

Modèle génératif pour l'interprétation géologique d'images tomographiques en Géosciences

Théo SANTOS^{1,2} Ferréol SOULEZ¹ Thomas BODIN²

¹Univ Lyon, Univ Lyon1, Ens de Lyon, CNRS, Centre de Recherche Astrophysique de Lyon UMR5574, F-69230, Saint-Genis-Laval, France

²Univ Lyon, Université Lyon 1, ENS de Lyon, CNRS, Laboratoire de Géologie de Lyon - Terre Planètes Environnement, UMR 5276, F-69622, Villeurbanne, France

Résumé – La résolution d'un problème inverse en sciences physiques peut passer par l'exploration de l'ensemble des modèles possibles. Nous proposons ici d'entraîner un réseau de neurones à reproduire des simulations physiques. Ce réseau est ensuite utilisé pour échantillonner l'ensemble des modèles expliquant les données observées. Nous le mettons en œuvre pour ajuster des modèles géologiques de manteau supérieur appelés *gâteaux marbrés* sur des données tomographiques. Nous montrons que les images générées sont statistiquement proches des modèles théoriques et nous obtenons le résultat de l'inversion sous forme d'une distribution.

Abstract – Solving an inverse problem in physical sciences can involve exploring all possible models. We propose here to train a neural network to reproduce physical simulations. This network is then used to sample the set of models explaining the observed data. We use this approach to fit geological models of upper mantle called *marble cakes* on tomographic data. We show that the generated images are statistically close to the theoretical models and we get the result of the inversion as a distribution.

1 Introduction

La tomographie sismique permet d'imager la structure interne de la terre à partir de l'analyse des ondes sismiques produites par des tremblements de terre. Cependant, le champ d'onde observé est souvent limité en fréquence. Les échelles plus petites que la longueur d'onde minimum du champ d'onde observé ne peuvent donc pas être reconstruites, et les images tomographiques ne représentent qu'un milieu effectif lisse, tel que vu à de grandes longueurs d'onde.

Pourtant, les modèles géologiques qu'on cherche à retrouver sont bien plus complexes et multi-échelles [1] que les images reconstruites en tomographie classique. Cela est dû à plusieurs causes. Tout d'abord, les observations sismiques sont à bande limitée et peu sensibles aux petites échelles. Ensuite, il est difficile de définir des modèles géologiques paramétriques dont on estimerait les paramètres d'après les données dans une approche inverse. Lorsqu'ils existent, ces modèles géologiques sont soit trop simplistes soit très coûteux numériquement. De plus ces modèles sont très mal conditionnés, c'est-à-dire qu'ils peuvent beaucoup changer pour des faibles variations des paramètres de simulation.

Dans cette étude, nous proposons une approche d'apprentissage pour interpréter des images tomographiques lisses en terme de modèles géologiques à plus petites échelles. La méthode que nous employons consiste à faire apprendre *a priori* sur les modèles géologiques par un réseau de neurones génératifs. Une fois entraîné, le réseau est capable de générer à la volée, à partir d'un vecteur de faible dimension (vecteur latent), un nouveau modèle géologique statistiquement similaire aux modèles d'entraînement. Cette génération a un faible coût numérique, et assure d'obtenir un résultat correspondant à *a priori* géologique. A l'aide de ce réseau, nous échantillons, par Markov-chain Monte-Carlo (McMC) [9], l'ensemble

des modèles possible qui expliquent les observations tomographiques.

Une approche similaire a été étudiée par Bloem *et al.* [5] pour la tomographie de systèmes fluviaux en tresse et de paraséquences marines.

2 Modèle direct

2.1 Modèle géologique

Le manteau supérieur est composé principalement de couches successives d'harzburgite et de basalte. Le mélange de ces couches par les mouvements convectifs crée des structures très lamellées, multi-échelles, anisotropes ressemblant en coupe 2D à des *gâteaux marbrés* [1] (voir figure 1). Alder *et al.* [4] proposent une méthode de simulation géodynamique numérique, permettant d'obtenir des modèles de ces structures géologiques harzburgite/basalte. Les modèles obtenus sont statistiquement proches des structures du manteau supérieur, dont la Densité Spectrale de Puissance (DSP) attendue en théorie est proportionnelle à $1/k$ (k étant la fréquence spatiale) [3]. Nous utilisons les résultats de ces simulations comme modèles de référence pour les modèles géologiques.

2.2 Problème direct : homogénéisation

Pour une coupe 2D du manteau, une reconstruction tomographique fournit 9 images 2D des 3×3 coefficients du tenseur élastique du milieu. Ce ne sont pas des images fidèles des structures du manteau mais uniquement des images du milieu effectif. Ce milieu effectif est un milieu lisse équivalent pour les grande longueurs d'ondes. Le calcul de ce milieu effectif à partir du tenseur élastique des milieux harzburgite et basalte

est appelé homogénéisation [2] [8]. Cette opération, caractérisée par une certaine longueur d’onde n est pas linéaire et préserve une partie de l’information des petites structures via l’anisotropie ; les lamellages des structures à petites échelles étant responsables de l’anisotropie du milieu effectif.

Dans cette preuve de concept, on suppose l’étape de reconstruction tomographique comme parfaitement réalisée. Cette hypothèse n’est pas toujours vérifiée en pratique mais est suffisante pour cette preuve de concept. Nous supposons donc que l’étape d’homogénéisation est équivalente à l’opérateur tomographique, liant l’image de la structure du manteau aux images du tenseur élastique. Nous considérons alors les images tomographiques $\mathbf{d} \in \mathbb{R}^{N \times 9}$ comme les observations avec N le nombre de pixels. Ainsi, notre problème est de reconstruire l’image des *gâteaux marbrés* $\mathbf{m} \in \mathbb{R}^N$ d’après les reconstructions tomographiques. Le modèle direct de notre problème est uniquement l’homogénéisation $\mathcal{H} : \mathbb{R}^N \rightarrow \mathbb{R}^{N \times 9}$ décrite par :

$$\mathbf{d} = \mathcal{H}(\mathbf{m}) \quad (1)$$

Dans les figures de ce papier, pour un souci de lisibilité, nous ne montrerons pas à chaque fois la totalité du tenseur élastique mais uniquement l’image du premier coefficient c_{11} .

3 Génération de *gâteaux marbrés*

3.1 Réseaux génératifs adversariaux

Pour explorer l’espace des *gâteaux marbrés*, nous utilisons un modèle génératif par réseaux de neurones. Plus précisément, nous utilisons un Generative Adversarial Network [6], capable de générer des images semblables aux images d’entraînement à partir d’un vecteur d’un espace de petite dimension (l’espace latent). Le modèle de GAN que nous utilisons est le WGAN avec *gradient penalty* [7], que nous entraînons avec un espace latent de dimension 30. Cette dimension a été choisie comme compromis entre la fidélité des images générées et la nécessité d’un nombre de paramètres réduit pour l’échantillonnage.

Le tableau 1 présente la structure du GAN ; chaque couche est une couche de convolution transposée avec un noyau de taille 4×4 , une fonction d’activation ReLU, sauf la dernière couche qui a pour fonction d’activation une tanh. Le générateur $\mathcal{G} : \mathbb{R}^{30} \rightarrow \mathbb{R}^N$ génère des images \mathbf{m} de $N = 128 \times 128$ pixels à partir d’un vecteur latent \mathbf{z} de dimension 30 :

$$\mathbf{m} = \mathcal{G}(\mathbf{z}) \quad (2)$$

| Taille d’entrée | Taille de sortie | Stride | Padding |
|---------------------------|---------------------------|--------|---------|
| $1 \times 1 \times 30$ | $4 \times 4 \times 2048$ | 1 | 0 |
| $4 \times 4 \times 2048$ | $8 \times 8 \times 1024$ | 2 | 1 |
| $8 \times 8 \times 1024$ | $16 \times 16 \times 512$ | 2 | 1 |
| $16 \times 16 \times 512$ | $32 \times 32 \times 256$ | 2 | 1 |
| $32 \times 32 \times 256$ | $64 \times 64 \times 128$ | 2 | 1 |
| $64 \times 64 \times 128$ | $128 \times 128 \times 1$ | 2 | 1 |

TABLE 1 : Structure du GAN

Le GAN a été entraîné sur un jeu de données de 20 000 images durant ≈ 750 époques.

3.2 Qualité des images générées

Qualitativement, la morphologie des structures générées sont similaires au jeu de données d’entraînement. Toutefois, on constate quelques différences. Par exemple, la connexité de ces structures est en général moins importante dans les images générées.

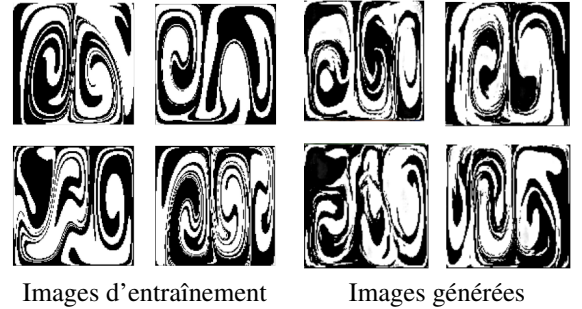


FIGURE 1 : Exemple d’images d’entraînement et d’images générées par le GAN

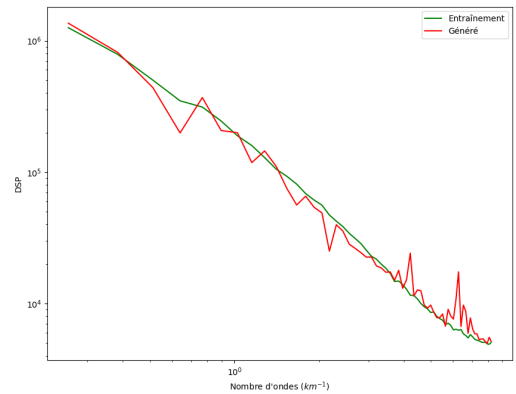


FIGURE 2 : Composantes horizontales de la DSP des images d’entraînement et des images générées

Pour évaluer la qualité du GAN quantitativement, nous nous sommes intéressés aux densités spectrales de puissance (DSP), afin de s’assurer que les structures statistiques soient les mêmes. Nous avons calculé la DSP moyenne de 2000 images d’entraînement et de 2000 images générées aléatoirement par le GAN. Puis nous affichons leurs composantes horizontales à fréquence nulle, visible en figure 2. Sur cette figure, on voit que le GAN arrive bien à reproduire les statistiques des images d’entraînement. Les DSP des images générées et d’apprentissage sont similaires, avec la même décroissance en $1/k$ comme attendue en théorie. On observe toutefois des biais pour certaines fréquences spatiales. Leur lien avec les échelles du GAN reste à investiguer.

4 Formulation bayésienne

Etant donné la faible résolution des images tomographiques, il peut exister un grand nombre *gâteaux marbrés* les expliquant. Nous cherchons à caractériser cet ensemble de solutions. Nous utilisons pour cela une formulation bayésienne. Nous cherchons donc à déterminer la distribution des vecteurs latents \mathbf{z}

et des images $\mathcal{G}(z)$ décrivant les observations \mathbf{d} données par :

$$\mathbf{d} = \mathcal{H}(\mathcal{G}(z)) + \epsilon, \quad (3)$$

où $\epsilon \in \mathbb{R}^{N \times 9}$ est un terme d'erreur supposée gaussienne centrée. Cette erreur prend en compte à la fois le bruit de la reconstruction tomographique et l'erreur de représentation du GAN. L'erreur de représentation traduit la difficulté du GAN à reproduire exactement la totalité des images possibles de simulation.

Pour chaque vecteur latent z , on cherche à estimer la densité de probabilité *a posteriori* $\Pr(z | \mathbf{d})$ du modèle sachant les données. D'après le théorème de Bayes :

$$\Pr(z | \mathbf{d}) = \frac{\Pr(z) \Pr(\mathbf{d} | z)}{\int \Pr(\mathbf{d} | z) \Pr(z) dz} \quad (4)$$

ainsi :

$$\Pr(z | \mathbf{d}) \propto \Pr(z) \Pr(\mathbf{d} | z) \quad (5)$$

Le terme de vraisemblance $\Pr(\mathbf{d} | z)$ décrit l'adéquation du modèle aux données tomographiques. Le terme d'*a priori* $\Pr(z)$ est la distribution des variables latentes z tel que $\mathcal{G}(z)$ décrit la distribution des images du jeu d'entraînement. Par construction, cela correspond à la distribution normale centrée $\mathcal{N}(0, \mathbf{I})$ des vecteurs latents utilisée pendant l'entraînement :

$$\Pr(z) \propto \exp\left(-\frac{1}{2} \|z\|_2^2\right) \quad (6)$$

Dans l'équation 3, on suppose la distribution des erreurs ϵ normale centrée. On a donc :

$$\Pr(\mathbf{d} | z) \propto \exp\left(-\frac{1}{2\sigma^2} \|\mathbf{d} - \mathcal{H}(\mathcal{G}(z))\|_2^2\right) \quad (7)$$

Comme ϵ représente à la fois le bruit des données et les erreurs de représentation du GAN, on décompose la variance de ces erreurs : $\sigma^2 = \sigma_b^2 + \sigma_{\text{GAN}}^2$. Le bruit des données σ_b^2 est connu dans nos simulations. Nous estimons l'erreur de représentation σ_{GAN} en recherchant les vecteurs latents z_i dont les *gâteaux marbrés* générés sont les plus proches d'images de *gâteaux marbrés* simulés numériquement \mathbf{m}_i :

$$z_i = \arg \min_z \{\|\mathcal{G}(z) - \mathbf{m}_i\|_2^2\} \quad (8)$$

Nous propageons les \mathbf{m}_i et $\mathcal{G}(z_i)$ à travers l'homogénéisation, et nous pouvons en estimer la variance empirique σ_{GAN} à partir des erreurs moyennes entre les $\mathcal{H}(\mathcal{G}(z_i))$ et les $\mathcal{H}(\mathbf{m}_i)$ sur $M = 100$ échantillons.

$$\sigma_{\text{GAN}}^2 = \frac{1}{M} \sum_{i=1}^M \|\mathcal{H}(\mathbf{m}_i) - \mathcal{H}(\mathcal{G}(z_i))\|_2^2 \quad (9)$$

Ainsi, pour chaque vecteur latent z , nous pouvons estimer sa probabilité $\Pr(z | \mathbf{d})$. Nous l'estimons en échantillonnant l'espace latent par la méthode Markov-chain Monte Carlo (MCMC)[9].

5 Résultats

Pour cette expérience, un gâteau marbré a été homogénéisé puis bruité (SNR = 20 dB). Le résultat obtenu (voir figure 3)

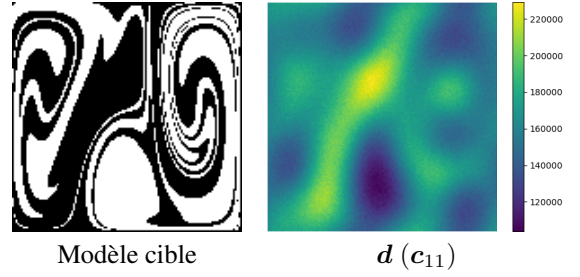


FIGURE 3 : Modèle cible à retrouver et observations (un seul coefficient du tenseur)

est l'observation \mathbf{d} . Le problème inverse que nous considérons consiste à déterminer la distribution des *gâteaux marbrés* expliquant les observations \mathbf{d} .

Un MCMC, avec le GAN utilisé comme générateur de modèles, a été lancé sur 1 600 000 itérations. Le résultat obtenu est une distribution sur l'espace latent à 30 dimensions. Un extrait des marginales 1D et 2D de cette distribution est visible en figure 4. La distribution obtenue a une variance plus réduite que l'*a priori*, ce qui confirme l'apport d'information par les observations tomographiques. Elle semble n'avoir que 2 modes principaux, visibles sur plusieurs dimensions de l'espace latent.

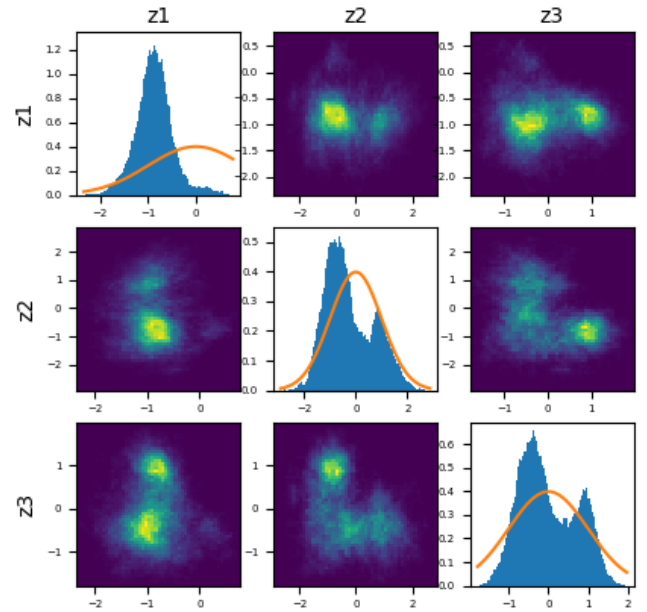


FIGURE 4 : Résultats du MCMC - distribution dans l'espace latent (extrait). La probabilité *a priori* est également affichée en orange sur les marginales 1D.

Pour interpréter cette distribution, nous affichons l'image du *maximum a posteriori* $\mathcal{G}(z_{\text{MAP}})$, et son modèle tomographique $\mathcal{H}(\mathcal{G}(z_{\text{MAP}}))$, l'espérance des images échantillonnées $\mu = \langle \mathcal{G}(z) \rangle_z$, et l'image générée par la médiane de la distribution $\mathcal{G}(z_{\text{med}})$.

On constate que le modèle cible est retrouvé par l'inversion pour les grandes échelles et la structure globale (i.e. les données simulées du MAP sont très proches de \mathbf{d}). Les petites échelles ne sont pas retrouvées avec exactitude. Nous retrouvons les deux modes observés dans la distribution des vecteurs latents par un algorithme K-means à deux clusters. Les images

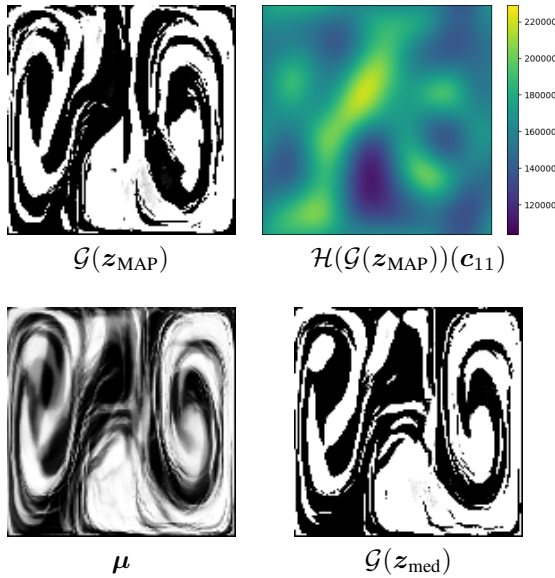


FIGURE 5 : Images caractéristiques de la distribution

de ces deux modes sont visibles sur la figure 6. Leurs empreintes sont perceptible sur l'image μ , dont les zones grises correspondent à l'alternance, durant l'échantillonnage, entre les deux modes. C'est notamment le cas dans la partie en haut à gauche de l'image.

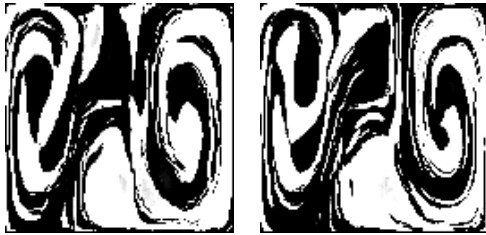


FIGURE 6 : Images des deux modes principaux de la distribution

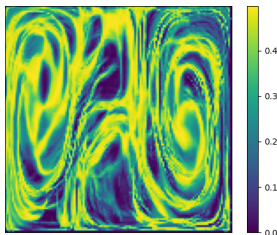


FIGURE 7 : Écart-type par pixel

La figure de l'écart-type par pixel (figure 7) n'est pas homogène. Les zones de l'image où il est élevé sont également des zones où les modèles retrouvés ne correspondent pas au modèle cible. Nous pouvons utiliser cette image écart-type comme un estimateur des incertitudes.

6 Discussion

La méthode que nous proposons est d'intégrer l'*a priori* géologique dans l'analyse d'images tomographiques à grandes échelles. L'ensemble des images géologiques à petites échelles est exploré grâce à un GAN et une approche MCMC.

Nous avons montré qu'il est possible de générer avec un GAN des modèles de *gâteaux marbrés* statistiquement proches des simulations numériques, et reproduisant le spectre attendu. Nous avons également montré qu'il est possible d'utiliser ce GAN pour échantillonner l'ensemble des modèles géologiques expliquant des images tomographiques. Grâce à une formulation bayésienne, la solution est une distribution de probabilité dans l'espace latent qui peut être exploité de différentes manières (incertitudes, modes).

Il reste quelques points qui peuvent être améliorés. Tout d'abord, le GAN présente des erreurs de représentation et les modèles générés ont moins de variabilité que les images d'entraînement. Le GAN n'est pas capable de reproduire certaines structures de gâteau marbré (par exemple, un gâteau marbré à 4 cellules) pourtant présentes dans le jeu d'entraînement. Nous avons plusieurs pistes (architecture du GAN, taille de l'espace latent, taille et distribution du jeu d'entraînement...) d'amélioration concernant ce point. Un autre enjeu est la quantification de ces erreurs de représentation. Elle permettrait de prendre en compte ces erreurs dans l'inversion, et les propager dans la solution probabiliste.

Une fois cette approche maîtrisée, cela ouvrira des perspectives pour l'appliquer à d'autres problèmes inverses, dans de nombreux domaines de la physique.¹

Références

- [1] C ALLÈGRE et D TURCOTTE : Implications of a two-component marble-cake mantle. *Nature*, 323(6084):123–127, 1986.
- [2] G E BACKUS : Long-wave elastic anisotropy produced by horizontal layering. *Journal of Geophysical Research*, 67(11):4427–4440, 1962.
- [3] G K BATCHELOR : Small-scale variation of convected quantities like temperature in turbulent fluid part 1. general discussion and the case of small conductivity. *Journal of fluid mechanics*, 5(1):113–133, 1959.
- [4] C Alder *et al.* : Quantifying seismic anisotropy induced by small-scale chemical heterogeneities. *Geophysical Journal International*, 211(3):1585–1600, 09 2017.
- [5] H Bloem *et al.* : Introducing conceptual geological information into bayesian tomographic imaging, 2022.
- [6] I Goodfellow *et al.* : Generative adversarial nets, in 'advances in neural information processing systems 27', curran associates, 2014.
- [7] I Gulrajani *et al.* : Improved training of wasserstein gans. *In Advances in Neural Information Processing Systems*, volume 30. Curran Associates, Inc., 2017.
- [8] Y Capdeville *et al.* : Fast fourier homogenization for elastic wave propagation in complex media. *Wave Motion*, 54:170–186, 2015.
- [9] W. K. HASTINGS : Monte Carlo sampling methods using Markov chains and their applications. *Biometrika*, 57(1): 97–109, 04 1970.

¹Ce travail a été soutenu par l'Action Spécifique Haute Résolution Angulaire (ASHRA) du CNRS/INSU co-financée par le CNES.