

# Un réseau thermodynamique récurrent (RTAN) avec encodage de Maxwell pour modéliser la réponse de matériaux viscoélastiques

N. Pistenon<sup>1</sup>, P. Kerfriden<sup>1</sup>, J.L. Bouvard<sup>2</sup>, D. Pino Muñoz<sup>2</sup>, S. Cantournet<sup>1</sup>

✉ Nicolas Pistenon [nicolas.pistenon@minesparis.psl.eu](mailto:nicolas.pistenon@minesparis.psl.eu)

<sup>1</sup> Mines Paris, Université PSL, Centre des Matériaux (MAT), UMR7633 CNRS, 91003 Evry, France

<sup>2</sup> Mines Paris, Université PSL, Centre de Mise en Forme des Matériaux (CEMEF), UMR7635 CNRS, 06904 Sophia Antipolis, France

**Résumé** — Une méthodologie d'apprentissage automatique est mise en place pour avoir un modèle non linéaire de loi de comportement dépendant de l'histoire. Elle repose sur deux biais inductifs mécaniques : le premier étant un modèle de Maxwell Généralisé linéaire utilisé comme couche cachée fournissant des variables cachées d'histoire. Le deuxième biais repose sur la formulation d'un réseau de neurones récurrent thermodynamique. Ce dernier montre des résultats prometteurs avec un bon apprentissage, un faible nombre de données et une bonne capacité à généraliser à d'autres trajets de chargements.

**Mots clés** — Apprentissage profond, série temporelle, réseau de neurones récurrents thermodynamique.

## 1 Introduction

Les réseaux de neurones sont de plus en plus utilisés pour concevoir des lois de comportement en mécanique [1]. En effet, les modèles micro-mécaniques qui fournissent une réponse haute fidélité pour des volumes représentatifs peuvent avoir un coût numérique prohibitifs. Les réseaux de neurones peuvent alors être utilisés comme substitut plus rapide. Alternativement, les réseaux de neurones peuvent être utilisés pour modéliser un comportement mécanique où les approches phénoménologiques échouent, c'est-à-dire lorsqu'une forme analytique de la loi de comportement n'est pas trouvée pour représenter une grande diversité de cas de chargements du matériau. C'est par exemple le cas pour le comportement viscoélastique de polymères qui reste aujourd'hui difficile à décrire précisément. Les approches phénoménologiques développées actuellement ne permettent pas de modéliser le comportement non linéaire du matériau en petite déformation comme par exemple l'effet Payne pour des élastomères renforcés [4, 13].

Dans le cadre de l'apprentissage automatique des lois de comportement dépendant de l'histoire du chargement, l'état de l'art ne montrent pas de méthodologies ayant des capacités interpolatoires et extrapolatoires consistantes avec les capacités expérimentales actuelles. Des réflexions sur des approches à petites données sont menées. Il s'agit alors d'introduire des biais inductifs dans les méthodologies d'apprentissage automatique. Les connaissances mécaniques sont utilisées dans la littérature pour introduire des biais inductifs, comme par exemple en ajoutant des termes de régularisation [3, 14], ou encore en introduisant ces connaissances mécaniques dans la structure même des réseaux de neurones [11].

Dans cette étude, une nouvelle méthodologie de réseau de neurones est présentée. L'approche multi-fidélité développée utilise différentes connaissances mécaniques actuelles comme biais inductifs structurels. Les connaissances mécaniques actuelles sont utilisées d'une part pour introduire des variables cachées au réseau de neurones. Et d'autre part, un réseau thermodynamique récurrent est formulé en introduisant la contrainte comme dérivée d'un potentiel.

L'approche multifidélité est présentée succinctement dans la section 2. Le réseau thermodynamique récurrent est plus particulièrement développé dans la section 3. Le cadre thermodynamique et la formulation des réseaux de neurones récurrents y sont tout d'abord posés. Puis notre formulation est confrontée à de données synthétiques non-linéaires en petites déformations. Les résultats sont validés par des tests d'interpolations au sein de l'espace paramétrique d'entraînement et des tests d'extrapolations à d'autres cas de chargements.

## 2 Approche multifidélité

Notre modèle combine un modèle de Maxwell Généralisé linéaire avec un réseau de neurones récurrents thermodynamique comme illustré sur la figure 1. Le modèle de Maxwell Généralisé est classiquement utilisé pour modéliser le domaine linéaire de matériaux viscoélastiques. Ce couplage a un double impact. Tout d'abord, la connexion résiduelle fait de la partie réseau de neurones une correction de la réponse mécanique du modèle de Maxwell. De plus, le modèle de Maxwell sert de couche d'encodage permettant d'introduire des variables cachées à la partie réseaux de neurones. Le réseau de neurones récurrents n'encode pas seul l'histoire du matériau. Ainsi, le modèle de Maxwell Généralisé fournit deux variables explicatives supplémentaires (une donnée de contrainte totale et la variable interne visqueuse du modèle de Maxwell) à la partie réseau de neurones comme illustré sur la figure 1. La description en déformation en est ainsi enrichie.

Les deux parties du modèle sont entraînées simultanément avec les algorithmes classiques de rétro-propagation du gradient. Ceci permet de déterminer les paramètres du modèle de Maxwell Généralisé et du réseau de neurones récurrents thermodynamique.

Dans toute la suite, une attention particulière est portée sur le réseau de neurones thermodynamique.

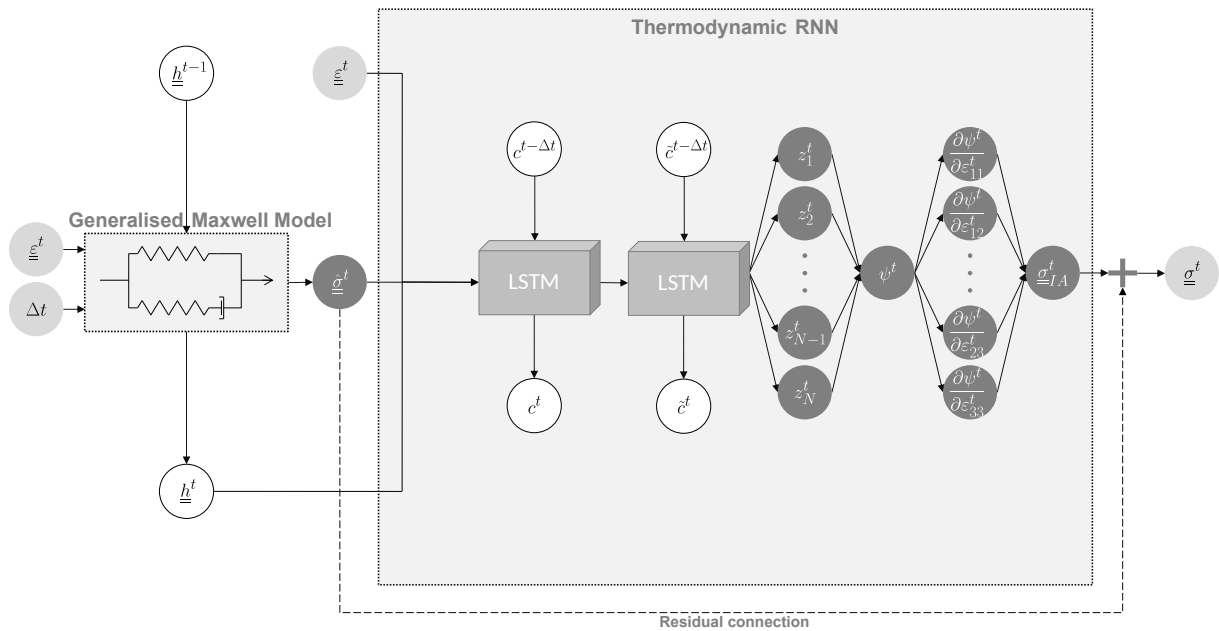


FIGURE 1 – Schéma du modèle couplé modèle de Maxwell Généralisé linéaire et réseau de neurones récurrents thermodynamiques. RNN : *Recurrent Neural Network*, LSTM : *Long Short-Term Memory* [8]

## 3 Réseau thermodynamique récurrent

Pour un matériau viscoélastique et de manière plus général, pour un matériau dont la réponse mécanique dépend du chemin suivi, la réponse en contrainte dépend de l'ensemble du chemin temporel de la déformation. Ceci correspond à rechercher une fonction  $f$  qui dépend de la déformation actuelle  $\underline{\underline{\epsilon}}^t$  à l'instant  $t$  et de son histoire :

$$\underline{\underline{\sigma}}^t = f\left(\left\{\underline{\underline{\epsilon}}^{t'}; t' \leq t\right\}\right) \quad (1)$$

Traditionnellement en mécanique, le principe de l'état local introduit des variables internes  $\mathbf{V}_k$  qui permettent de représenter l'histoire du matériau en un nombre fini de variables :

$$\underline{\underline{\sigma}}^t = f\left(\underline{\underline{\epsilon}}^t, \mathbf{V}_k\right) \quad (2)$$

$\mathbf{V}_k$  désigne un ensemble de variables internes de natures différentes (scalaires, vecteurs, tenseurs) en fonction de la description,  $\mathbf{V}_k = \{V_i; \forall i \in \llbracket 1, k \rrbracket\}$ .

Cette méthodologie de définition des variables internes peut être rapprochée des réseaux de neurones récurrents. MEMOIRE CACHEES DES RNNs, voir section 3.1

### 3.1 Mémoires cachées des réseau de neurones récurrents

Pour des variables temporelles où le principe de causalité est un biais inductif primordial, les réseaux de neurones récurrents (RNN) sont des réseaux de neurones qui ont la spécificité d'avoir des mémoires (ou états) cachées. Ces mémoires cachées  $c^t$  représentent l'histoire temporelle de la variable d'entrée  $(x^{t'})_{t' < t}$  pour expliquer la variable de sortie  $y^t$ . En désignant par  $NN_\theta$  un réseau de neurones, une cellule d'un RNN peut simplement s'écrire sous la forme :

$$\begin{aligned} y^t &= NN_\theta(x^t, c^{t-1}) \\ c^{t-1} &= \tilde{NN}_\theta(x^{t-1}, c^{t-2}) = \tilde{NN}_\theta(\{x^{t'}; t' < t\}) \end{aligned} \quad (3)$$

Un RNN est composé d'un grand nombre de ces cellules qui sont mises en parallèles et en séries. Différentes versions de RNN existent dans la littérature comme les Gated Recurrent Unit (GRU) [5] ou encore les Long Short-Term Memory (LSTM) [8]. Dans cette étude, les LSTM ont été utilisés.

Ces mémoires cachées peuvent être vues comme des variables internes d'un modèle de mécanique. Ainsi, pour les matériaux dépendant de l'histoire, les RNN sont naturellement introduits pour les modéliser [12, 7, 9, 15]. Maia et al. [10] reformulent cette structure récurrente en introduisant des considérations physiques. Bonatti et al. [2] modifient la structure classique des LSTM pour décroître le nombre de mémoires cachées tout en conservant une structure profonde.

Dans cette étude, ces réseaux de neurones récurrents sont utilisés dans un cadre thermodynamique qui est présenté dans la prochaine section.

### 3.2 Réseau thermodynamique récurrent pour les comportements non-linéaires

La déformation est supposée se décomposer de façon additive en une déformation thermoélastique  $\underline{\underline{\varepsilon}}^e$  et une déformation anélastique  $\underline{\underline{\varepsilon}}^{an}$  :

$$\underline{\underline{\varepsilon}} = \underline{\underline{\varepsilon}}^e + \underline{\underline{\varepsilon}}^{an} \quad (4)$$

La déformation anélastique correspond pour les matériaux plastiques aux déformations plastiques  $\underline{\underline{\varepsilon}}^p$ , ou pour les matériaux viscoélastiques aux déformations visqueuses  $\underline{\underline{\varepsilon}}^v$ .

La première loi de la thermodynamique exprime la conservation de l'énergie en petites déformations comme

$$\rho \dot{e} = \underline{\underline{\sigma}} : \underline{\underline{\dot{\varepsilon}}} + r - \text{div } \underline{\underline{q}} \quad (5)$$

avec  $e$  l'énergie interne spécifique par unité de masse,  $\rho$  la masse volumique du matériau supposée constante ici,  $\underline{\underline{\sigma}}$  le tenseur des contraintes,  $\underline{\underline{\varepsilon}}$  le tenseur des déformations en petites déformations,  $r$  la source volumique de chaleur et  $\underline{\underline{q}}$  le taux de flux de chaleur par unité de volume and ":" correspond au produit doublement contracté de tenseurs.

La seconde loi de la thermodynamique peut être formulée localement comme

$$\rho \dot{s} + \text{div} \left( \frac{\underline{\underline{q}}}{\Theta} \right) - \frac{r}{\Theta} \geq 0 \quad (6)$$

avec  $s$  l'entropie spécifique par unité de masse et  $\Theta$  la température. En substituant les termes relatifs à la source de chaleur dans (6) avec la forme de (5), les deux premiers principes permettent d'écrire l'inégalité de Clausius-Duhem :

$$\underline{\underline{\sigma}} : \underline{\underline{\dot{\varepsilon}}} - \rho (\dot{\psi} + s\dot{\Theta}) - \frac{\underline{\underline{q}} \cdot \text{grad}\Theta}{\Theta} \geq 0 \quad (7)$$

avec  $\psi := e - \Theta s$  étant l'énergie libre de Helmholtz.  $\psi$  est une fonction dépendant des déformations, de la température et des variables internes :

$$\psi = \psi(\underline{\underline{\varepsilon}}, T, \mathbf{V}_k) \quad (8)$$

et donc la dérivée temporelle de  $\psi$  s'écrit :

$$\dot{\psi} = \frac{\partial \psi}{\partial \underline{\underline{\varepsilon}}} : \dot{\underline{\underline{\varepsilon}}} + \frac{\partial \psi}{\partial \Theta} \dot{\Theta} + \frac{\partial \psi}{\partial \mathbf{V}_k} \cdot \dot{\mathbf{V}}_k \quad (9)$$

avec "." le produit adapté en fonction de la nature des variables internes  $\mathbf{V}_k$ . Ainsi, l'inégalité de Clausius-Duhem (7) peut être formulée comme

$$\left( \underline{\underline{\sigma}} - \rho \frac{\partial \psi}{\partial \underline{\underline{\varepsilon}}} \right) : \dot{\underline{\underline{\varepsilon}}} - \rho \left( s + \frac{\partial \psi}{\partial \Theta} \right) \dot{\Theta} - \rho \frac{\partial \psi}{\partial \mathbf{V}_k} \cdot \dot{\mathbf{V}}_k - \frac{q \cdot \text{grad} \Theta}{\Theta} \geq 0 \quad (10)$$

Cette expression doit être vraie pour tous les chemins, y compris pour un chemin de déformation élastique à température constante, ainsi :

$$\underline{\underline{\sigma}} = \rho \frac{\partial \psi}{\partial \underline{\underline{\varepsilon}}^e}(\underline{\underline{\varepsilon}}^e) \quad (11)$$

avec  $\underline{\underline{\varepsilon}}^e$  qui désigne un chemin de déformation élastique.

La décomposition additive de la déformation permet de réécrire la relation (11) avec la déformation totale en posant la variable interne  $V_0 = \underline{\underline{\varepsilon}}^{an}$  :

$$\underline{\underline{\sigma}} = \rho \frac{\partial \psi}{\partial \underline{\underline{\varepsilon}}} \quad (12)$$

avec  $\psi$  une fonction dépendant de :

$$\psi = \psi(\underline{\underline{\varepsilon}}, T, \mathbf{V}_k) \quad (13)$$

Dans toute la suite, il est posé  $\rho \psi = \psi$  sans perdre en généralité.

Notre réseau thermodynamique récurrent s'écrit :

$$\begin{cases} \psi^t &= RNN_{\theta} \left( \left\{ \underline{\underline{\varepsilon}}^{t'} ; t' \leq t \right\} \right) = RNN_{\theta} \left( \underline{\underline{\varepsilon}}^t, c^{t-1} \right) \\ \underline{\underline{\sigma}}^t &= \frac{\partial \psi^t}{\partial \underline{\underline{\varepsilon}}^t} \end{cases} \quad (14)$$

$RNN_{\theta}$  dénote le réseau de neurones récurrents, avec  $c^{t-1}$  l'état de ces mémoires cachées représentant l'histoire des déformations.

Ce système correspond à la définition de notre modèle de réseau thermodynamique récurrent (RTAN). Ce modèle est testé dans la section suivante à partir de données synthétiques sur des essais de cisaillement.

### 3.3 Expériences numériques

#### 3.3.1 Génération de données

Le modèle est testé sur des données synthétiques de cisaillement. Cependant, la méthodologie développée est directement applicable à des données expérimentales.

Le modèle synthétique est issu du travail de Colombo et al. [6] qui modélise l'influence de l'apport de renfort dans une matrice élastomère. Il combine des interactions physiques avec un adoucissement des contraintes pour introduire une non-linéarité de la réponse en contrainte en fonction du chargement en déformation.

Ce modèle non-linéaire est utilisé pour générer des essais de cisaillement cyclique : ces essais sont classiquement réalisés pour caractériser un polymère lors d'une analyse mécanique dynamique (DMA). Ces essais dépendent de la fréquence et de l'amplitude de la déformation.

Le choix des paramètres du modèle non-linéaire de génération de données et le choix des paramètres d'amplitudes et de fréquences des cycles sont faits de manière à balayer la non-linéarité du matériau.

### 3.3.2 Résultats numériques

Le modèle de réseau thermodynamique récurrent est testé ici avec les données issues de la méthodologie de la section 3.3.1.

L'influence de la quantité de données sur les résultats de notre modèle est très important pour identifier les capacités d'inférences du réseau de neurones. Ainsi, le modèle est entraîné avec quatre ensembles différents de données, avec respectivement 10, 40, 160 et 2560 données. Les figures présentées ci-dessous montrent systématiquement les inférences de chacun de ces modèles entraînés.

Les apprentissages ont été faits sur des courbes sinusoïdales dont étaient modifiées les amplitudes et les fréquences de déformations. Ainsi, la figure 2 montre un résultat d'interpolation avec la fréquence et l'amplitude moyenne de l'entraînement pour différents ensembles d'apprentissage.

Toutes les composantes du tenseur symétrique des contraintes  $y$  sont présentées. Il peut y être observé une très bonne interpolation de la contrainte du matériau non linéaire.

Dans le domaine extrapolatoire, la figure 3 présente la réponse du modèle à un essai de relaxation, un autre type de chargement que ceux vus durant l'entraînement (des essais cycliques). Ce chargement vient solliciter tous les mécanismes de déformations du matériau, ainsi tous les temps caractéristiques de relaxation sont représentés. Il est observé sur la figure 3 un bon accord entre la réponse en contrainte des RTAN entraînés et la réponse réelle du matériau. Ceci montre la capacité de notre modèle à généraliser l'entraînement cyclique à d'autres types de chargement.

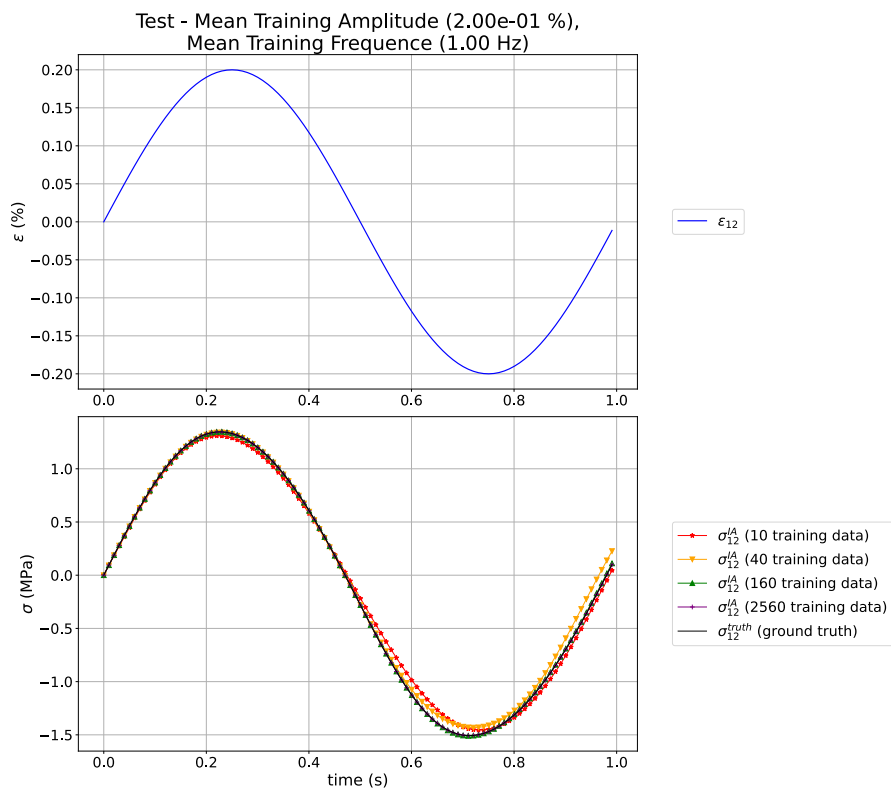
Il a été vu dans la section 3.3.1 que le modèle synthétique utilisé pour entraîner notre modèle correspond à un modèle non-linéaire dont la réponse en contrainte dépend de l'amplitude des déformations. La non-linéarité peut-être mise en évidence en étudiant la réponse en contrainte à différentes amplitudes de déformations. Lors d'un essai cyclique, il est extrait la dernière boucle de contrainte-déformation. A fréquence fixe, la figure 4a montre cette dernière boucle de contrainte-déformation à différentes amplitudes de déformation pour le modèle synthétique non linéaire.

Les autres courbes de la figure 4 montre la capacité de notre modèle à reproduire cette non linéarité.

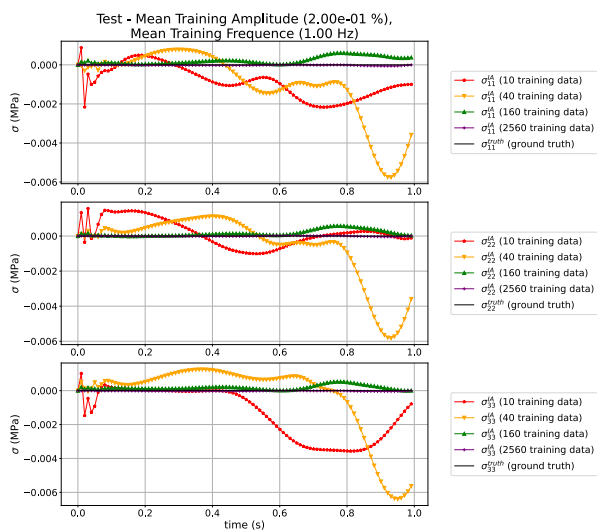
## 4 Conclusion

Le réseau thermodynamique récurrent développé ici propose une modélisation de matériau avec dépendance à l'histoire de déformation. Ce modèle se repose sur le cadre thermodynamique défini par les deux premiers principes. Les résultats numériques mettent en évidence les très bonnes capacités d'apprentissage de ce modèle pour l'interpolation. Le changement de type de chargement permet de montrer les capacités extrapolatoires de notre modèle. La non-linéarité modélisée a été bien inférée par le RTAN.

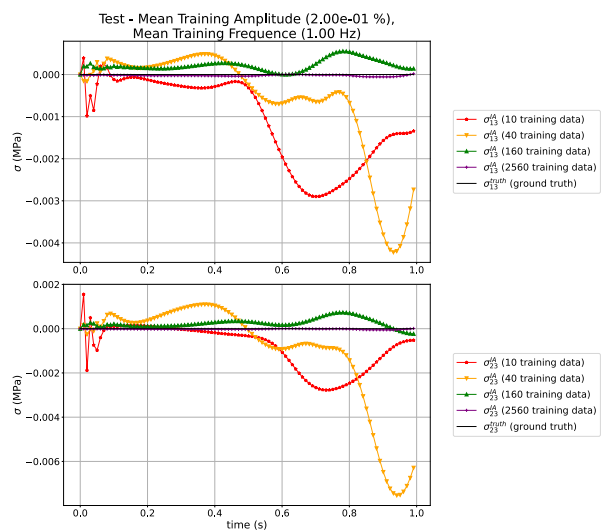
Ce modèle a été développé dans le cadre général 3D permettant de s'intéresser par la suite à des chargements complexes et à développer des méthodologies d'apprentissage pour traiter des cas en plus grandes dimensions.



(a) Direction de cisaillement



(b) Composantes diagonales du tenseur



(c) Composantes hors diagonales du tenseur

FIGURE 2 – Inférence des modèles entraînés pour la fréquence et l'amplitude moyenne des entraînements.

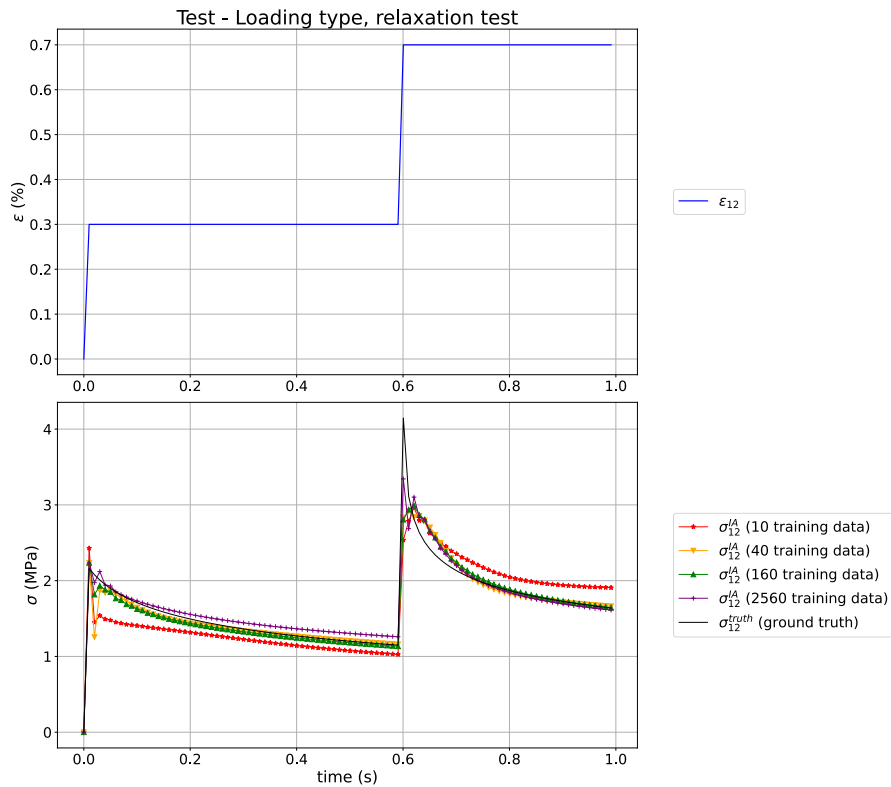


FIGURE 3 – Inférence des modèles sur un essai de relaxation dans la direction du cisaillement.

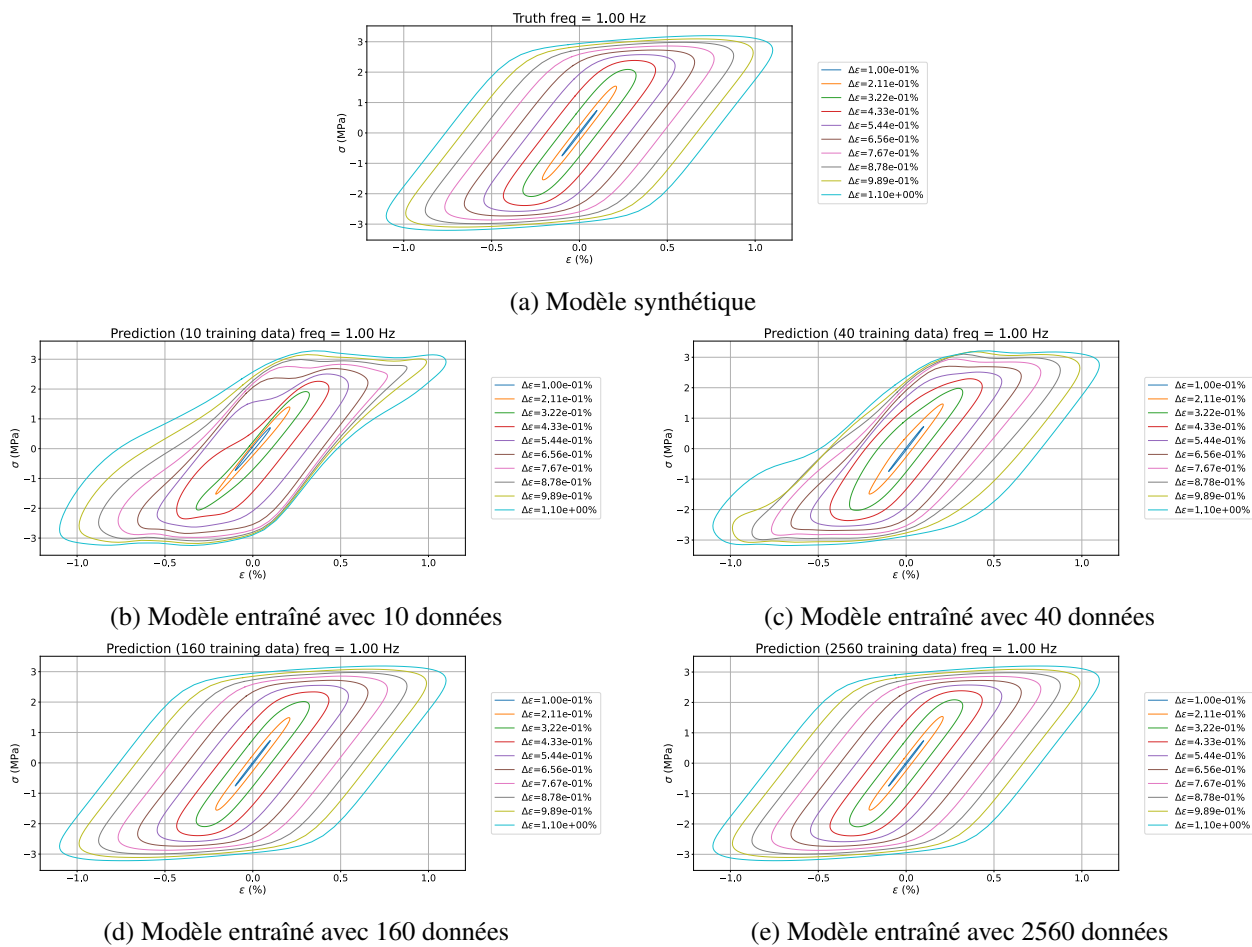


FIGURE 4 – Dernière boucle d'essai cyclique avec une fréquence de 1 Hz à différentes amplitudes

## Références

- [1] Frederic E. BOCK, Roland C. AYDIN, Christian J. CYRON, Norbert HUBER, Surya R. KALIDINDI et Benjamin KLUSEMANN. “A Review of the Application of Machine Learning and Data Mining Approaches in Continuum Materials Mechanics”. In : *Frontiers in Materials* 6 (15 mai 2019), p. 110.
- [2] Colin BONATTI et Dirk MOHR. “One for all : Universal material model based on minimal state-space neural networks”. In : *Science Advances* 7.26 (25 juin 2021), eabf3658.
- [3] L. BORKOWSKI, C. SORINI et A. CHATTOPADHYAY. “Recurrent neural network-based multiaxial plasticity model with regularization for physics-informed constraints”. In : *Computers & Structures* 258 (jan. 2022), p. 106678.
- [4] Jonathan CHAMPAGNE. “Modélisation physique du comportement mécanique linéaire et non-linéaire des élastomères renforcés”. These de doctorat. Université Paris sciences et lettres, 11 déc. 2020.
- [5] Kyunghyun CHO, Bart van MERRIENBOER, Caglar GULCEHRE, Dzmitry BAH DANAU, Fethi BOUGARES, Holger SCHWENK et Yoshua BENGIO. “Learning Phrase Representations using RNN Encoder-Decoder for Statistical Machine Translation”. In : (2014). Publisher : arXiv Version Number : 3.
- [6] Davide COLOMBO, Hélène MONTES, François LEQUEUX et Sabine CANTOURNET. “Thermo-mechanical modeling of a filled elastomer based on the physics of mobility reduction”. In : *Mechanics of Materials* 143 (avr. 2020), p. 103319.
- [7] Maysam B. GORJI, Mojtaba MOZAFFAR, Julian N. HEIDENREICH, Jian CAO et Dirk MOHR. “On the potential of recurrent neural networks for modeling path dependent plasticity”. In : *Journal of the Mechanics and Physics of Solids* 143 (oct. 2020), p. 103972.
- [8] Sepp HOCHREITER et Jürgen SCHMIDHUBER. “Long Short-Term Memory”. In : *Neural Computation* 9.8 (1<sup>er</sup> nov. 1997), p. 1735-1780.
- [9] Arnd KOEPE, Franz BAMER et Bernd MARKERT. “An efficient Monte Carlo strategy for elasto-plastic structures based on recurrent neural networks”. In : *Acta Mechanica* 230.9 (sept. 2019), p. 3279-3293.
- [10] M.A. MAIA, I.B.C.M. ROCHA, P. KERFRIDEN et F.P. VAN DER MEER. “Physically recurrent neural networks for path-dependent heterogeneous materials : Embedding constitutive models in a data-driven surrogate”. In : *Computer Methods in Applied Mechanics and Engineering* 407 (mars 2023), p. 115934.
- [11] Filippo MASI, Ioannis STEFANO, Paolo VANNUCCI et Victor MAFFI-BERTHIER. “Thermodynamics-based Artificial Neural Networks for constitutive modeling”. In : *Journal of the Mechanics and Physics of Solids* 147 (fév. 2021), p. 104277.
- [12] M. MOZAFFAR, R. BOSTANABAD, W. CHEN, K. EHMANN, J. CAO et M. A. BESSA. “Deep learning predicts path-dependent plasticity”. In : *Proceedings of the National Academy of Sciences* 116.52 (26 déc. 2019), p. 26414-26420.
- [13] Martin RENDEK et Alexander LION. “Amplitude dependence of filler-reinforced rubber : Experiments, constitutive modelling and FEM – Implementation”. In : *International Journal of Solids and Structures* 47.21 (oct. 2010), p. 2918-2936.
- [14] I.B.C.M. ROCHA, P. KERFRIDEN et F.P. VAN DER MEER. “On-the-fly construction of surrogate constitutive models for concurrent multiscale mechanical analysis through probabilistic machine learning”. In : *Journal of Computational Physics : X* 9 (jan. 2021), p. 100083.
- [15] Ling WU, Van Dung NGUYEN, Nanda Gopala KILINGAR et Ludovic NOELS. “A recurrent neural network-accelerated multi-scale model for elasto-plastic heterogeneous materials subjected to random cyclic and non-proportional loading paths”. In : *Computer Methods in Applied Mechanics and Engineering* 369 (sept. 2020), p. 113234.