



Prédiction du comportement mécanique non linéaire de matériaux dissipatifs sous chargements non proportionnels avec des Réseaux de Neurones Récurrents Thermodynamiquement Consistants

Aymen Danoun, Etienne Prulière, Yves Chemisky

► To cite this version:

Aymen Danoun, Etienne Prulière, Yves Chemisky. Prédiction du comportement mécanique non linéaire de matériaux dissipatifs sous chargements non proportionnels avec des Réseaux de Neurones Récurrents Thermodynamiquement Consistants. 15ème colloque national en calcul des structures, Université Polytechnique Hauts-de-France [UPHF], May 2022, 83400 Hyères-les-Palmiers, France. hal-03717684

HAL Id: hal-03717684

<https://hal.science/hal-03717684>

Submitted on 8 Jul 2022

HAL is a multi-disciplinary open access archive for the deposit and dissemination of scientific research documents, whether they are published or not. The documents may come from teaching and research institutions in France or abroad, or from public or private research centers.

L'archive ouverte pluridisciplinaire **HAL**, est destinée au dépôt et à la diffusion de documents scientifiques de niveau recherche, publiés ou non, émanant des établissements d'enseignement et de recherche français ou étrangers, des laboratoires publics ou privés.

Prédiction du comportement mécanique non linéaire de matériaux dissipatifs sous chargements non proportionnels avec des Réseaux de Neurones Récurrents Thermodynamiquement Consistants

Aymen Danoun¹, Etienne Prulière², Yves Chemisky¹

¹ Univ. Bordeaux, I2M, UMR 5295 351 Cours de la Libération, 33400 Talence, France, {aymen.danoun,yves.chemisky}@u-bordeaux.fr

² École Nationale Supérieure d'Arts et Métiers, Campus de Bordeaux, I2M, Esplanade des Arts et Métiers, Etienne.PRULIERE@ENSAM.eu

Résumé — Le présent travail vise à proposer un modèle hybride basé sur la physique et l'IA utilisant des réseaux de neurones récurrents thermodynamiquement consistants (ThC-RNN) pour prédire le comportement mécanique non linéaire de matériaux dissipatifs sous des trajets de chargement multi-axiaux non proportionnels. L'effet de la consistance thermodynamique sur la fiabilité du modèle est testé sur un modèle de plasticité à travers l'évaluation de différentes quantités d'intérêts (contraintes, variables internes, énergies mécaniques réversible, irréversible et dissipée).

Mots clés — Réseaux de neurones récurrents, Plasticité, Thermodynamique, Matériaux dissipatifs.

1 Introduction

Le développement des approches dites "Data-Driven" suscite l'intérêt de la communauté de mécanique pour leurs capacités à surmonter les limitations liées aux temps de calcul des simulations actuelles. En tant que technique d'apprentissage automatique, l'utilisation de réseaux neuronaux artificiels (ANN) comme métamodèle s'est avérée être un outil fiable pour plusieurs applications, notamment en analyse de structures, formulation de lois de comportement et modélisation multi-échelle. Ce succès repose sur le fait que dans le groupe des métamodèles, les ANN sont des algorithmes d'approximation universel [1]. Différents modèles et architectures ANN ont été développés pour résoudre des problèmes d'apprentissage tels que la reconnaissance de la parole et des images et la prédiction des tendances des données. Le plus courant est le perceptron multicouche (Multi Layer Perceptron MLP) [2]. L'architecture MLP a été utilisée par plusieurs auteurs comme métamodèle de la modélisation constitutive de matériaux. Par exemple, Furukawa et al [3] ont utilisé le MLP pour modéliser les comportement viscoplastique des matériaux et Waszczyszyn [4] a employé cette même architecture pour l'analyse de la flexion des poutres élastoplastiques. De nombreuses études ont également été menées pour montrer le potentiel des ANN en tant que technique puissante pour le gain de temps de calcul. Par exemple, le facteur de gain de temps de calcul obtenu par Settgast et al [5] pour la modélisation ANN des mousses poreuses a été estimé à 10 000. De plus, un facteur identique a été évalué par Ali et al [6] lors de l'utilisation de MLP comme architecture pour prédire les comportements mécaniques et l'évolution de la texture de métaux polycristallins sous sollicitations de traction-cisaillement. Cependant, nous tenons à préciser que ces études ont été menées sur un cas très particulier qui est la proportionnalité des trajets de chargement, conduisant ainsi à une simplification importante du problème.

Bien que le perceptron multicouche s'est avéré approprié pour traiter plusieurs problèmes mécaniques, il a été démontré que cette architecture n'est pas adaptée à la modélisation constitutive où la réponse mécanique du matériaux dépend du trajet de chargement appliqué [7]. Par exemple, l'utilisation d'un MLP comme métamodèle de loi de comportement non linéaire pour des matériaux dissipatifs sous des chemins de chargement non proportionnels présente une difficulté : il n'existe pas de relation directe entre l'état mécanique à l'étape actuelle et à l'étape précédente en considérant l'état de sollicitation actuel uniquement. Ce comportement dépend de l'historique du chargement. Cette difficulté peut être surmontée en introduisant une architecture ANN spécifique connue sous le nom de Réseau neuronal récurrent (Recurrent Neural Network RNN). Les RNN [8] sont conçus pour traiter des séquences temporelles par

l'introduction d'une mémoire. Contrairement aux MLP, les entrées des RNN ne sont pas seulement liées aux entrées de l'incrément du temps actuel, mais aussi aux sorties de tout les états précédents. Ce type d'architecture a été largement appliqué à des tâches telles que la reconnaissance vocale, la génération de texte ou la traduction automatique où le contexte est pris en considération par la mémoire. Pour les applications en science des matériaux, les architectures basées sur les RNN ont été récemment introduites pour accélérer les simulations multi-échelles par éléments finis [9, 10, 11]. En outre, ce type d'architecture a été également utilisé pour la modélisation constitutive de la plasticité et de la thermo-viscoplasticité [7, 12].

Bien que les méthodes basées sur les RNN aient montré leur fort potentiel en tant qu'outil fiable pour la modélisation des comportements dépendant de l'histoire de chargement, nous rappelons que les approches Data Driven en général, y compris les RNN, sont des algorithmes d'approximation considérés comme des "boîtes noires". Une "boîte noire" signifie que le modèle ne prend pas nécessairement en considération des lois d'évolution explicites issues de l'analyse physique pour faire des prédictions. En particulier, il n'a aucune raison de respecter les principes de la mécanique newtonienne et de la thermodynamique. Ce manque de cohérence physique peut présenter des limitations majeures sur la fiabilité du modèle lorsqu'il est entraîné sur un nombre limité de données, car la taille de la base de données est considérée comme la principale source d'enrichissement du modèle. Ainsi, l'utilisation de modèles "boîte noire" pour des applications d'ingénierie peut parfois être remise en question lorsque le respect des lois physiques n'est pas garanti. L'incorporation de la physique dans les modèles basés sur l'IA est un nouveau domaine de recherche prometteur qui tente de combiner les principes physiques avec les techniques d'apprentissage automatique. Le développement de tels modèles hybrides avec une utilisation complète des données et de la physique peut potentiellement fournir des résultats fiables en termes de capacité de généralisation des ANN ainsi qu'une cohérence physique. Dans cette étude, un réseau neuronal récurrent hybride thermodynamiquement consistant ThC-RNN est développé en incorporant des fonctions coûts basées sur la physique avec un modèle Stacked LSTM (Stacked Long Short Term Memory). Par conséquent, l'objectif principal de cette étude est d'investiguer l'effet de la combinaison des lois physiques et en particulier des lois thermodynamiques avec des modèles basés sur les RNN pour la modélisation constitutive des matériaux dissipatifs, en particulier pour une loi élastoplastique avec écrouissage isotrope.

2 Présentation du modèle et méthodologie

2.1 Présentation de l'architecture du RNN utilisé

Dans cette étude, une architecture Stacked LSTM a été utilisée au lieu d'un modèle RNN classique. Ce choix s'explique par la raison suivante : Lorsque les données d'entrée sont composées de longues séquences, l'apprentissage des RNN peut souffrir d'un phénomène courant connu sous le nom du Vanishing Gradient. Pour résoudre ce problème, le réseau neuronal LSTM (Long Short Term Memory) a été introduit par [13]. L'architecture du LSTM est basée sur un mécanisme de portes, permettant au réseau de contrôler le flux d'informations. Une cellule LSTM de base se compose de quatre portes principales, le Input Gate (pour la mise à jour et l'ajout de nouvelles informations), le Forget Gate (pour décider quelles informations sont conservées ou oubliées parmi les informations accumulées précédemment), le Memory Gate (pour donner au réseau une mémoire à long terme des événements passés) et le Output Gate (pour prédire le prochain état caché). Pour chaque pas de temps, le flux d'information est régulé dans et hors de la cellule LSTM en utilisant les quatre portes. La figure 1 représente l'architecture du modèle utilisée dans cette étude. Le modèle est composé de deux cellules LSTM₁ et LSTM₂ ainsi que d'une couche *Dense* pour la prédiction des données de sortie. Chaque porte de couche LSTM est régie par les équations du tableau de la figure 2. $W^{1,2}$, $U^{1,2}$ et $b^{1,2}$ désignent respectivement la matrice des poids et les vecteurs de biais des gates LSTM₁ et LSTM₂. W^D et b^D sont respectivement la matrice de poids et le biais correspondants à la couche *Dense*. σ est la fonction sigmoïde $\tanh$ la fonction tangente hyperbolique.

$$\sigma(x) = \frac{1}{1 + e^{-x}}$$

$$\tanh(x) = \frac{e^x - e^{-x}}{e^x + e^{-x}}$$

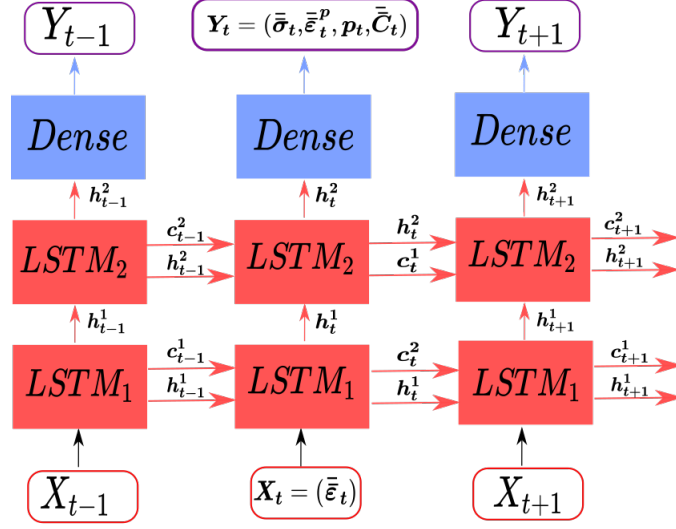


FIGURE 1 – Architecture de réseau neuronal Stacked LSTM pour la prédiction du comportement mécanique d'un matériau élastoplastique avec écrouissage isotrope.

Stacked LSTM Layers	LSTM ₁ Layer	LSTM ₂ Layer	Dense Layer
Input	X_t	h_t^1	h_t^2
Equations	$f_t^1 = \sigma(W_f^1 X_t + U_f^1 h_{t-1}^1 + b_f^1)$ $i_t^1 = \sigma(W_i^1 X_t + U_i^1 h_{t-1}^1 + b_i^1)$ $o_t^1 = \sigma(W_o^1 X_t + U_o^1 h_{t-1}^1 + b_o^1)$ $\tilde{c}_t^1 = \tanh(W_c^1 X_t + U_c^1 h_{t-1}^1 + b_c^1)$ $c_t^1 = f_t^1 \odot c_{t-1}^1 + i_t^1 \odot \tilde{c}_t^1$ $h_t^1 = o_t^1 \odot \tanh(c_t^1)$	$f_t^2 = \sigma(W_f^2 h_t^1 + U_f^2 h_{t-1}^2 + b_f^2)$ $i_t^2 = \sigma(W_i^2 h_t^1 + U_i^2 h_{t-1}^2 + b_i^2)$ $o_t^2 = \sigma(W_o^2 h_t^1 + U_o^2 h_{t-1}^2 + b_o^2)$ $\tilde{c}_t^2 = \tanh(W_c^2 h_t^1 + U_c^2 h_{t-1}^2 + b_c^2)$ $c_t^2 = f_t^2 \odot c_{t-1}^2 + i_t^2 \odot \tilde{c}_t^2$ $h_t^2 = o_t^2 \odot \tanh(c_t^2)$	$Y_t = \tanh(W^D h_t^2 + b^D)$

FIGURE 2 – Équations liées au modèle Stacked LSTM

Les paramètres d'entrée du modèle à chaque pas de temps t sont les composantes incrémentales du tenseur de déformation $\mathbf{X}_t = (\bar{\epsilon}_t)$. La sortie $\mathbf{Y}_t$ est donnée par l'ensemble de variables $\mathbf{Y}_t = (\bar{\sigma}_t, \zeta_t, \bar{C}_t)$ qui comprend les composantes du tenseur des contraintes $\bar{\sigma}_t$, les variables internes de la loi constitutive ζ_t et la matrice tangente $\bar{C}_t$. Pour la loi de comportement élastoplastique avec écrouissage isotrope, ζ_t contient la déformation plastique accumulée p_t et les composantes du tenseur de déformation plastique $\bar{\epsilon}_t^p$, donc $\zeta_t = (\bar{\epsilon}_t^p, p_t)$.

2.2 Apprentissage thermodynamiquement consistant

Le processus d'apprentissage du RNN est basé sur la technique de rétropropagation dans le temps (Back Propagation Through Time BPTT) qui est l'application de l'algorithme de rétropropagation [14] aux séquences temporelles. La technique de rétropropagation est appliquée pour l'apprentissage du réseau ANN par la règle de la chaîne en calculant la dérivée partielle d'une fonction coût par rapport aux paramètres du modèle (poids et biais). Après chaque passage dans la direction avant du réseau, la rétropropagation effectue un passage en arrière tout en ajustant les poids et le biais par la technique de descente de gradient. Pour mesurer la précision de prédiction du RNN, une fonction coût $\mathcal{L}$ est définie. Pour chaque pas de temps t , $\mathcal{L}_t$ correspond à l'erreur entre les valeurs prédites par le RNN y^p et les

données de sortie réelles y^{true} . Pour le problème étudié, la fonction de perte MSE (Mean Squared Error) est utilisée :

$$\mathcal{L}_t = \frac{1}{2N} \sum \left(y_t^{(p)} - y_t^{(true)} \right)^2 \quad (1)$$

N est le nombre total de données d'apprentissage au pas de temps t

La fonction coût totale $\mathcal{L}$ de l'ensemble du réseau peut être définie comme la somme des fonctions coût à chaque pas de temps $\mathcal{L}_t$.

$$\mathcal{L} = \sum \mathcal{L}_t \quad (2)$$

En ce qui concerne le processus d'apprentissage des modèles RNN à l'aide de la BPPT, nous souhaitons souligner que la minimisation d'une seule fonction objective telle que MSE (voir Eq 1) ne garantit pas nécessairement des prédictions cohérentes avec la connaissance des lois physiques. Dans cette étude, une nouvelle fonction de coût basée sur la physique $\mathcal{L}^\Phi$ est proposée pour assurer la cohérence thermodynamique du modèle.

$$\mathcal{L}^\Phi = \sum \mathcal{L}_t^\Phi \quad (3)$$

Pour chaque pas de temps, $\mathcal{L}_t^\Phi$ est défini comme suit :

$$\mathcal{L}_t^\Phi = \frac{1}{2N} \sum (y_t^{(p)} - y_t^{(t)})^2 + \lambda Relu(-\gamma_{loc}) \quad (4)$$

où γ_{loc} est la dissipation intrinsèque déterminée dans le matériau à chaque pas de temps t nécessairement positive, cette condition correspond à la forme forte du second principe de la thermodynamique.

Pour la loi de comportement élastoplastique avec écrouissage isotrope suivant une loi puissance, γ_{loc} est exprimée :

$$\gamma_{loc} = \sigma : \dot{\epsilon}^p - H p^n \dot{p}$$

H and n sont le paramètre d'écrouissage isotrope et exposant d'écrouissage respectivement.

La fonction $ReLU$ est l'unité linéaire rectifiée définie comme suit :

$$Relu(x) = \begin{cases} 0 & \text{for } x < 0 \\ x & \text{for } x \geq 0 \end{cases}$$

D'après l'équation 4, $\mathcal{L}_t^\Phi$ peut être divisé en deux termes : Le premier terme $\frac{1}{2N} \sum (y_p^{(t)} - y_{true}^{(t)})^2$ est la MSE qui mesure l'erreur quadratique moyenne entre les valeurs prédites et les valeurs réelles. Le deuxième terme est celui où la consistance thermodynamique est prise en compte pendant la phase d'apprentissage de l'ANN. Il peut être considéré comme une partie physique qui contraint le modèle à respecter le deuxième principe de la thermodynamique pendant la phase d'apprentissage pour tous les pas de temps. λ est un hyperparamètre qui contrôle le degré de régularisation du modèle. L'effet du terme physique pendant l'apprentissage peut être résumé comme suit :

- si $\gamma_{loc} > 0$ then $Relu(-\gamma_{loc}) = 0$ (Aucune pénalité sur la fonction coût)
- Si $\gamma_{loc} < 0$ then $Relu(-\gamma_{loc}) = -\gamma_{loc}$ (Ajout d'un terme de pénalité pour contraindre la fonction coût lorsqu'une incohérence physique se produit)

3 Résultats et discussion

Le modèle ThC-RNN est développé en associant un Stacked LSTM avec la fonction coût $\mathcal{L}_t^\Phi$ pendant le processus d'apprentissage. Le modèle ThC-RNN est comparé à un modèle Stacked LSTM standard pour évaluer l'effet de la consistance thermodynamique sur la fiabilité du modèle. Les deux modèles partagent la même architecture du réseau de neurones comme décrit dans la figure 1 mais diffèrent alors en terme de processus d'apprentissage adopté. Les modèles sont testés dans le cas d'un point matériel soumis à un trajet de chargement non proportionnel dans l'espace des déformations. Le matériau testé

est un alliage de titane Ti-6Al-4V dont l'écrouissage isotrope est décrit par une loi de puissance. Les propriétés mécaniques correspondantes sont données dans le tableau de la figure 3. Ensuite, la stratégie de préparation de la base de données est décrite comme suit : 10000 échantillons de trajets de chargement multiaxiaux et non proportionnels sont générés à l'aide de la librairie *Simcoon* [15]. Les paramètres d'entrée sont les composantes du tenseur de déformation incrémentales $\bar{\epsilon}^t = (\epsilon_{11}^t, \epsilon_{22}^t, \epsilon_{12}^t)$. La figure 4 montre quelques exemples de configurations de trajets de chargement utilisés comme données d'apprentissage. Les échantillons générés sont ensuite divisés en ensembles de données d'apprentissage (90 % du total des échantillons) et de prédiction (10%).

Material parameter	Value
Young's modulus E	113800 MPa
Poisson's ratio ν	0.34
Yield Stress σ_Y	1000 MPa
Hardening parameter H	1600 Mpa
Plastic hardening exponent n	0.5

FIGURE 3 – Propriétés mécanique de l'alliage du titane Ti-6Al-4V

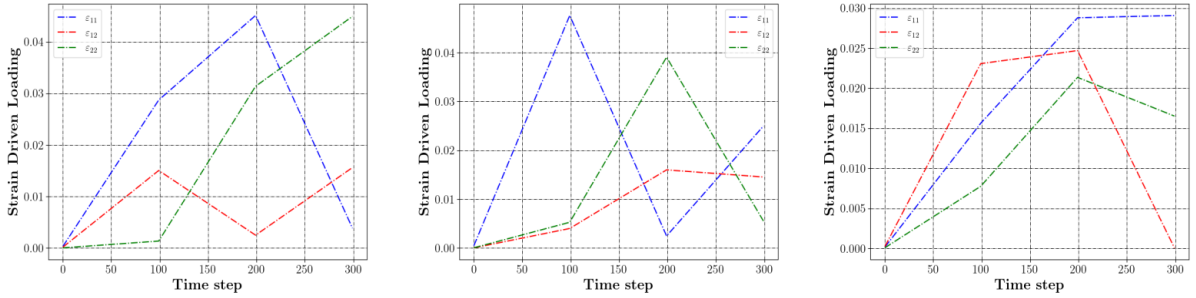


FIGURE 4 – Exemples de trajets de chargement multiaxiaux non proportionnelles utilisées comme données d'entraînement

3.1 Modèle Stacked LSTM standard

Pour évaluer les capacités de généralisation et de prédiction du modèle Stacked LSTM standard après la phase d'apprentissage, le modèle est testé sur de nouvelles configurations de trajets de chargement multiaxiaux et non proportionnelles afin de prédire la réponse mécanique résultante. La figure 5 montre une comparaison entre les réponses contraintes-déformation prédites par le modèle et les valeurs cibles obtenues via les simulations *Simcoon*. À première vue, on peut observer que la tendance générale des composantes de contraintes est bien capturée par le Stacked LSTM malgré la complexité du trajet de chargement. La quantification de la qualité prédictive globale du modèle peut être démontrée à l'aide des diagrammes de corrélation présentés dans la figure 5 où le coefficient de détermination R^2 est très proche de 1 pour toutes les réponses prédites. De plus, le modèle est également capable de prédire la transition de la courbe contrainte-déformation du domaine élastique au domaine plastique. Ces premiers résultats prouvent le potentiel du RNN en tant que substitut capable de modéliser des comportements mécaniques dépendant de l'histoire du chargement, donnant la possibilité de remplacer potentiellement les schémas itératifs tels que Newton Raphson par un modèle basé sur l'IA qui offre un avantage significatif en termes de gain de temps de calcul si on se place dans un cadre multi-échelles.

La vérification de la cohérence thermodynamique du modèle est mise en évidence par la prédiction des différents travaux mécaniques par rapport à la loi de comportement élastoplastique avec écrouissage isotrope. La figure 6 présente une comparaison entre les travaux mécaniques réversibles, irréversibles et dissipatifs prédits par le modèle Stacked LSTM standard et les valeurs cibles. On observe sur la figure

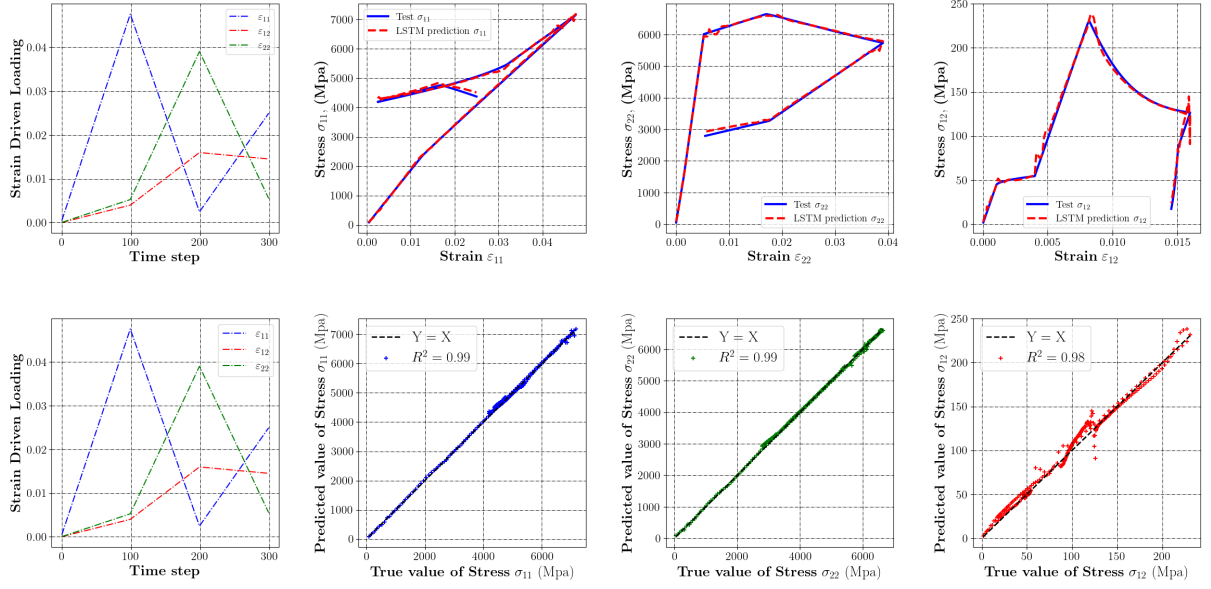


FIGURE 5 – Prédiction des réponses contraintes-déformations à l’aide d’un modèle Stacked LSTM standard sur un ensemble de données de test & Graphiques de corrélation entre les valeurs prédites du modèle et les valeurs cibles

6 que le travail mécanique dissipatif prédit est négatif pendant certains pas de temps même si la prédiction globale suit toujours la tendance générale de la dissipation cible. Ce résultat contredit le deuxième principe de la thermodynamique qui stipule que l’entropie locale générée γ_{loc} doit toujours être positive pendant un mécanisme dissipatif. Ainsi, le respect du second principe n’est pas assuré en utilisant le modèle Stacked LSTM standard avec seulement une MSE comme fonction coût. Les oscillations engendrées sur le comportement observées notamment lors d’un changement de trajet peuvent nuire à la convergence d’un modèle numérique. En utilisant un modèle RNN guidé par la physique ‘ThC-RNN’, nous avons l’objectif d’étudier l’effet de l’incorporation de contraintes basées sur la physique pendant le processus d’apprentissage du Stacked LSTM afin de satisfaire potentiellement les lois physiques gouvernantes et d’atteindre une consistance thermodynamique afin d’éviter ces perturbations.

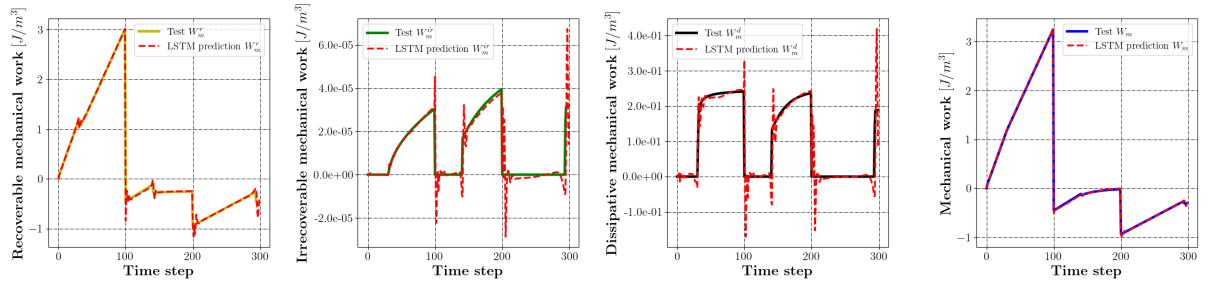


FIGURE 6 – Évaluation de divers énergies mécaniques avec un modèle Stacked LSTM standard sur un exemple d’ensemble de données de test.

3.2 Modèle Stacked LSTM Thermodynamiquement Consistant ThC-RNN

L’architecture ThC-RNN développée est basée sur un modèle Stacked LSTM avec une fonction coût basée sur la physique dont l’expression à chaque pas de temps t est donnée par l’équation 4. Le ThC-RNN est entraîné avec la même quantité de données que le modèle précédent. L’objectif est de vérifier les capacités du ThC-RNN à prédire avec précision le comportement mécanique de la loi de comportement élastoplastique ainsi que du respect des principes thermodynamiques lors du calcul des valeurs énergétiques mécaniques.

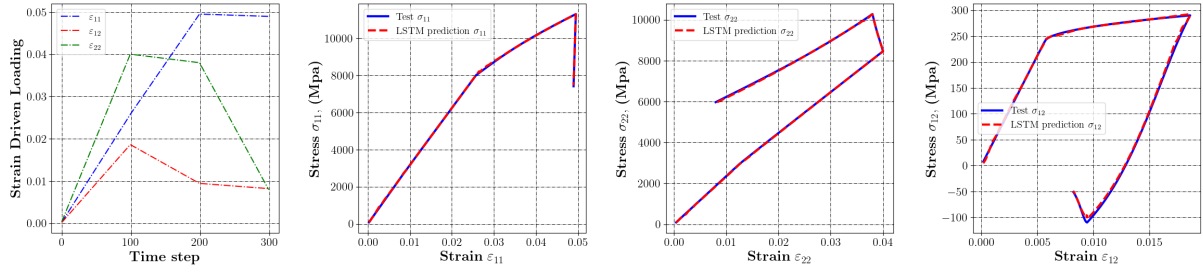


FIGURE 7 – Prédiction des réponses en contrainte-déformation avec le modèle ThC-RNN sur un exemple d'ensemble de données de test

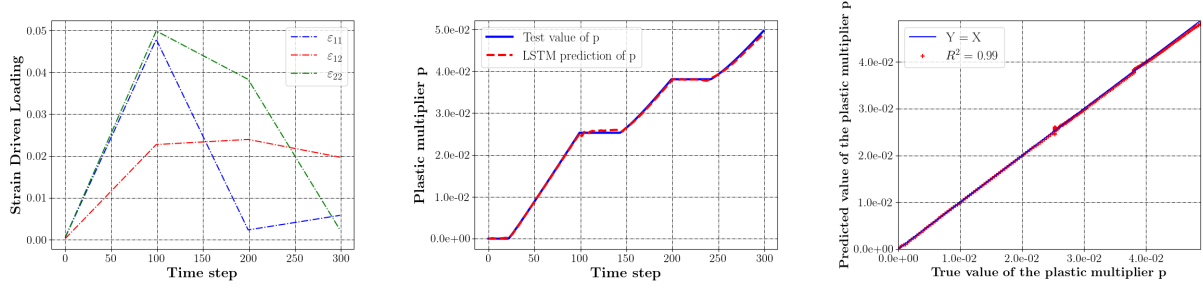


FIGURE 8 – Prédiction de la déformation plastique accumulée p avec le modèle ThC-RNN sur un exemple d'ensemble de données de test

La figure 7 représente les contraintes prédites par le modèle ThC-RNN et la réponse cible pour un exemple de trajet de chargement de données de test. Comme pour le modèle précédent, nous pouvons déduire que le modèle ThC-RNN assure toujours une performance très élevée lors de la prédiction des réponses contrainte-déformation. D'autre part, la figure 8 décrit l'évolution de la variable interne prédite p qui est le multiplicateur plastique. Comme le montre le graphique de corrélation de la figure 8, un excellent accord est observé entre les valeurs prédites par ThC-RNN et les cibles pour la déformation plastique accumulée, le modèle ThC-RNN est donc capable de capturer avec précision l'évolution de cette variable interne. De plus, les zones correspondants à un changement de trajet sont correctement prédites contrairement au modèle précédent. Ainsi, l'incorporation de cette loi physique pendant la phase d'apprentissage a un effet bénéfique important sur la précision du modèle.

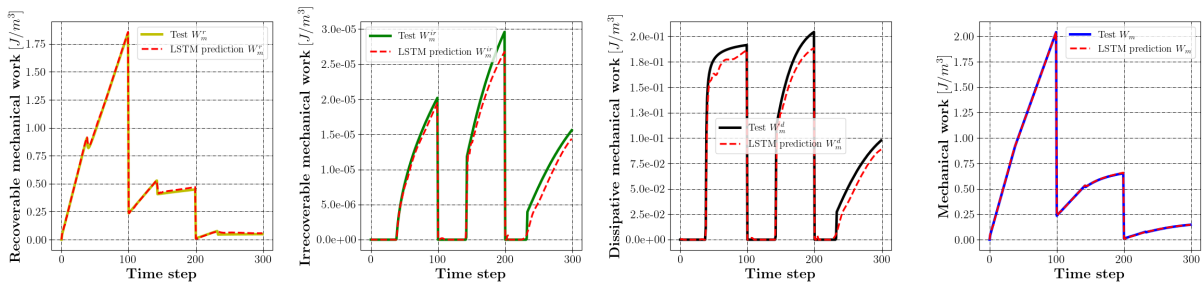


FIGURE 9 – Évaluation de divers énergies mécaniques avec le modèle ThC-RNN sur un exemple d'ensemble de données de test.

Pour bien mettre en évidence l'effet de la physique sur la fiabilité du modèle, les différentes valeurs énergétiques mécaniques sont évaluées comme précédemment. La figure 9 montre une comparaison entre les énergies mécaniques prédites par le modèle ThC-RNN et les valeurs cibles. En ce qui concerne l'énergie dissipée, on peut remarquer que le modèle est thermodynamiquement consistant puisque la dissipation intrinsèque est toujours supérieure ou égale à zéro par rapport à tous les incréments de temps. Les résultats obtenus avec ThC-RNN prouvent que l'incorporation de fonctions objectives basées sur

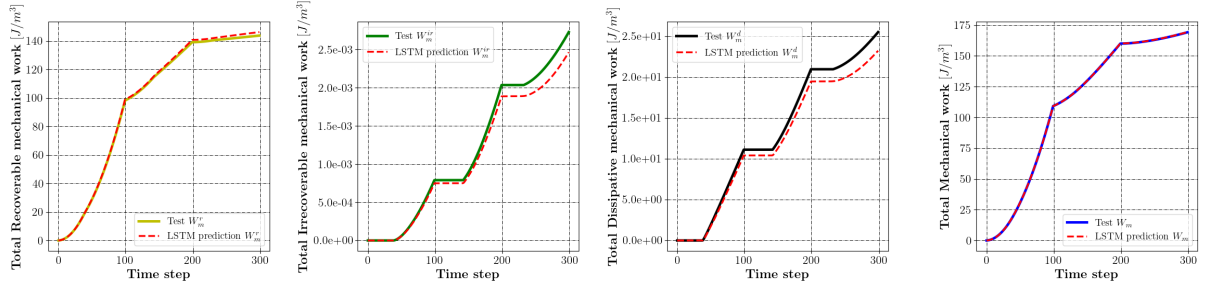


FIGURE 10 – Évaluation des grandeurs énergétiques totales avec le modèle ThC-RNN sur un exemple d'ensemble de données de test.

la physique et la présence de données suffisantes peut réduire de manière significative l'impact des incohérences physiques sur la précision du modèle. Cependant, lorsque nous avons analysé la qualité de la prédiction pour toutes les données test, nous avons observé la présence de certaines erreurs sur les énergies mécaniques irréversibles et dissipatifs prédites pendant certains incréments, cet écart peut également être mis en évidence sur la figure 10 où l'erreur entre la dissipation totale prédite et la valeur cible est d'environ 10 %. La principale hypothèse pour expliquer cette différence est la propagation des erreurs car le calcul de W_m^r et W_m^d fait intervenir plusieurs quantités dont les variables internes, les composantes de la déformation plastique et du tenseur des contraintes. Les erreurs de prédiction de valeurs de sorties, même de faibles valeurs, peuvent induire une amplification de l'erreur lors du calcul de W_m^r et W_m^d . Par conséquent, le processus d'apprentissage du modèle ThC-RNN utilisant la fonction de perte $\mathcal{L}_t^\varphi$ assure certainement le respect de la condition de positivité de la dissipation mais ne garantit pas nécessairement que le travail mécanique dissipatif prédit correspondra à la cible. Une solution potentielle pour surmonter ce problème sera d'ajouter un terme à la fonction objectif $\mathcal{L}_t^\varphi$ qui minimise également l'erreur entre les valeurs énergétiques mécaniques prédites et cibles de manière similaire au composantes du tenseur des contraintes et des variables internes.

4 Conclusions

Dans cette étude, le potentiel d'un modèle hybride physique-RNN (ThC-RNN) est étudié afin de servir comme métamodèle fiable et cohérent pour la modélisation constitutive du comportement de matériaux dissipatifs. Ce modèle hybride a été développé en combinant une architecture RNN spécifique appelée Stacked LSTM et une fonction coût basée sur la physique résultant de la thermodynamique des processus irréversibles. Les résultats obtenus ont montré les excellentes capacités de prédiction de ThC-RNN par rapport à plusieurs quantités d'intérêts (réponses contrainte-déformation, énergies mécaniques, variables internes) par rapport à un modèle RNN standard non consistant. La robustesse du ThC-RNN provient principalement de la performance du modèle Stacked LSTM qui peut être considéré comme un outil puissant pour capturer des comportements dépendants des trajets de chargement. D'autre part, l'incorporation des lois physiques pendant la phase d'apprentissage a fourni des résultats pertinents en ce qui concerne la cohérence thermodynamique. Le présent travail donne un premier aperçu des capacités des techniques d'apprentissage profond guidées par la physique pour traiter les problèmes d'ingénierie en mécanique des solides. Bien que cette étude ne se soit concentrée que sur la modélisation constitutive des matériaux dissipatifs, il pourrait être intéressant de tester l'approche proposée dans différents domaines d'application où les lois physiques gouvernantes peuvent être obtenues sous une forme explicite. Notez que même si les gains de temps de calcul ne sont pas mis en évidence au cours de cette étude, étant donné que les méthodes numériques existantes sont déjà assez rapides, ce travail peut être considéré comme une preuve de concept pour nos travaux futurs potentiels où les RNN guidés par la physique sont appliqués pour l'analyse multi-échelle par éléments finis. Dans ce cas, nous pensons que les coûts de calcul peuvent être réduits de manière drastique comme démontré par les études effectués sur des trajets de chargement proportionnels.

Références

- [1] Kurt Hornik, Maxwell Stinchcombe, and Halbert White. Multilayer feedforward networks are universal approximators. *Neural networks*, 2(5) :359–366, 1989.
- [2] G. Bebis and M. Georgiopoulos. Feed-forward neural networks. *IEEE Potentials*, 13(4) :27–31, October 1994.
- [3] Tomonari Furukawa and Genki Yagawa. IMPLICIT CONSTITUTIVE MODELLING FOR VISCOPLASTICITY USING NEURAL NETWORKSR. Technical report, 1998.
- [4] Zenon Waszczyszyn and Leonard Ziemiański. Neural networks in mechanics of structures and materials - New results and prospects of applications. *Computers and Structures*, 79(22-25) :2261–2276, sep 2001.
- [5] Christoph Settgastr, Martin Abendroth, and Meinhard Kuna. Constitutive modeling of plastic deformation behavior of open-cell foam structures using neural networks. *Mechanics of Materials*, 131(January) :1–10, 2019.
- [6] Usman Ali, Waqas Muhammad, Abhijit Brahme, Oxana Skiba, and Kaan Inal. Application of artificial neural networks in micromechanics for polycrystalline metals. *International Journal of Plasticity*, 120(May) :205–219, 2019.
- [7] Maysam B. Gorji, Mojtaba Mozaffar, Julian N. Heidenreich, Jian Cao, and Dirk Mohr. On the potential of recurrent neural networks for modeling path dependent plasticity. *Journal of the Mechanics and Physics of Solids*, 2020.
- [8] L. C. Jain and L. R. Medsker. *Recurrent Neural Networks : Design and Applications*. CRC Press, Inc., USA, 1st edition, 1999.
- [9] F. Ghavamian and A. Simone. Accelerating multiscale finite element simulations of history-dependent materials using a recurrent neural network. *Computer Methods in Applied Mechanics and Engineering*, 2019.
- [10] Bin Li and Xiaoying Zhuang. Multiscale computation on feedforward neural network and recurrent neural network. *Frontiers of Structural and Civil Engineering*, jan 2021.
- [11] Ling Wu, Van Dung Nguyen, Nanda Gopala Kilinger, and Ludovic Noels. A recurrent neural network-accelerated multi-scale model for elasto-plastic heterogeneous materials subjected to random cyclic and non-proportional loading paths. *Computer Methods in Applied Mechanics and Engineering*, 2020.
- [12] Diab W. Abueidda, Seid Koric, Nahil A. Sobh, and Huseyin Sehitoglu. Deep learning for plasticity and thermo-viscoplasticity. *International Journal of Plasticity*, 136 :102852, jan 2021.
- [13] Sepp Hochreiter and Jürgen Schmidhuber. Long Short-Term Memory. *Neural Computation*, 9(8) :1735–1780, nov 1997.
- [14] Yann Lecun. A theoretical framework for back-propagation. 08 2001.
- [15] Yves Chemisky. SIMCOON : OPEN SOURCE SOFTWARE IN MECHANICS AND MATERIALS | Yves Chemisky | 2 updates | Research Project. 2018.