressing uncertainty in seismic imaging by using deep learning surrogate model for reverse time migration

Rodolfo S. M. Freitas<sup>1</sup>, Carlos H. S. Barbosa<sup>2</sup>, Gabriel M. Guerra<sup>3</sup>, Álvaro L. G. A. Coutinho<sup>2</sup>, Fernando A. Rochinha<sup>1</sup>

<sup>1</sup>*Department of Mechanical Engineering, COPPE/Federal University of Rio de Janeiro R. Hélio de Almeida - Cidade Universitária, 21941-614, Rio de Janeiro, Brazil rodolfoismfreitas@gmail.com, faro@mecanica.coppe.ufrj.br*

<sup>2</sup>*Department of Civil Engineering, COPPE/Federal University of Rio de Janeiro R. Hélio de Almeida - Cidade Universitária, 21941-598, Rio de Janeiro, Brazil c.barbosa@nacad.ufrj.br, alvaro.lga.coutinho@gmail.com*

<sup>3</sup>*Department of Mechanical Engineering, Federal Fluminense University Rua Passo da Pátria, 156 Bl. D, Campus da Praia Vermelha, 24210-240, Niterói, Brazil gguerra@id.uff.br*

**Abstract.** In seismic exploration, many decisions are based on interpretations of seismic images, which are affected by the presence of multiple sources of uncertainty. Uncertainties exist in data measurements, source positioning, and subsurface geophysical properties. Reverse time migration (RTM) is a high-resolution depth migration approach useful for extracting information such as reservoir localization and boundaries. RTM, however, is time-consuming and data-intensive as it requires computing twice the wave equation to generate and store an imaging condition. RTM, when embedded in an uncertainty quantification algorithm (like the Monte Carlo method), shows a many-fold increase in its complexity and cost due to the high input-output dimensionality with computationally intensive requirements'. Hence, one of the main challenges facing uncertainty quantification in seismic imaging is reducing the computational cost of the analysis. In this work, we propose an encoder-decoder deep learning surrogate model for RTM under uncertainty. Inputs are an ensemble of velocity fields, expressing the uncertainty, and outputs the seismic images. We show by numerical experimentation that the surrogate model can reproduce the seismic images accurately, and, more importantly, the uncertainty propagation from the input velocity fields to the image ensemble.

**Reverse time migration, Deep Learning, Surrogate Modeling, Uncertainty Quantification**

## 1 Introduction

Seismic imaging is employed to delineate the salient geological features of the Earth subsurface. Imaging methods are popular in the Oil & Gas industry as they are designed to be focused on the more essential characteristics: the horizons bounding the regions of interest. They can also be used in conjunction with inverse methods such as Full Waveform Inversion [1]. Imaging methods are designed and built departing from the integration of specialized optical (illuminating) principles and physics-based models describing the wave propagation through heterogeneous media. A critical aspect arising from such arrangement is the potential computational cost required, as a large domain is to be illuminated, which implies solving partial differential equations (PDEs) associated with the wave models in that area. The situation tends to be more complicated as the excitation signals bear high-frequency content, which demands very fine grids in space and time. Such time-consuming tasks often hamper the use of high-fidelity codes constructed upon physics-based models. That becomes a more critical issue whenever one faces many-query applications like sensitivity analysis, design, optimization, or uncertainty quantification.

In this work, we develop a machine-learning model to alleviate computational costs to provide seismic images with quantified uncertainty [2]. In this context, we propose a Monte Carlo method (MC) to sweep a large ensemble of plausible velocity fields obtained by approximate methods, and, therefore, prone to uncertainties, to compute an ensemble of images aiming at characterizing the propagated uncertainties along with the seismic image processing. Here, we apply a deep learning architecture based on generative adversarial network (GAN) [3] to obtain a surrogate model for the RTM imaging under uncertainty. The main aspect of this approach is to train simultaneously two models. A generative model that returns the probabilistic characterization of the data, and a discriminative model that estimates the probability that a sample came from the training data rather than a generative model. More specifically, we construct the surrogate model using a generative model as baseline and the encoder-decoder architecture proposed by [4, 5] designed for problems cast as image-to-image regressions. Also, we build the discriminative model using the architecture proposed by [6]. Such surrogates are nonlinear mappings linking the uncertain velocity field to the seismic images. It is worth to highlight that differently from usual surrogates, we do not replace only a forward solver associated to a PDE, but the whole more

expensive imaging process. The surrogate can handle the high-dimensional inputs (velocity fields) and outputs (seismic images), leading to cost savings in processing and memory storage. We demonstrate through one example that such an approach enables producing seismic images with quantified uncertainty. Indeed, it can accurately reproduce the ensemble of images resulting from the MC uncertain propagation with much less computational effort. Moreover, it uses a limited training data as expected, which was confirmed by our results and efficiency estimation.

## 2 Reverse Time Migration under Uncertainty

RTM is a high-resolution depth migration technique providing useful subsurface images for extracting information such as reservoir localization and boundaries [7]. RTM relies on the two-way wave propagation equation, resulting in an imaging condition (IC) [8] computed over the space-time domain to be imaged. More specifically, the wave equation is solved twice, the first time to compute the forward-propagated wave due to seismic sources, followed by the computation of the backward-propagated wave induced by the recorded seismograms. Both solutions are needed to compute the IC. We calculate the forward wave isotropic acoustic case by solving,

$$\begin{aligned} \nabla^2 p(\mathbf{r}, t) - \frac{1}{v^2(\mathbf{r})} \frac{\partial^2 p(\mathbf{r}, t)}{\partial t^2} &= f(\mathbf{r}_s, t), \\ p(\mathbf{r}, t) = 0 \quad \text{on } \partial\Omega_D \quad \text{and} \quad \mathcal{B}p(\mathbf{r}, t) &= 0 \quad \text{on } \partial\Omega_{inf}, \\ p(\mathbf{r}, 0) = 0 \quad \text{and} \quad \frac{\partial p(\mathbf{r}, 0)}{\partial t} &= 0, \quad \mathbf{r} \in \Omega, \end{aligned} \quad (1)$$

where  $\Omega \subset \mathbb{R}^3$  denotes the domain to be imaged,  $\partial\Omega = \partial\Omega_D \cup \partial\Omega_{inf} \subset \mathbb{R}^2$  is the domain boundary and  $\partial\Omega_D$  and  $\partial\Omega_{inf}$  are non-overlapping boundary partitions.  $\partial\Omega_D$  is the portion of the boundary where Dirichlet boundary conditions are applied, representing, for instance, the free-surface. The operator  $\mathcal{B}$  represents the non-reflecting boundary condition [9] applied on  $\partial\Omega_{inf}$ . The pressure (the forward-propagated source wavefield)  $p(\mathbf{r}, t)$  is defined at the position  $\mathbf{r} = (r_x, r_y, r_z) \in \Omega$  and time  $t \in [0, T]$ . Moreover,  $v(\mathbf{r})$  is the compressional wave velocity spatial field, and  $f(\mathbf{r}_s, t)$  is the seismic source. The vector  $\mathbf{r}_s$  represents the seismic source position. The backward-propagated wavefield is calculated solving,

$$\begin{aligned} \nabla^2 \bar{p}(\mathbf{r}, \tau) - \frac{1}{v^2(\mathbf{r})} \frac{\partial^2 \bar{p}(\mathbf{r}, \tau)}{\partial \tau^2} &= s(\mathbf{r}_r, \tau), \\ \bar{p}(\mathbf{r}, \tau) = 0 \quad \text{on } \partial\Omega_D \quad \text{and} \quad \mathcal{B}\bar{p}(\mathbf{r}, \tau) &= 0 \quad \text{on } \partial\Omega_{inf} \\ \bar{p}(\mathbf{r}, 0) = p(\mathbf{r}, T) \quad \text{and} \quad \frac{\partial \bar{p}(\mathbf{r}, 0)}{\partial \tau} &= \frac{\partial p(\mathbf{r}, T)}{\partial \tau}, \quad \mathbf{r} \in \Omega. \end{aligned} \quad (2)$$

which is an equation similar to (1), but with a different source  $s(\mathbf{r}_r, \tau)$ , that is, the recorded signals at the receivers positioned in  $\mathbf{r}_r$ . Besides, the evolution in Eq. (2) is over the reverse time  $\tau = T - t$ . Thus, the backward-propagated wavefield  $\bar{p}(\mathbf{r}, \tau)$  is defined in  $\Omega$  and  $\tau \in [0, T]$ .

The IC dictates the quality and fidelity of the final RTM image. There are several possibilities, for instance, excitation ICs [10, 11], extend ICs [12, 13], wavefield decomposition ICs [14], and the zero-lag cross-correlation ICs [7, 15]. We have chosen the zero-lag cross-correlation between the forward and backward propagated waves at each point in  $\Omega$ ,

$$I(\mathbf{r}) = \int_0^T p(\mathbf{r}, t) \bar{p}(\mathbf{r}, \tau) dt. \quad (3)$$

As we wrap RTM into a sampling method, for taking into consideration the input uncertainties, the overall computational cost rises proportionally to  $N_{MC}$ , the cardinality of the ensemble of possible velocity fields. For each MC iteration, we solve the wave equation twice, one for the seismic source and other for the seismograms associated with it. The computation of the imaging condition uses both solutions (source wavefield, and receiver wavefield), retrieving from persistent storage the source wavefield to build the migrated seismic section and stacking the partial results over time ( $\mathbf{I}_{\Sigma_r}$ ), and over the number of seismograms ( $\mathbf{I}_{\Sigma_{shot\_id}}$ ). At the end of MC we have the discrete imaging condition set  $\{\mathbf{I}_1, \mathbf{I}_2, \dots, \mathbf{I}_{N_{MC}}\}$  where each  $\mathbf{I}_i$  is a vector in  $\mathbb{R}^N$ .

## 3 Deep Learning Surrogate

The main goal of surrogate models is to replicate the multivariate input-output mapping provided by physical models governed by PDEs to save computational costs. Performing uncertainty quantification in such conditions is often hampered whenever one faces high-dimensionality, the so-called curse of dimensionality. As pointed out in the literature, DNNs have proved successful in such situations by exploiting low-dimensional latent spaces and sophisticated training

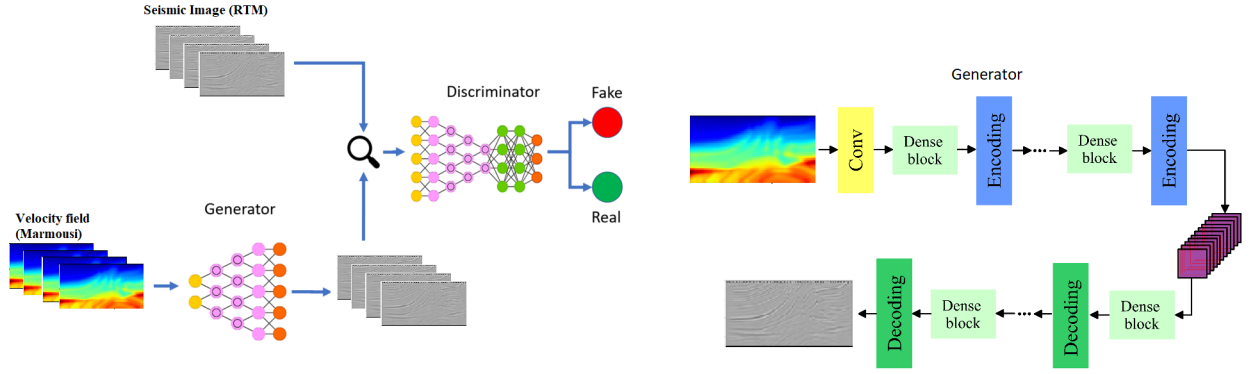


Figure 1. RTM deep convolutional encoder-decoder generative adversarial network architecture. Inputs: ensemble of velocity fields  $\mathbf{v} \in \mathbb{R}^N$ . Outputs: corresponding ensemble of image conditions  $\mathbf{I} \in \mathbb{R}^N$ .

methods [16]. We aim to construct and evaluate the performance of a generative adversarial network (GAN) [3] acting as a surrogate model for the RTM imaging under uncertainty. The main aspect of this approach is to train simultaneously two models. More specifically, a generative model that returns the probabilistic characterization of the data, and a discriminative model that estimates the probability that a sample came from the training data rather than a generative model. Here, we construct the generative model using as baseline the encoder-decoder architecture proposed by [4] and designed for problems cast as image-to-image regressions. Also, we build the discriminative model using the architecture proposed by [6].

Figure 1, provides the big picture of our end-to-end solution, depicting the main components of the encoder-decoder network. In our particular application, the input for the deep learning surrogate is the ensemble of heterogeneous velocity fields, and the outputs are the corresponding imaging conditions given by Eq. (3) for each sample of the velocity field ensemble. Inputs and outputs are high-dimensional spatial fields, and the surrogate modeling cast as a field-to-field regression. This approach is image-inspired. Then, it relies on connecting each pixel of the input field to an output pixel, where pixels correspond to grid points in the input computational mesh and output fields. The trained network maps the velocity  $\mathbf{v} \in \mathbb{R}^N$  into the IC field  $\mathbf{I} \in \mathbb{R}^N$ .

The surrogate  $\mathbf{g}$  is expressed formally in a compact notation as,

$$\hat{\mathbf{y}} = \mathbf{g}(\mathbf{x}; \mathbf{w}), \quad (4)$$

where  $\hat{\mathbf{y}}$  is the surrogate prediction (imaging condition,  $\mathbf{I}$ ) for an input  $\mathbf{x}$  (that is, a velocity field  $\mathbf{v}$ ), and vector  $\mathbf{w}$  contains the parameters of the neural network. Training the neural network means learning parameters  $\mathbf{w}$  using data from the set  $\mathcal{D} = \{(\mathbf{x}_i, \mathbf{y}_i)\}, i = 1 \dots N_{train}$  obtained from simulations with the RTM algorithm, where  $N_{train}$  is the number of samples in the training set. The stochastic gradient descent algorithm computes the unknown network elements for a given loss function [17]. The adversarial networks are trained jointly in two phases for each iteration. Firstly, the generative model is trained using the following loss function,

$$L_{ED} = \frac{1}{N_{train}} \sum_{i=1}^{N_{train}} \|\hat{\mathbf{y}}_i - \mathbf{y}_i\|_2^2 + \alpha L_G \quad (5)$$

where  $\hat{\mathbf{y}}_i = \mathbf{g}_i(\mathbf{x}_i, \mathbf{w})$ . Here,  $L_G$  measures the generative model ( $\cdot$ ) ability to fool the discriminative model  $\mathcal{D}(\cdot)$ , and it is given by

$$L_G = -\frac{1}{N_{train}} \sum_{i=1}^{N_{train}} \log \{D[G(\mathbf{y}_i)]\} \quad (6)$$

where  $\alpha$  represents a weight factor balancing between the two losses and a value of  $\alpha = 0.01$  is used. Also, the discriminative model is trained following the given loss function,

$$L_D = -\frac{1}{N_{train}} \sum_{i=1}^{N_{train}} \{\log [D(\mathbf{y}_i)] + \log [1 - D[G(\mathbf{y}_i)]]\}, \quad (7)$$

to distinguish the true sample from the fake sample produced by the generative model  $G(\mathbf{y}_i)$ .

## 4 Numerical experiments

Here, we present an example to demonstrate the ability of the deep convolutional encoder-decoder generative adversarial surrogate in replacing the original two-way wave equation RTM algorithm efficiently. The example deals with

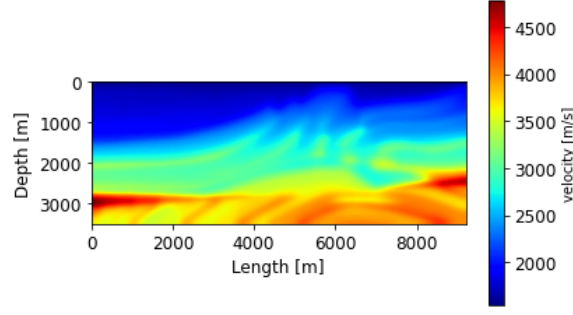


Figure 2. Marmousi Geologic model.

the velocity field of the Marmousi model [18]. The seismic source term considered in the present work is a Ricker-type wavelet [19]. More specifically, we use the Marmousi geologic model, as shown on Fig. 2. Also, the deep convolutional encoder-decoder generative adversarial is constructed using the open platform Tensorflow [20]. The Adam optimizer algorithm [21] is employed for parameter learning, considering a learning rate of  $1 \times 10^{-3}$  for the generative model and a learning rate of  $1 \times 10^{-4}$  for the discriminative model. We compute a total of 1200 samples by considering the velocity magnitude constant in the interior of each geological layer. Therefore, each velocity field in the ensemble has the form,

$$\mathbf{v} = \sum_{l=1}^{n_l} \bar{v}_l (1 + \sigma_l \xi_l) \mathbf{P}_l \quad (8)$$

where  $n_l$  is the number of geological layers,  $\bar{v}_l$  is the mean velocity within each geological layer,  $\sigma_l$  is the standard deviation, here assumed as 5%,  $\xi_l \sim \mathbb{U}[-1, 1]$  is a random variable following a uniform distribution, and  $\mathbf{P}_l$  is a  $N$ -dimensional vector containing 1 in the components corresponding to the  $l$ -th geological layer grid points and 0 otherwise. After that, we apply a moving harmonic average to  $\mathbf{v}$  with a window length of 20 grid points to mimic the velocity fields computed by parameter model building techniques like tomography or full-waveform inversion. To analyze the surrogate training convergence, out of 1200 samples computed with the original RTM model, we create three training datasets with 300, 600, and 900 samples each. We used the remaining  $N_{test} = 300$  samples to test the trained network.

Accuracy is measured using the distance between predictions with the surrogate model and those computed with the RTM original model. To evaluate the surrogate model quality, we consider the coefficient of determination ( $R^2$ -score) [22] and the root mean squared error ( $RMSE$ ) metrics,

$$R^2 = 1 - \frac{\sum_{i=1}^{N_{test}} \|\mathbf{y}_i - \hat{\mathbf{y}}_i\|_2^2}{\sum_{i=1}^{N_{test}} \|\mathbf{y}_i - \bar{\mathbf{y}}\|_2^2} \quad (9)$$

and

$$RMSE = \sqrt{\frac{1}{N_{train}} \sum_{i=1}^{N_{train}} \|\mathbf{y}_i - \hat{\mathbf{y}}_i\|_2^2}. \quad (10)$$

where  $\bar{\mathbf{y}} = \frac{1}{N_{test}} \sum_{i=1}^{N_{test}} \mathbf{y}_i$  is the mean of test samples. Also, we measure the structural similarity (SSIM) between the predicted seismic image and original RTM, and it is computed over local windows of the image [23]

$$SSIM(\mathbf{e}_l, \mathbf{e}_k) = \frac{2\mu_{\mathbf{e}_l}\mu_{\mathbf{e}_k} + c_1 2\sigma_{\mathbf{e}_l}\sigma_{\mathbf{e}_k} + c_2}{\mathbf{e}_l^2 + \mathbf{e}_k^2 + c_1\sigma_{\mathbf{e}_l}^2 + \sigma_{\mathbf{e}_k}^2 + c_2} \quad (11)$$

Also, we intend that the surrogate model returns not only good predictions of seismic images but also accurate estimations of quantities that characterize the uncertainties in such images. To measure the degree of uncertainty in the seismic images, we follow the approach proposed in [24]. In their approach, the degree of uncertainty is expressed by a confidence index that represents the pointwise normalized standard deviation, where low values represent high variabilities and high values the opposite. The confidence index is,

$$c(\mathbf{r}) = \frac{\sigma_{max} - \sigma(\mathbf{r})}{\sigma_{max} - \sigma_{min}}, \quad (12)$$

where  $c(\mathbf{r})$  is the confidence index at position  $\mathbf{r}$ , and  $\sigma_{min}$  and  $\sigma_{max}$  are the minimum and maximum field standard deviations, respectively. Another form of measuring the degree of uncertainty is the coefficient of variation, defined as the pointwise ratio between the standard deviation and the mean,

$$c_v(\mathbf{r}) = \frac{\sigma(\mathbf{r})}{\mu(\mathbf{r})} \quad (13)$$

where  $\mu(\mathbf{r})$  is the expected value at position  $\mathbf{r}$ .

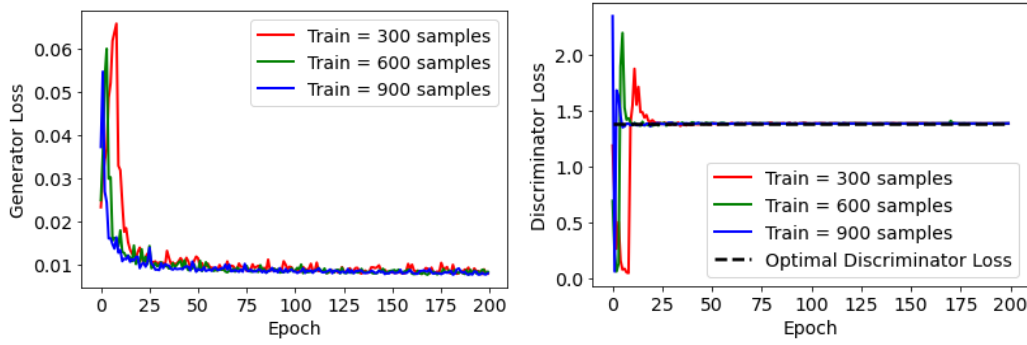


Figure 3. Generator and discriminator loss values with number of epochs.

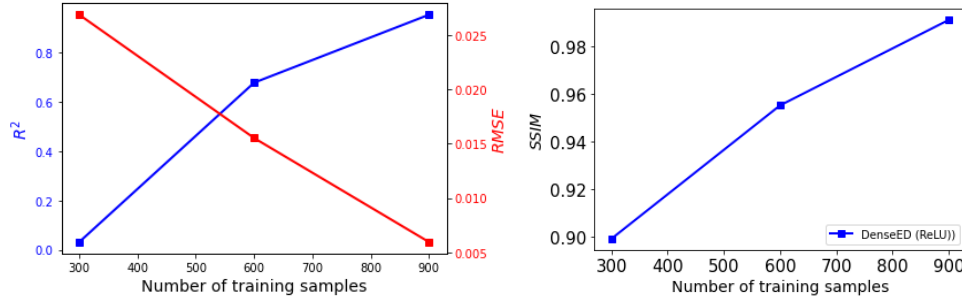


Figure 4. Accuracy metrics as a function of the number of training samples.

Figure 3 shows the decay of the generator loss (5) as a function of the number of epochs during the training process, for training data sets ranging from 300 to 900 samples. Note that the  $L_{ED}$  stabilizes after 50 epochs for all cases and that for smaller data sets, we see higher error values. Here, we also observe that the discriminator loss quickly converge to the optimal value [3]. Also, we calculate the coefficient of determination  $R^2$  to assess the accuracy of the machine learning model with the remaining 300 samples. We observe that the surrogate model accuracy is very satisfactory for the training with 900 samples, reaching  $R^2 \geq 0.95$ , as shown in Fig. 4. Furthermore, Fig. 4 also depicts the structural similarity (SSIM) between the predicted seismic image and the original RTM. Here is it worth remarking that the machine learning model accuracy is good even in small training data scenarios.

To further illustrate how the surrogate model approximates the predictions of the original model with good accuracy, Fig. 5 shows comparisons for three realizations chosen randomly from the test set. We observe that the surrogate presents good results, returning good predictions of the seismic image amplitudes. We also depict a comparison between the confidence index,  $c(\mathbf{r})$ , and the coefficient of variation,  $c_v(\mathbf{r})$ , predicted by the original and surrogate models with  $N_{testing} = 300$  testing samples, see Fig. 6. Besides, we introduce a discrete version of a  $L_2$  relative error between two spatial fields  $g$ , one computed with the RTM and the other by the surrogate as,

$$e_g^2 = \frac{1}{N} \sum_{i=1}^N \left( \frac{g_{RTM}^i - g_s^i}{g_{RTM}^i} \right)^2 \quad (14)$$

where the subscripts refer to how we compute the field  $g$ . This index is an average of the pointwise relative error for all  $N$  grid points. The visual resemblance of the images in Figure 5 is quantified using equation (14), leading to relative errors that stay below 2%. For the indexes in Figure 6, the relative errors computed with equation (14) are less than 2%.



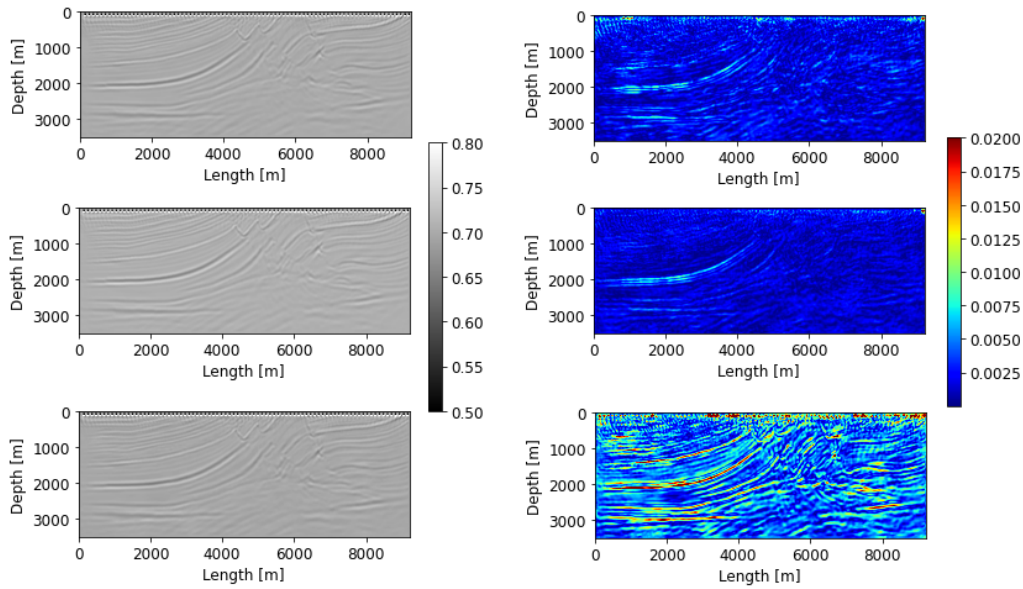


Figure 5. Randomly selected samples of seismic images in the test data set computed by the RTM (left) and the surrogate model (right). The relative errors in the image condition,  $e_g^2$ , are lower than 2%.

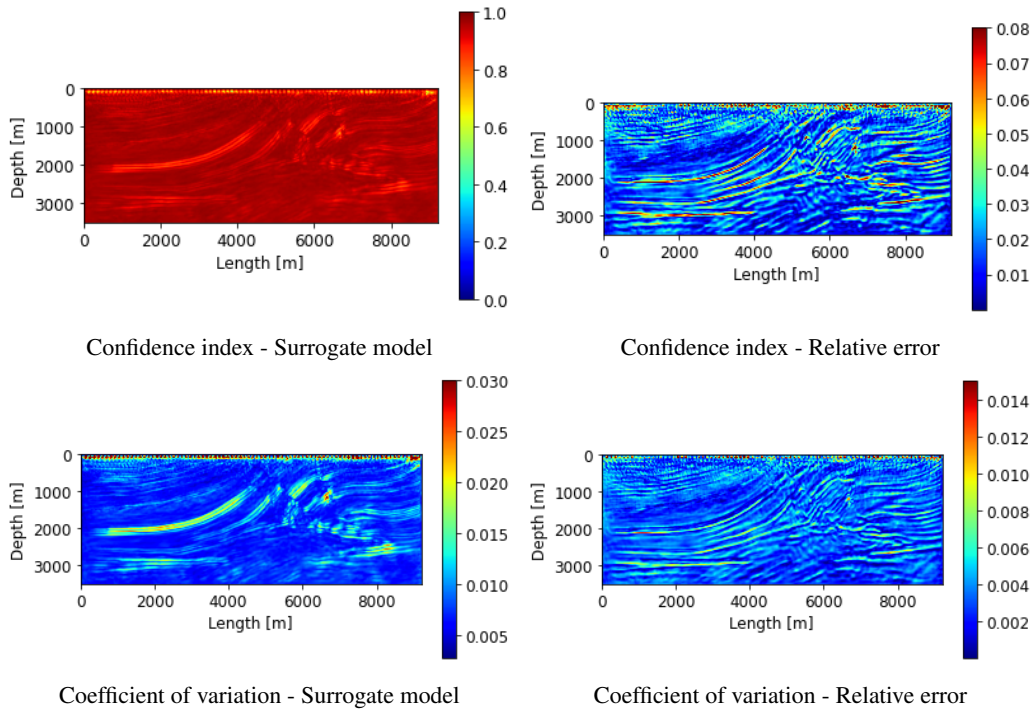


Figure 6. UQ indexes - confidence index,  $c(\mathbf{r})$ , and coefficient of variation,  $c_v(\mathbf{r})$  - surrogate model (left) and relative errors (right). The relative errors to the RTM model are lower than 2%.

## 5 Conclusions

We propose a deep convolutional encoder-decoder generative adversarial network architecture to replace the costly RTM technique on producing seismic images. This surrogate model builds a scalable image-to-image mapping, coping with the high dimensionality of both the heterogeneous velocity fields that serve as inputs and images outputs. Such surrogate has revealed to be very efficient in the context of UQ many-query tasks, as demonstrated by our numerical example. Also, we place our contribution in the emerging area of physics-informed machine learning, where the final model, in many different ways, blends two main components: often expensive computational models relying on first principles and phenomenological closure equations, and machine learning data-driven tools.

**Acknowledgements.** This study was financed in part by CAPES, Brasil Finance Code 001. This work is also partially supported by FAPERJ (grant E-26/203.018/2017), CNPq (grant 302489/2016-9), and Petrobras (grant 2017/00148-9).

**Authorship statement.** The authors hereby confirm that they are the sole liable persons responsible for the authorship of this work, and that all material that has been herein included as part of the present paper is either the property (and authorship) of the authors, or has the permission of the owners to be included here.

## References

- [1] Y. Chen, K. Gao, E. S. Davis, D. N. Sinha, C. Pantea, and L. Huang. Full-waveform inversion and least-squares reverse-time migration imaging of collimated ultrasonic-beam data for high-resolution wellbore integrity monitoring. *Applied Physics Letters*, vol. 113, n. 7, pp. 071903, 2018.
- [2] C. H. Barbosa, L. N. Kunstmann, R. M. Silva, C. D. Alves, B. S. Silva, M. Mattoso, F. A. Rochinha, A. L. Coutinho, and others. A workflow for seismic imaging with quantified uncertainty. *Computers & Geosciences*, vol. 145, pp. 104615, 2020.
- [3] I. J. Goodfellow, J. Pouget-Abadie, M. Mirza, B. Xu, D. Warde-Farley, S. Ozair, A. Courville, and Y. Bengio. Generative adversarial networks, 2014.
- [4] Y. Zhu and N. Zabaras. Bayesian deep convolutional encoder-decoder networks for surrogate modeling and uncertainty quantification. *Journal of Computational Physics*, vol. 366, pp. 415–447, 2018.
- [5] R. S. M. Freitas, C. H. S. Barbosa, G. M. Guerra, A. L. G. A. Coutinho, and F. A. Rochinha. An encoder-decoder deep surrogate for reverse time migration in seismic imaging under uncertainty. *Comput Geosci*, vol. 25, pp. 1229–1250, 2021.
- [6] A. Radford, L. Metz, and S. Chintala. Unsupervised representation learning with deep convolutional generative adversarial networks, 2016.
- [7] H.-W. Zhou, H. Hu, Z. Zou, Y. Wo, and O. Youn. Reverse time migration: A prospect of seismic imaging methodology. *Earth-Science Reviews*, vol. 179, pp. 207–227, 2018.
- [8] G. T. Schuster. *Seismic inversion*. Society Exploration Geophysicists, 1 edition, 2017.
- [9] C. Cerjan, D. Kosloff, R. Kosloff, and M. Reshef. A nonreflecting boundary condition for discrete acoustic and elastic wave equations. *Geophysics*, vol. 50, n. 4, pp. 705–708, 1985.
- [10] W.-F. Chang and G. A. McMechan. Reverse-time migration of offset vertical seismic profiling data using the excitation-time imaging condition. *Geophysics*, vol. 51, n. 1, pp. 67–84, 1986.
- [11] B. D. Nguyen and G. A. McMechan. Excitation amplitude imaging condition for prestack reverse-time migration. *Geophysics*, vol. 78, n. 1, pp. S37–S46, 2013.
- [12] P. Sava and S. Fomel. Time-shift imaging condition in seismic migration. *Geophysics*, vol. 71, n. 6, pp. S209–S217, 2006.
- [13] B. Wang, C. Mason, M. Guo, K. Yoon, J. Cai, J. Ji, and Z. Li. Subsalt velocity update and composite imaging using reverse-time-migration based delayed-imaging-time scan. *Geophysics*, vol. 74, n. 6, pp. WCA159–WCA166, 2009.
- [14] F. Liu, G. Zhang, S. A. Morton, and J. P. Leveille. An effective imaging condition for reverse-time migration using wavefield decomposition. *Geophysics*, vol. 76, n. 1, pp. S29–S39, 2011.
- [15] S. Chattopadhyay and G. A. McMechan. Imaging conditions for prestack reverse-time migration. *Geophysics*, vol. 73, n. 3, pp. S81–S89, 2008.
- [16] I. Goodfellow, Y. Bengio, and A. Courville. *Deep Learning*. MIT Press. <http://www.deeplearningbook.org>, 2016.
- [17] F. Chollet. *Deep Learning with Python*, volume 1. Manning Publications Company, 2017.
- [18] R. Versteeg. The marmousi experience: Velocity model determination on a synthetic complex data set. *The Leading Edge*, vol. 13, n. 9, pp. 927–936, 1994.
- [19] N. Ricker. The form and laws of propagation of seismic wavelets. *Geophysics*, vol. 18, n. 1, pp. 10–40, 1953.
- [20] M. Abadi, A. Agarwal, P. Barham, E. Brevdo, Z. Chen, C. Citro, G. S. Corrado, A. Davis, J. Dean, M. Devin, S. Ghemawat, I. Goodfellow, A. Harp, G. Irving, M. Isard, Y. Jia, R. Jozefowicz, L. Kaiser, M. Kudlur, J. Levenberg, D. Mané, R. Monga, S. Moore, D. Murray, C. Olah, M. Schuster, J. Shlens, B. Steiner, I. Sutskever, K. Talwar, P. Tucker, V. Vanhoucke, V. Vasudevan, F. Viégas, O. Vinyals, P. Warden, M. Wattenberg, M. Wicke, Y. Yu, and X. Zheng. TensorFlow: Large-scale machine learning on heterogeneous systems. Software available from tensorflow.org, 2015.
- [21] P. K. Diederik and B. Jimmy. Adam: A method for stochastic optimization. preprint arXiv:1412.6980, 2014.
- [22] S. Weisberg. *Applied Linear Regression*. John Wiley & Sons, Inc., 2005.
- [23] Z. Wang, A. Bovik, H. Sheikh, and E. Simoncelli. Image quality assessment: from error visibility to structural similarity. *IEEE Transactions on Image Processing*, vol. 13, n. 4, pp. 600–612, 2004.
- [24] Y. Li and J. Sun. 3d magnetization inversion using fuzzy c-means clustering with application to geology differentiation. *Geophysics*, vol. 81, pp. J61–J78, 2016.