Mechanics of Materials 212 (2026) 105525

Contents lists available at ScienceDirect

Mechanics of Materials

journal homepage: www.elsevier.com/locate/mecmat

Research paper

Enhancing nonlinear viscoelastic modeling of elastomers through neural networks: A deep rheological element

Federico Califano [a,*,1], Jacopo Ciambella [b,1]

[a] Department of Mechanical and Aerospace Engineering, Sapienza University of Rome, Via Eudossiana 18, Rome, 00184, Italy
[b] Department of Structural and Geotechnical Engineering, Sapienza University of Rome, Via Eudossiana 18, Rome, 00184, Italy

ARTICLE INFO

Invited Editor Jean-F Molinari

Keywords:
Viscoelasticity
Neural network
Constitutive equations
Finite strains
Elastomers
Soft tissues
Data-driven mechanics
Physics-informed machine learning

ABSTRACT

An established method for incorporating inelastic constitutive equations into finite element software is the use of rheological elements, assembled in series or in parallel, to describe the constitutive response at each material point. This approach can be extended to finite strains by exploiting the multiplicative decomposition of the deformation gradient. In this study, we propose an hybrid approach that integrates traditional elements with elements, with constitutive equations defined through deep neural networks (DNNs), into an assemblage of standard rheological elements. We formulate DNNs to guarantee thermodynamic consistency, enabling us to model the time-dependent, large strain response of elastomers and predict the Payne effect in filled rubber. This effect, characterized by deformation-enhanced shear thinning, poses unique modeling challenges. Additionally, we discuss data augmentation procedures to address the data-intensive nature of training neural networks, showcasing the effectiveness of utilizing ordinary dynamic mechanical analysis (DMA) tests for this purpose.

1. Introduction

Viscoelastic materials are extensively utilized across various engineering fields, including polymer processing and structural design, due to their unique time-dependent mechanical behavior. The mechanical response of these materials is influenced by both the magnitude and rate of deformation, making their characterization complex. For small and slow deformations, linear viscoelastic models provide a simple and effective means to describe the stress–strain relationship (Ciambella et al., 2011). However, these models fall short in capturing the nonlinear, time-dependent effects that become significant under large and rapid deformations (Gent, 2012). To address these limitations, advanced modeling techniques are required to account for the finite deformation behavior of viscoelastic materials. Among these approaches, the assemblage of rheological elements has proven to be a robust framework for capturing the time-dependent response of materials, which can be extended to large strains (Hurtado et al., 2013) and enriched with elastoplastic elements to enhance accuracy.

A key concept that facilitates the extension of rheological element assemblages to finite deformations is the multiplicative decomposition of the deformation gradient $\mathbf{F} = \mathbf{F}_e \mathbf{F}_v$ into elastic $\mathbf{F}_e$ and inelastic $\mathbf{F}_v$ components (Sadik and Yavari, 2017). This approach generalizes the classical additive decomposition of total deformation, commonly used at small strains, where deformation is split between spring and dashpot contributions, as illustrated in Fig. 1. Under the assumption of small strains, the multiplicative decomposition naturally reduces to the additive form. Originally derived from the theory of growth (DiCarlo and Quiligotti, 2002), this concept has been widely adopted in various fields, such as viscoplasticity (Ciambella and Nardinocchi, 2021), electroelasticity (Ghosh and Lopez-Pamies, 2021), and magnetoelasticity (Brancati et al., 2019). In many of these applications, multiple material phases coexist within a single material point, where one phase exhibits purely elastic behavior, while others display viscoplastic behavior, allowing for flow at zero stress. These material phases are constrained by a fundamental kinematic assumption that they undergo the same overall deformation, ensuring consistency within the model.

In the framework of viscoplasticity, the different literature proposals primarily differ in their constitutive formulations for the inelastic elements and, in particular, in viscoelasticity, for the constitutive assumptions governing the viscosity functions associated with the dashpot (Reese and Govindjee, 1998; Bergström and Boyce, 1998; Kumar and Lopez-Pamies, 2016; Hurtado et al., 2013). Accurately defining this function is essential to capture the complex mechanical behavior of materials, which exhibit a nonlinear dependence of viscosity on both strain and strain rate, such as elastomers that exhibit a marked deformation-enhanced shear thinning, a phenomenon known in the rubber industry as the Payne effect (Chazeau et al., 2000). Capturing this effect within constitutive models remains a significant challenge.

* Corresponding author.
E-mail addresses: federico.califano@uniroma1.it (F. Califano), jacopo.ciambella@uniroma1.it (J. Ciambella).
[1] All authors contributed equally to all aspects of this joint paper.

https://doi.org/10.1016/j.mechmat.2025.105525
Received 29 January 2025; Received in revised form 11 September 2025; Accepted 8 October 2025
Available online 16 October 2025

A promising and rapidly evolving approach in the field of computational mechanics involves the integration of data-driven techniques directly into the governing equations of mechanical problems. These data-driven methods can be broadly classified into two distinct categories (Ghaderi et al., 2020). The first category, known as model-free approaches, bypasses the need for an explicit constitutive law or a pre-trained model that is separate from the data. Instead, these strategies embed the experimental dataset directly into the solution process of the mechanical problem (Conti et al., 2018): the equilibrium solution is obtained by optimizing a compatibility measure between the governing constraints and the available data points, eliminating the requirement for a predefined stress–strain relationship. While this approach is conceptually appealing, it heavily relies on the availability of extensive and representative datasets and may face challenges when extrapolating beyond the observed data (Eggersmann et al., 2019; Salahshoor and Ortiz, 2023).

The second category of data-driven methods encompasses a range of approaches, with a strong influence from the resurgence of neural networks (NNs) in constitutive modeling, a field that has been explored in earlier studies as well (Furukawa and Yagawa, 1998; Ghaboussi and Sidarta, 1998; Jung and Ghaboussi, 2006). These NN-based models can be further classified into three fundamental types (Faroughi et al., 2024): (i) black-box neural networks, (ii) physics-informed networks that weakly incorporate physical principles, and (iii) physics-encoded networks that strongly enforce governing laws. Black-box neural networks rely solely on data to infer constitutive relationships without consideration of the underlying physics. In contrast, physics-informed and physics-encoded networks aim to integrate essential physical principles, with a particular emphasis on satisfying the second law of thermodynamics. Physics-informed neural networks (PiNNs) incorporate physical constraints into the loss function by adding penalty terms for deviations from governing equations (Taç et al., 2022b; Zhang et al., 2022). Physics-encoded neural networks (PeNNs) enforce physical constraints by embedding governing principles directly into the neural architecture. PeNNs have recently gained significant attention (Taç et al., 2024a).

In the context of hyperelasticity, physical requirements are inherently incorporated into the model by adopting invariant-based strain energy functions to ensure objectivity, while architectures utilizing convex activation functions are employed in PeNNs to enforce polyconvexity (Amos et al., 2017; Klein et al., 2022; As'ad et al., 2022; Chen and Guilleminot, 2022; Linden et al., 2023; Linka and Kuhl, 2021, 2023; Taç et al., 2022a, 2024b). In the case of dissipative models, additional constraints must be imposed on the dissipation potential to satisfy thermodynamic consistency (Kestin and Rice, 1970). Neural Ordinary Differential Equations (NODEs) (Chen and Guilleminot, 2022; Taç et al., 2022b) have been successfully extended to viscoelasticity, serving as surrogates for evolution equations, as demonstrated in the works of Jones et al. (2022) and in the distinctive approach proposed by Taç et al. Taç et al. (2023), which focuses on ensuring increasing monotonicity and non-negativity of the dissipation potential derivative, thereby guaranteeing the convexity of the dissipation potential. Moreover, Input Convex Neural Networks (ICNNs) and Convex Artificial Neural Networks (CANNs) have been employed to represent the dissipation potential. In the ICNN-based viscoelastic framework proposed by As'ad and Farhat (2023) (As'ad and Farhat, 2023), the strain energy function and the dissipation potential are both represented by Input Convex Neural Networks (ICNNs). ICNNs enforce convexity by construction through nonnegative weights and convex, non-decreasing activation functions. Similarly, the iCANN framework introduced by Holthusen et al. (2024) (Holthusen et al., 2024) integrates CANNs with a recurrent neural network (RNN) layer to effectively capture path-dependent and inelastic responses. Unlike standard neural networks, CANNs incorporate activation functions derived from expert-constructed models, specifically designed to preserve convexity. These activation functions operate on polynomial expansions, enabling the construction of convex, non-decreasing functions to represent the dissipation potentials. Recurrent Neural Networks (RNNs) are a class of neural networks designed to process sequential data by maintaining a memory of previous inputs, allowing them to capture temporal dependencies and path-dependent behaviors effectively (Sherstinsky, 2020).

Building upon these advancements, we propose a hybrid approach that complements traditional rheological elements with Deep Neural Networks (DNNs) (Goodfellow et al., 2016) to enhance the predictive capabilities of viscoelastic material models. Specifically, we introduce a novel component, termed the Deep Rheological Element (DRE), wherein the DNN is utilized exclusively to model the material's viscosity function. This approach ensures thermodynamic consistency, particularly adherence to the second law, through the imposition of minimal constraints on the DNN architecture and so address one of the key challenges of data-driven constitutive modeling. In contrast to conventional approaches that employ neural networks to directly predict the overall stress–strain relationship (Zhang et al., 2022; Chen, 2021), our methodology integrates classical strain energy density functions with a neural network-driven viscosity function. This strategic combination not only mitigates the dependency on extensive datasets but also preserves the model's capability to accurately capture the complex, nonlinear viscoelastic behaviors exhibited by materials such as elastomers.

The structure of the paper is as follows. Section 2 derives the governing equations of a continuum with an assemblage of parallel elements, emphasizing the main kinematic assumptions underlying the framework. It also presents the proposed data-driven approach and compares it with the traditional model-based formulation. The training of the neural network and the data augmentation procedure are discussed, highlighting their ability to reproduce experimental observations. Finally, conclusions and further developments are drawn in Section 3.

2. Model formulation

In this section, we outline the essential modeling framework that forms the basis for our Deep Rheological Element. Our work builds upon the well-established generalized Maxwell model (GMM), which we extend through the integration of deep neural networks. While we provide only the key elements of the formulation here, a complete theoretical derivation based on the Principle of Virtual Power and the Clausius–Duhem inequality is presented in Appendix A. The GMM represents the mechanical response of materials through an assemblage of rheological elements, typically consisting of springs and dashpots. For viscoelastic materials under finite strain, we employ a generalized Maxwell model (Fig. 1), where the deformation gradient $\mathbf{F}$ is multiplicatively decomposed into elastic and viscous components:

$$\mathbf{F} = \mathbf{F}_e \mathbf{F}_v \tag{2.1}$$

This decomposition extends the additive decomposition found in linear viscoelasticity to finite strains, where $\mathbf{F}_e$ represents the elastic deformation that instantaneously unloads the current configuration to a stress-free state, and $\mathbf{F}_v$ represents the viscous contribution.

To derive the constitutive equation of this viscoelastic solid, we start by assuming the existence of a specific (per unit mass) strain energy density, denoted as $\hat{\varphi}$, which is dependent on both the elastic deformation $\mathbf{B}_e = \mathbf{F}_e \mathbf{F}_e^T$ and the total deformation $\mathbf{B} = \mathbf{F}\mathbf{F}^T$. Guided by the generalized Maxwell model of Fig. 1, we assume that there are two components capable of storing elastic energy during a deformative process:

- an equilibrium network represented by a hyperelastic spring with strain energy density $\varphi_\infty(\mathbf{B})$, which determines the long-term equilibrium response of the material;

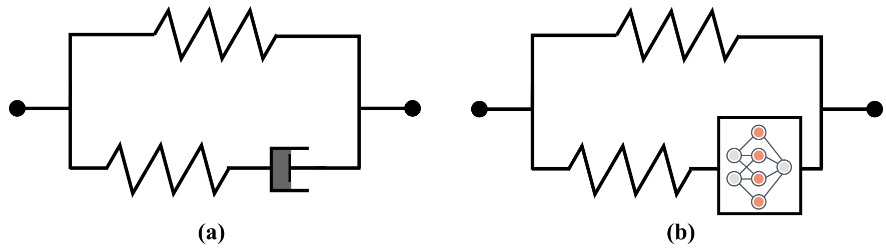

Fig. 1. (a) Assemblage of a Maxwell element in parallel with an equilibrium spring. (b) Proposed framework in which the inelastic element is substituted by Deep Neural Networks.

- a non-equilibrium network consisting of a Maxwell element, with a hyperelastic spring energy density $\varphi(\mathbf{B}_e)$ in series with a viscous dashpot.

that in turn yields the following additive decomposition of the energy:

$$\hat{\varphi}(\mathbf{B},\mathbf{B}_e) = \varphi_\infty(\mathbf{B}) + \varphi(\mathbf{B}_e)\,. \tag{2.2}$$

Accordingly, the total Cauchy stress $\mathbf{T}$ is given by the sum of contributions from both networks (derivations in Appendix), each defined in terms of the corresponding strain energy densities:

$$\mathbf{T} = \mathbf{T}_\infty + \mathbf{T}_e, \qquad \mathbf{T}_\infty = 2\rho_t\,\frac{\partial\varphi_\infty}{\partial\mathbf{B}}\mathbf{B}, \quad \mathbf{T}_e = 2\rho_t\,\frac{\partial\varphi}{\partial\mathbf{B}_e}\mathbf{B}_e\,. \tag{2.3}$$

where ρ_t is the mass density in the current configuration. The thermodynamic consistency requirement, expressed through the Clausius–Duhem inequality, can be used to introduce a constitutive equation for the stresses that guarantees positive dissipation for every strain rate. Most models commonly employed in the literature (see Califano and Ciambella, 2023; Gouhier and Diani, 2024) assume the specific dissipation density to be a quadratic function of the rate of deformation. In this case one may write

$$\rho_t\,\delta = \frac{1}{4}J^{-1}\,\mathbb{D}[\mathrm{sym}(\mathbf{B}_e^\triangledown\mathbf{B}_e^{-1})]\cdot\mathrm{sym}(\mathbf{B}_e^\triangledown\mathbf{B}_e^{-1}) \geq 0 \tag{2.4}$$

where $J = \det\mathbf{F}$, and $\mathbf{B}_e^\triangledown := \dot{\mathbf{B}}_e - \mathbf{L}\mathbf{B}_e - \mathbf{B}_e\mathbf{L}^T$ is the "codeformational derivative" or *Oldroyd* derivative of $\mathbf{B}_e$. The tensor $\mathbb{D}$ is a fourth-order positive definite viscosity tensor, which for a isotropic material has the following representation formula:

$$\mathbb{D} = 2\,\eta_D\big(\mathbb{I} - \frac{1}{3}\mathbf{I}\otimes\mathbf{I}\big) + 3\,\eta_V\big(\frac{1}{3}\mathbf{I}\otimes\mathbf{I}\big) \tag{2.5}$$

Here, $\mathbb{I}$ is the fourth-order identity tensor, i.e., $\mathbb{I}[\mathbf{A}] = \mathbf{A}$, and $(\mathbf{I}\otimes\mathbf{I})\mathbf{A} = (\mathbf{I}\cdot\mathbf{A})\mathbf{I}$. The parameters η_D and η_V represent the deviatoric and volumetric viscosities (per unit of reference volume), which may depend on the elastic or total deformation. Thermodynamic consistency is automatically satisfied if these viscosity functions are positive:

$$\eta_D \geq 0, \qquad \eta_V \geq 0 \tag{2.6}$$

By considering that elastomers and many other soft materials can be treated as incompressible ($J = \det\mathbf{F} = 1$ and $\det\mathbf{F}_e = 1$), the constitutive prescription (2.4) together with a microforce-like balance, derived in the Appendix, gives the evolution equation governing the viscous behavior:

$$\mathrm{sym}\{\mathbf{B}_e^\triangledown\mathbf{B}_e^{-1}\} = -\frac{1}{\eta_D}\mathbf{T}_e^D \tag{2.7}$$

in which $\mathbf{T}_e^D$ is the deviatoric part of the elastic Cauchy stress tensor. In our previous work (Califano and Ciambella, 2023), we noted that, aside from assumptions on the specific energy density, the main differences among models for filled rubber in the literature lie in how the viscosity function is formulated. Phenomenological expressions are commonly used for η_D, which often struggle to capture complex behaviors like the Payne effect characteristic of filled elastomers. The distinctive aspect of our approach is how we model the viscosity function η_D, relying on a neural network to represent it.

Model-based approach. The diversity of viscosity function forms found in the literature is substantial, with dozens of models each taking either phenomenological considerations or drawing inspiration from micromechanical perspectives (Hurtado et al., 2013; Reese and Govindjee, 1998; Bergström and Boyce, 1998; Kumar and Lopez-Pamies, 2016; McLeish and Larson, 1998; Drozdov, 2004; Rubin and Papes, 2011; Yoshida and Sugiyama, 2015; Kim et al., 2022; Gouhier and Diani, 2024). Elastomers, in particular, are noted for exhibiting deformation-induced shear thinning, leading to a viscosity function that varies dynamically-increasing with the magnitude of applied deformation and decreasing with the rate of deformation. This behavior contrasts sharply with the constant viscosity assumption seen in some models, like that in Reese and Govindjee (1998). In addressing this variable viscosity, the authors of Kumar and Lopez-Pamies (2016) introduce a function that escalates with applied deformation and diminishes with deformation rate. The model in Bergström and Boyce (1998) suggests that viscosity hinges on the magnitude of deviatoric Cauchy stress, a measure that indirectly accounts for strain rate, positing that macromolecules are capable of significant conformational adjustments under creep loading.

The encapsulated summary of these models in Table 1 serves as a concise reference, delineating the foundational assumptions and functional forms of η_D across different theoretical frameworks. Typically, the deviatoric viscosity η_D is expressed as a function of the total deformation tensor $\mathbf{B}$ and the elastic deformation tensor $\mathbf{B}_e$. However, considerations of frame-invariance, isotropy, and incompressibility imply that η_D can be represented in terms of an appropriate set of four scalar invariants of the unimodular tensors $\mathbf{B}$ and $\mathbf{B}_e$, leading to the general formulation:

$$\eta_D = \hat{\eta}_D(\mathcal{I}_1,\mathcal{I}_2,\mathcal{I}_3,\mathcal{I}_4). \tag{2.8}$$

In the existing literature, this general formulation is often simplified by considering the viscosity as a function of a subset of these scalar invariants. Specifically, common assumptions include defining the viscosity in terms of the invariant $\mathcal{I}_1 = \mathbf{B}\cdot\mathbf{B}_e^{-1}$, denoted as I_1^v, the invariant $\mathcal{I}_2 = \mathbf{I}\cdot\mathbf{B}_e^{-1}$, which corresponds to I_2^e, and the invariant $\mathcal{I}_3 = \|\mathbf{T}_e^D\|$, commonly referred to as J_2. The fourth invariant, $\mathcal{I}_4$, is not explicitly employed in the existing literature, yet different choices can be made; for example, we can define:

$$\mathcal{I}_4 = \frac{1}{2}\left(\left(\mathbf{B}\cdot\mathbf{B}_e^{-1}\right)^2 - \mathbf{B}\mathbf{B}_e^{-1}\cdot\mathbf{B}_e^{-1}\mathbf{B}\right).$$

In our previous work (Califano and Ciambella, 2023), we presented a comparative analysis of the models in Table 1.

Data-driven approach. In this paper, we adopt a data-driven approach to infer the viscosity function η_D that characterizes the time-dependent response of materials using a Deep Neural Network (DNN). Specifically, we propose that for an incompressible material, the deviatoric viscosity η_D can be expressed as:

$$\eta_D = f_{\mathrm{DNN}}(\mathbf{x};\theta) \tag{2.9}$$

where f_{DNN} represents the Deep Neural Network (DNN) with input features $\mathbf{x}$, and θ denotes the network's parameters. In this hybrid modeling approach, $\mathbf{x}$ can either be one or a subset of the invariants

Table 1
Viscoplastic/viscoelastic models from the literature incorporated in the proposed framework.
Source: From Califano and Ciambella (2023).

Constitutive model	Viscosity η_D
Reese & Govindjee (Reese and Govindjee, 1998)	Constant
Bergström & Boyce (Bergström and Boyce, 1998; Bergstrom and Boyce, 2000)	$J_2^{1-m}/A(\lambda^v - 1 + \xi)^C$ with $\lambda^v = \sqrt{I_1^v/3}$
Kumar & Lopez-Pamies (Kumar and Lopez-Pamies, 2016)	$\eta_\infty + \dfrac{\eta_0 - \eta_\infty + K_1(I_1^{v\,\beta_1} - 3^{\beta_1})}{1 + (K_2 J_2^2)^{\beta_2}}$
Strain Hardening Power Law (Hurtado et al., 2013)	$J_2/\dot{\varepsilon}^{cr}$ with $\dot{\varepsilon}^{cr} = \left(A J_2^n\,[(m+1)\bar{\varepsilon}^{cr}]^m\right)^{1/(m+1)}$
Drozdov (Drozdov, 2004) and Pom-Pom model (McLeish and Larson, 1998)	$c/\sqrt{2}\lVert\mathbf{D}\rVert$
Rubin & Papes (Rubin and Papes, 2011)	$(4/3)(c/\Gamma)I_2^e$

$\mathcal{I}_1, \mathcal{I}_2, \mathcal{I}_3$, or $\mathcal{I}_4$. The hyperelastic part of the model is defined through standard strain energy densities. The underlying idea is that the strain energy density φ can be easily fitted using readily available experimental data, whereas, accurately capturing the nonlinearities of the viscosity function presents a significant challenge. Additionally, within this approach, thermodynamic consistency is ensured by imposing the simple requirement that the viscosity function remains positive.

In what follows, we assume a neo-Hookean type strain energy function for both the equilibrium and non-equilibrium springs, given by:

$$\rho_0\,\varphi_\infty(\mathbf{B}) = \frac{\mu_\infty}{2}(\mathbf{I}\cdot\mathbf{B} - 3), \quad \rho_0\,\varphi(\mathbf{B}_e) \quad = \frac{\mu}{2}(\mathbf{I}\cdot\mathbf{B}_e - 3)\,. \tag{2.10}$$

where $\rho_0 = J\rho_t$ is the reference density.

2.1. Architecture selection: Feedforward neural network (FNN)

In the proposed computational framework, Feedforward Neural Networks (FNNs) are utilized for modeling the viscosity function. Feedforward Neural Networks (FNNs) are a class of artificial neural networks where information flows in a single direction, from input to output, without loops or cycles. By incorporating the generalized Maxwell model's ordinary differential equations (ODEs), this hybrid framework captures temporal dependencies, eliminating the need of Recurrent Neural Networks (RNNs) (Chen, 2021; Wang et al., 2023; Holthusen et al., 2024) or Long Short-Term Memory (LSTM) networks (Zhang et al., 2022), resulting in lower data-intensive requirements (Goodfellow et al., 2016; LeCun et al., 2015). The architecture of the FNN used in this paper has the following structure:

$$f_{\mathrm{DNN}}(x; W, b) = \frac{1}{2}\left(1 + \tanh\left(\ldots\tanh\left(W_1 x + b_1\right)\ldots\right)\right) W_f^2 + b_f^2 \tag{2.11}$$

By squaring the final weights and biases, and applying a shifted tanh activation only at the output layer, the network ensures a strictly positive output, thereby guaranteeing thermodynamic consistency through non-negative viscosity (see Fig. 2).

2.2. Data preparation and training

A fundamental challenge in modeling the viscosity function with neural networks is that viscosity measures are not directly available from mechanical testing but must instead be derived from time-dependent stress–strain data. Consequently, traditional supervised learning approaches, which depend on labeled training data, are not applicable to viscosity. To address this challenge, this paper adopts a two-step training strategy, leveraging the concept of transfer learning (Pan and Yang, 2010).

In the initial phase of the process, the function f_{DNN} is pre-trained using viscosity data obtained from established rheological models like (Bergström and Boyce, 1998) or (Kumar and Lopez-Pamies, 2016), which have been calibrated against real experiments. Typically, experimental data on viscoelastic materials is acquired through dynamic mechanical analysis (DMA) tests, and the outcomes are represented synthetically by storage and loss moduli (Ciambella et al., 2011), rather than time data. Thus, the pre-training phase enables the generation of additional data points, thereby increasing the density and coverage of the dataset.

A second step refines the model's predictive accuracy through targeted retraining based on minimizing the distance difference between stress histories as predicted by the DNN and the experimental measurements.

Experimental protocol

The data used in both the pre-training and re-training phases are taken from the experimental results reported by Chazeau et al. (2000), specifically those corresponding to dynamic mechanical analysis (DMA) amplitude sweep tests conducted under simple shear. In these tests, the material is subjected to a shear strain of the form:

$$\gamma(t) = \gamma_0 + \gamma_d \sin(2\pi f_d t) \tag{2.12}$$

where $f_d = 1\,\mathrm{Hz}$ is the constant imposed frequency, while γ_d denotes the dynamic amplitude, which is systematically varied to perform an amplitude sweep. The corresponding shear stress $T_{12}(t)$ is recorded over multiple cycles and subsequently processed to derive the storage G' and loss G'' moduli, commonly used to characterize the material response in DMA testing. These moduli are defined as:

$$\begin{aligned} G' &= \frac{2\,f_d}{N}\int_{t_i}^{t_i+N/f_d} \frac{T_{12}(t)}{\gamma_\mathrm{d}}\sin(2\pi\,f_d\,t)\,dt \\ G'' &= \frac{2\,f_d}{N}\int_{t_i}^{t_i+N/f_d} \frac{T_{12}(t)}{\gamma_\mathrm{d}}\cos(2\pi\,f_d\,t)\,dt \end{aligned} \tag{2.13}$$

Here, t_i denotes the starting time of the integration window, chosen to ensure that the material's response is measured in the steady-state regime, while N represents the number of cycles considered in the averaging process. The experimental G' and G'' in Fig. 5 exhibit a pronounced dependence on the dynamic amplitude, particularly noticeable in the storage modulus, which experiences over a tenfold decrease from the maximum (0.2) to the minimum (0.001) dynamic strain amplitudes investigated, clearly illustrating the Payne effect. Notably, a constant viscosity model, such as the one used in Reese and Govindjee (1998), would predict a much weaker dependence on the strain amplitude and thus fails to capture this nonlinear viscoelastic response, as also discussed in Califano and Ciambella (2023).

The experimental setup, consisting of symmetric shear sandwich fixtures with an aspect ratio of approximately 8:1, ensures true simple shear conditions, significantly minimizing nonuniform strain effects. This approach aligns with our previous validation work (Califano and Ciambella, 2023), in which a semi-analytical solution derived from a non-linear viscoplastic model, restricted to simple shear kinematics, was benchmarked against full 3D finite element simulations under identical loading and boundary conditions. The identification of the model parameters in the present work follows this semi-analytical strategy, thereby avoiding the computational burden of full-field finite element simulations. Key analytical steps and expressions of the solution strategy are summarized in Appendix A.

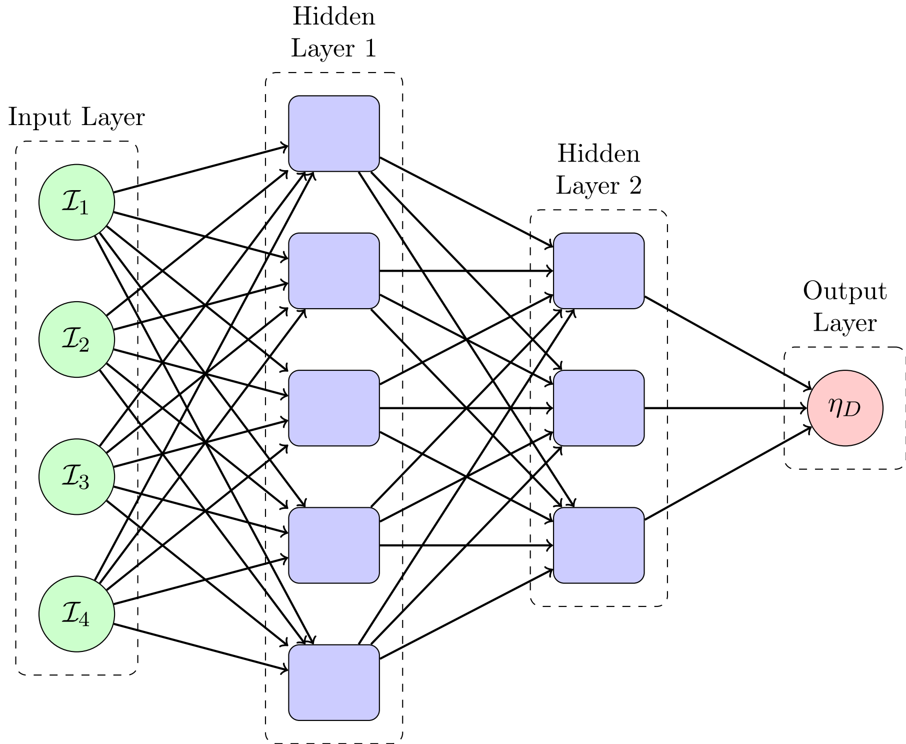

Fig. 2. Example of the architecture of a Feedforward Neural Network (FNN) for predicting the viscosity function η_D. The input layer consists of the scalar invariants $\mathcal{I}_1, \mathcal{I}_2, \mathcal{I}_3, \mathcal{I}_4$, as required by the considerations of frame-invariance, isotropy, and incompressibility. These invariants provide inputs to two fully connected hidden layers, leading to the output η_D. In the present study, however, only the third invariant $\mathcal{I}_3 = J_2$ is used as input, resulting in a simplified architecture with a single input neuron, two hidden layers with five neurons each, and one output neuron, corresponding to a 1-5-5-1 configuration.

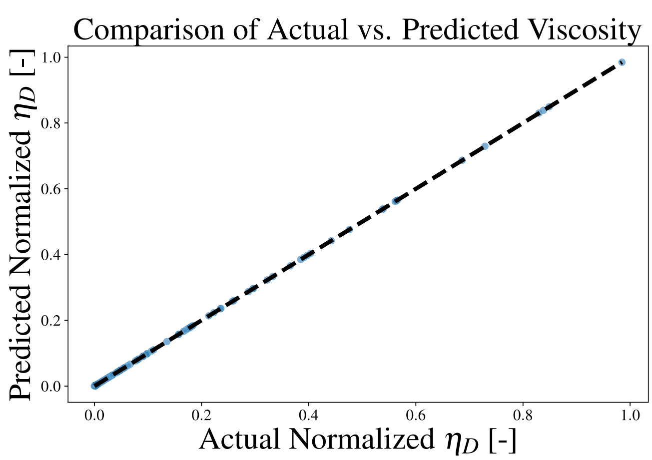

Fig. 3. Comparison of actual (blue dots) versus predicted viscosity on test data. The dashed black line represents the ideal identity line where prediction perfectly matches the actual values. (For interpretation of the references to color in this figure legend, the reader is referred to the web version of this article.)

Pre-training

The pre-training phase involves generating pseudo-experimental viscosity data using the pre-calibrated Strain Hardening Power Law (SHPL) model (Hurtado et al., 2013). Among the various models proposed in Table 1, the SHPL is selected for its simplicity and its ability to capture the Payne's effect with a minimal number of parameters. The viscosity function for this model is defined in Table 1 and A, m, and n are constitutive parameters to be calibrated against experimental data. As customary in literature, the parameter m is set to 0, reducing the number of constitutive parameters to four:

Table 2
Parameter values.

Parameter	Value
μ_∞	0.5 MPa
n	2.45
μ	3.55 MPa
A	100 MPa^{-n}s^{-m-1}

$$\eta_D = J_2/\dot{\varepsilon}^{cr} \text{ with } \dot{\varepsilon}^{cr} = A J_2^n \tag{2.14}$$

Although the pre-training is carried out with a model that only incorporates the $\mathcal{I}_3 = J_2$ dependence on viscosity, the DNN approach pursued here is capable of capturing more complex dependencies, such as those used in Bergström and Boyce (1998), Kumar and Lopez-Pamies (2016), or on temperature or magnetic field, which could significantly affect the material response (Cleément et al., 2005; Si et al., 2023). The SHPL model was calibrated using data described in previous subsection on experimental protocol and presented in Fig. 7. Model calibration was performed following the approach described in Califano and Ciambella (2023). The resulting parameters are listed in Table 2, whereas the model's performance is illustrated in Fig. 7. Although the SHPL model can closely match the storage modulus' dynamic amplitude dependence, it overpredicts the variation of the loss modulus in the same dynamic amplitude range.

The dataset used for training and analysis spans the full range of J_2 values observed during experimental testing. In this context, J_2, representing the norm of the non-equilibrium stress, is not directly measurable, and was estimated using the calibrated SHPL model. This approach enabled the generation of input–output pairs, where J_2 serves as the input and the corresponding viscosity η_D as the output. We generated 1000 such pairs spanning a J_2 range from 0.001 to 0.2.

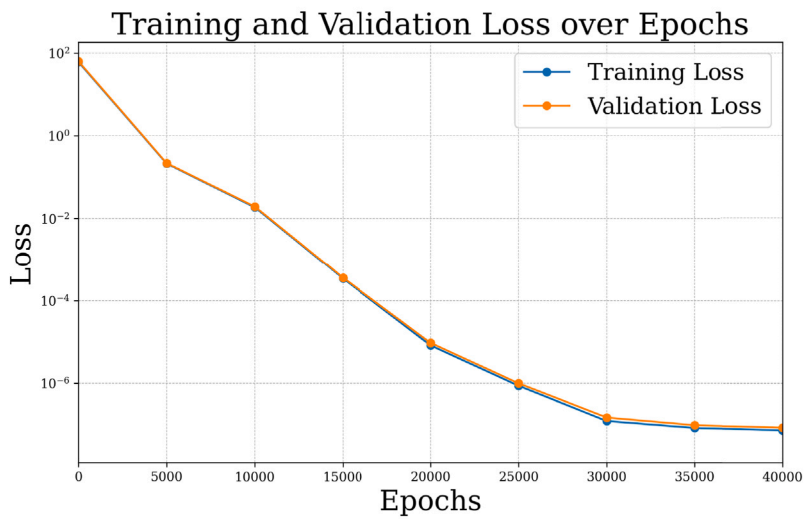

Fig. 4. Training and Validation Loss in logarithmic scale over first 40000 Epochs, computed using the loss function defined in Eq. (2.15).

The DNN used for the training of the selected dataset consists of an input layer with one neuron, followed by two hidden layers with five neurons each, and a final output layer with a single neuron, corresponding to a 1-5-5-1 configuration. The network parameters (weights and biases) were initialized by sampling from a standard normal distribution. The use of a single input node reflects the power law formulation, which depends exclusively on J_2. The training strategy was carried out by squaring J_2 values and log-transforming, with a small constant added to avoid issues from zeros. This transformation stabilized the input distribution across magnitudes, facilitating smoother convergence during training. We then applied Z-score normalization to the log-transformed J_2^2 values, centering them around a mean of 0 and scaling to unit variance. For the η values, we used Min-Max normalization: after identifying the minimum and maximum values, the data were linearly scaled to the range [0, 1]. This choice preserved the positivity of the target values, consistent with their physical meaning, and ensured compatibility with the neural network architecture, which generates positively encoded outputs (LeCun et al., 1998; Schraudolph, 1998).

The training process utilized a loss function based on the mean squared error (MSE), quantifying the discrepancy between the SHPL viscosity and the DNN prediction.

$$\mathrm{MSE} = \frac{1}{N}\sum_{i=1}^{N}\left(\eta_i^{\mathrm{SHPL}} - \eta_i^{\mathrm{DNN}}\right)^2 . \tag{2.15}$$

where N is the number of data points in the training set. A training over 100,000 epochs with a learning rate of 0.001 was conducted. The dataset underwent a division into training and test subsets, with 70% of the data allocated for training and the remaining 30% reserved for validation and testing. Additionally, the remaining data was further split into validation and test sets in a 50–50. The validation set was used to monitor performance during training and prevent overfitting, while the test set was reserved for a final evaluation of the model's generalization capabilities, ensuring reliable performance assessment on both seen and unseen data. Upon completing the training process and evaluating the model on the test set, the model exhibited a test loss of $2.69 \cdot 10^{-8}$. This low loss value indicates that the predicted viscosity closely matches the target viscosity, as also evidenced in Fig. 3.

Fig. 4 showcases the trend of training and validation losses throughout the epochs. The initial rapid decrease in both training and validation losses indicates a significant learning uptake by the network. As epochs progress, the loss values plateau, reflecting the optimization of the network's weights and biases. The alignment of the validation loss with the training loss suggests that the model has achieved a good fit to the data without overfitting, confirmed by the consistently low test loss reported.

Algorithm 1: Pre-training procedure for DNN viscosity prediction

Input: Experimental storage and loss modulus data, calibrated SHPL model
Output: Pre-trained DNN model predicting viscosity η_D

```
Step 1: Data Generation
    Calibrate the SHPL model on experimental data.
    Use the calibrated SHPL model to generate synthetic (J_2, η_D)
      pairs within the experimental J_2 range.
Step 2: Data Pre-processing
    Apply log-transformation to J_2^2 values, followed by Z-score
      normalization.
    Apply Min-Max normalization to η_D values.
Step 3: Neural Network Setup
    Define a feedforward neural network architecture.
    Initialize weights and biases from a standard normal
      distribution.
Step 4: Training
    Split the dataset into training, validation, and test sets.
    Initialize the optimizer with learning rate.
    for epoch = 0 to 100,000 do
        Perform forward pass to predict viscosity η_D.
        Compute the MSE loss between predicted and target
          viscosities.
        Compute gradients of the loss with respect to network
          parameters.
        Update network parameters using the optimizer.
    end
```

2.3. Re-training

After the pre-training phase, a re-training process was conducted to further enhance the DNN's performance. To increase the dataset's density, interpolation was performed on the storage and loss data, generating an additional 100 points within the experimental dynamic amplitude range as schematically represented in Fig. 5. For this purpose, spline interpolation was used, and the results were visually verified against expected rheological trends. Specifically, the storage modulus exhibited an initial plateau at low strain amplitudes and a

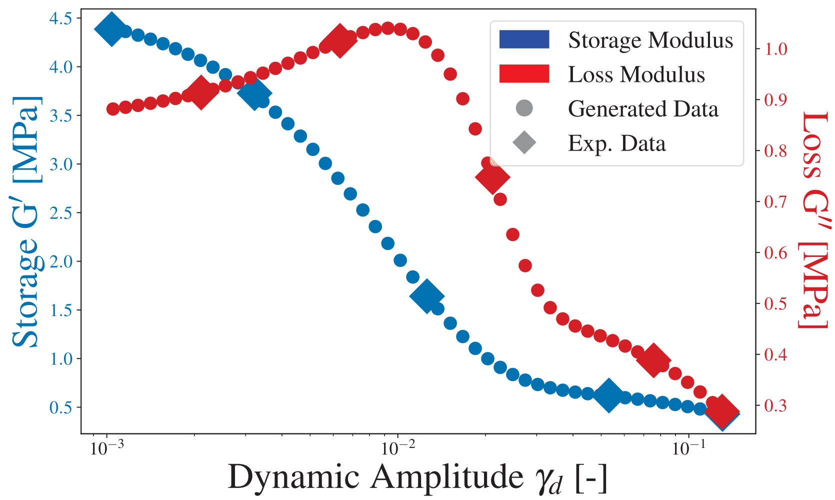

Fig. 5. Experimental data points (diamonds ♦) from (Chazeau et al., 2000) used for the pre-training. Circles • represent the additional data points used for the re-training and generated from the original dataset through spline interpolation.

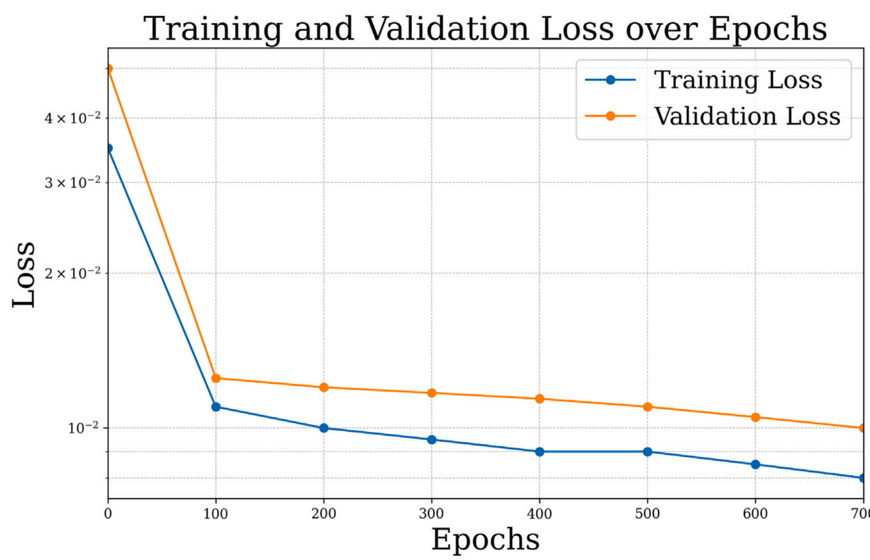

Fig. 6. Training and validation losses in logarithmic scale throughout the re-training epochs, computed using the loss function defined in Eq. (2.19).

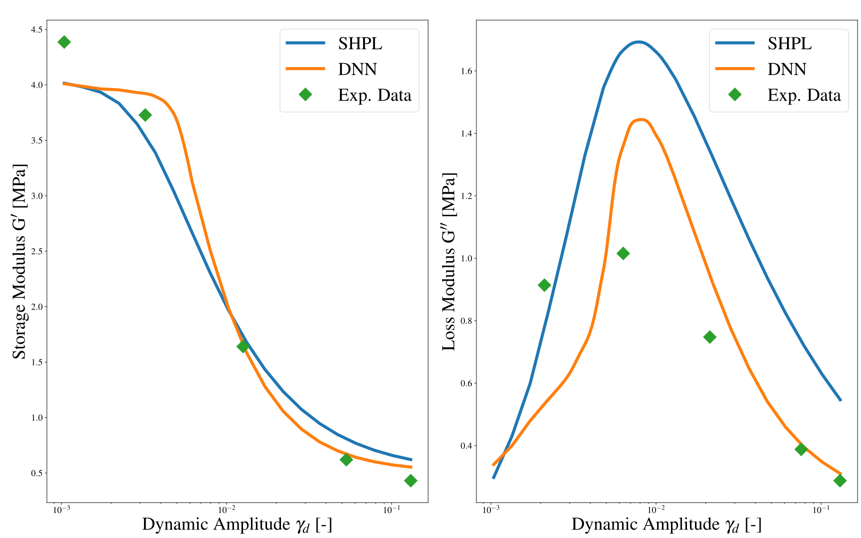

Fig. 7. Comparative analysis of storage modulus (left) and loss modulus (right) as functions of dynamic amplitude, predicted by SHPL and DNN models, against experimental data for filled silicone elastomer at 1 Hz, as reported by Chazeau et al. (2000).

subsequent decline, followed by a second plateau at higher amplitudes, while the loss modulus exhibited a peak, consistent with the characteristic viscoelastic behavior of such blends (Klüppel, 2008).

For each pair of storage and loss values corresponding to a specific dynamic amplitude, the time histories of the steady-state response $T_{12}(t)$ were reconstructed assuming a linear relationship between shear stress and shear strain history:

$$T_{12} = \mu_\infty \gamma_0 + \gamma_d G' \sin(2\pi f_d t) + \gamma_d G'' \cos(2\pi f_d t), \tag{2.16}$$

where 10 cycles at a frequency of 1 Hz were generated for each amplitude and subsequently employed in training. The reconstructed stress response was used as the target during the re-training phase to guide the neural network in predicting the viscosity.

Indeed, during re-training, the neural network input remains the scalar value J_2^2 (the squared norm of the stress deviator), which is log-transformed and normalized, consistent with the pre-training phase. The output is the predicted viscosity η_D, which is then integrated into the viscoelastic constitutive model, formulated in Section 2, to simulate the stress response, directly influencing the stress evolution obtained by solving the internal variable evolution equations (ODEs). Details of adapting this constitutive equation explicitly for simple shear conditions are provided in Appendix A. The loss function employed during re-training is based on a weighted mean squared error (MSE) between the stress histories predicted by the model $T_{12}^{\text{pred}}(t)$ and the reconstructed target histories $T_{12}^{\text{target}}(t)$. To ensure proper learning across the full spectrum of experimental conditions, particularly for low strain amplitudes, the weights are assigned inversely proportional to the dynamic stress amplitude of the ith target history τ_{d_i} (i.e., the amplitude of the oscillatory stress response), following the rationale that stress histories with smaller input amplitudes, typically harder to predict and prone to higher relative error, would otherwise contribute negligibly to the total loss. Specifically, for each loading history, the individual mean squared error MSE_i is:

$$\text{MSE}_i = \frac{1}{N_t} \sum_t \left(T_{12_i}^{\text{pred}}(t) - T_{12_i}^{\text{target}}(t) \right)^2 \tag{2.17}$$

computed over the last two cycles of the ith loading history. Here, N_t denotes the number of time steps within the interval. The overall weighted loss is given by:

$$\text{MSE}_{\text{retrain}} = \frac{\sum_{i=1}^n w_i \cdot \text{MSE}_i}{\sum_{i=1}^n w_i}, \qquad \text{with} \quad w_i \propto \frac{1}{\tau_{d_i}} \tag{2.18}$$

To allow for a dimensionless and scale-invariant evaluation, the resulting loss is normalized by the mean squared value of the target shear stress $T_{12}(t)$ within the same interval:

$$\text{W-MSE}_{\text{retrain}} = \frac{\text{MSE}_{\text{retrain}}}{\text{MST}(T_{12}) + \varepsilon}, \qquad \text{with} \quad \text{MST}(T_{12}) = \frac{1}{nN_t} \sum_{i=1}^n \sum_t (T_{12_i}(t))^2 \tag{2.19}$$

where n denotes the number of loading histories used in the retraining. The parameter ε is a small constant added to avoid division by near-zero values, thus preserving numerical stability.

The re-training process utilized an Adam optimizer (Kingma and Ba, 2014). The learning rate was set to 0.01. The available histories and test types were partitioned following the same strategy used during pre-training. Initially, a 70–30 split was performed to separate the training set from the combined validation and test set. Subsequently, the validation and test sets were obtained by evenly dividing the remaining data. Within the training loop, a patience-based early stopping mechanism was integrated to prevent overfitting, ensuring the robustness and generalization capability of the trained model (Prechelt, 2012). The model was trained iteratively, with periodic evaluations on the validation set.

Fig. 6 illustrates the trajectory of training and validation losses throughout the re-training epochs. An initial steep decline in both

Algorithm 2: Re-training procedure for improving DNN model performance

Input: Pre-trained DNN model, experimental storage and loss modulus data
Output: Re-trained DNN model predicting viscosity η_D

Step 1: Data Augmentation
  Interpolate storage and loss modulus data to generate additional points within the experimental dynamic amplitude range.
  For each interpolated point, reconstruct the corresponding stress history $T_{12}(t)$ over 10 cycles at 1 Hz, assuming a linear stress–strain relationship.

Step 2: Initialization
  Load pre-trained network parameters.
  Initialize the optimizer with learning rate.
  Partition data into training, validation, and test sets.

Step 3: Re-training Loop
  while *training has not stopped early and epoch count is below maximum* **do**
    Process the full batch of training histories.
    Generate the dynamic shear strain history $\gamma(t)$ based on test parameters.
    Predict viscosity η_D from normalized J_2^2 inputs.
    Inverse-normalize viscosity predictions.
    Simulate stress response $T_{12}^{\text{pred}}(t)$ by solving the viscoplastic constitutive model ODEs.
    Compute the weighted and normalized loss $\text{W-MSE}_{\text{retrain}}$ over the last two cycles.
    Compute gradients of the loss with respect to network parameters.
    Update network parameters using the Adam optimizer.
  end

training and validation losses is evident, signifying substantial improvement in the network's performance. As re-training progresses, the losses begin to level off, indicating that the network is approaching an optimized state where further adjustments to weights and biases yield diminishing returns in loss reduction. The convergence of training and validation losses, remaining close to each other, indicates that the re-training process has enhanced the model's generalization abilities as also demonstrated by the fitting results in Fig. 7. All computations, including model training, optimization, and differential equation solving, were implemented using `Python`, leveraging `JAX` for array operations, automatic differentiation, and just-in-time (JIT) compilation (Bradbury et al., 2018), `Optax` for optimization routines (DeepMind I. Babuschkin et al., 2020), and `Diffrax` for solving differential equations (Kidger, 2021). For all training and evaluation phases, time-integration was carried out with Dopri5 solver (Dormand-Prince 5th order).

The deep neural network (DNN) model demonstrates comparable performance to the Strain Hardening Power Law (SHPL) model in matching the storage modulus, but significantly outperforms it in predicting the loss modulus. This improvement is evident from the two-fold decrease in the error associated with the loss modulus, as reported in Table 3. However, it is noteworthy that the initial loss value is still underestimated, suggesting that both models fail to capture the energy losses of the real material at low strain rates. To highlight the effect of the re-training on the viscosity function, Fig. 8 compares the deviatoric viscosity $\eta_D(J_2)$ obtained with the SHPL model and with the DNN after re-training (log–log scale). The DNN deviates from the simple power law in the regions of J_2 explored in the DMA tests, which provides a better description of the shear-thinning behavior associated with

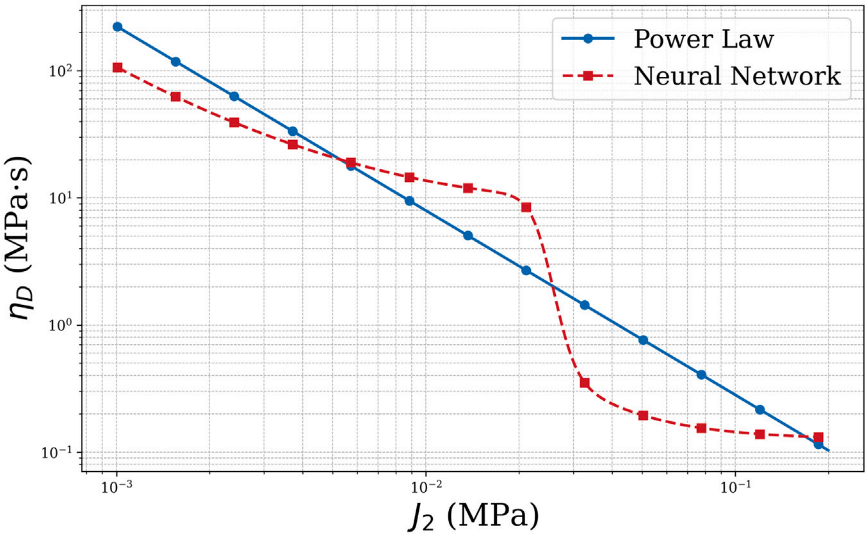

Fig. 8. Deviatoric viscosity η_D as a function of J_2 after re-training: comparison between the SHPL model and the DNN (log–log scale). The adaptation of the DNN in the range of J_2 relevant to DMA tests explains the improved prediction of the loss modulus.

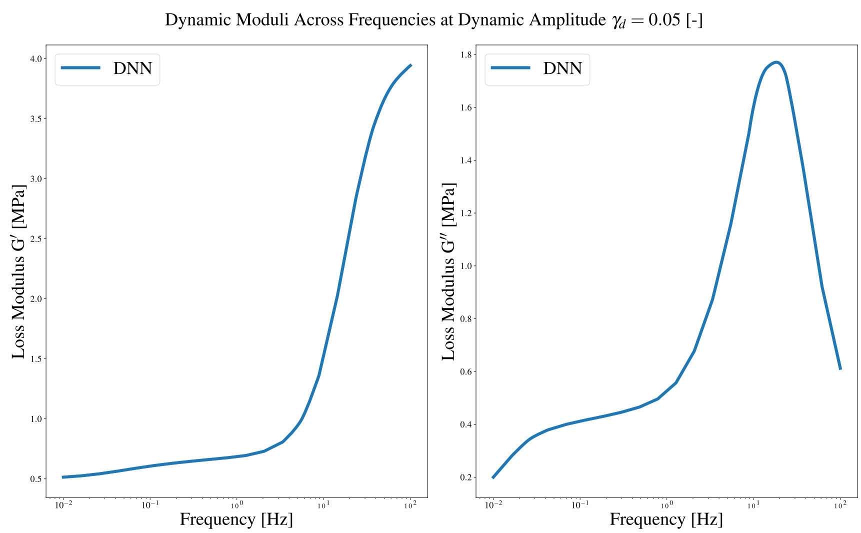

Fig. 9. Frequency extrapolation at a fixed dynamic amplitude of $\gamma_d = 0.05$ for (a) storage and (b) loss moduli.

Table 3
Comparison of Mean Relative Errors (MRE) for Deep Neural Network and Strain Hardening Power Law Predictions.

Parameter	DNN	SPHL
MRE for Storage	0.1219	0.1349
MRE for Loss	0.2701	0.6556

the Payne effect. Importantly, positivity of η_D is preserved, ensuring thermodynamic consistency.

To further evaluate the DNN's performance, we assess its extrapolation capabilities beyond the experimental range of the fitting data. This evaluation is crucial for ensuring the model's physical behavior and reliability. Extrapolation tests help verify whether the model can generalize its learned patterns and relationships to unseen data points or scenarios. In this regard, Fig. 9 presents the predicted storage and loss moduli over a four order of magnitude frequency range. The DNN predicts a storage modulus that increases with frequency, leveling off around 100 Hz, and a loss modulus that reaches a maximum value around 10 Hz, which aligns with expectations for this type of material (Chazeau et al., 2000; Klüppel, 2008).

Additional testing was conducted by imposing the cyclic strain history depicted in Fig. 10(a), comprising five cycles at 5% strain and an additional five cycles at 10% strain. The stress–strain plot (Fig. 10(c)) reveals hysteresis loops characteristic of viscoelastic materials, which reach equilibrium after few initial cycles. This step aims to assess the DNN's capability to accurately simulate the material behavior under conditions not explicitly encountered during the training phase.

3. Conclusions

In this study, we have proposed a novel rheological element with a constitutive response based on a Deep Neural Network to enhance the constitutive modeling of elastomers under finite strains. While the use of an assemblage of standard rheological elements has been widely adopted for capturing the elastic and viscoelastic behavior of elastomers, its traditional approach relies on a phenomenological viscosity model for the dashpot-like element. However, this approach exhibits

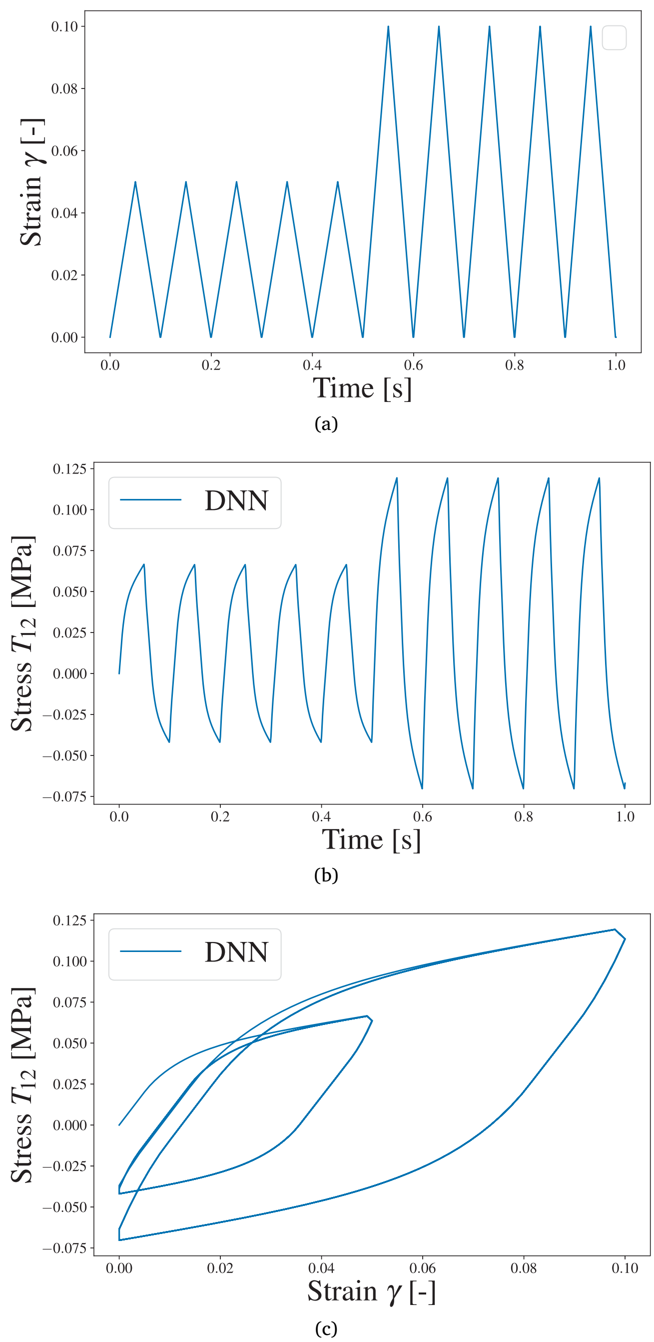

Fig. 10. Analysis of cyclic strain and material response: (a) Cyclic strain history with triangular waves at amplitudes 0.05 and 0.1 over a $0.1\,\mathrm{s}$ period. (b) Response from the DNN-calibrated constitutive model, depicting stationarity. (c) Material's hysteresis in the stress–strain plane, with stationarity achieved.

limitations in accurately capturing nonlinear effects observed in experimental data, such as the Payne effect. In our proposed approach, we replaced the phenomenological viscosity model with a deep neural network, resulting in what we term "Deep Rheological Element".

To train the neural network, we employed a two-step strategy aimed at alleviating the requirement for a large amount of available data. In the initial step, the viscosity neural network was trained with the viscosity function of a well-established rheological model, such as the strain hardening power law, previously calibrated against experimental data. This pre-training phase facilitated a rapid initialization of the DNN and ensured an accurate starting point. Subsequently, in the second step, we conducted a retraining of the network using stress–strain histories derived from dynamic mechanical analysis tests. To achieve this, we transformed data on storage and loss moduli into time–history data, which served as the actual training dataset. This approach allowed us to leverage a significant number of data points for training while utilizing only a limited number of real experimental tests.

The analysis demonstrated that the deep neural network model successfully replicated the performance of the strain hardening power law model, which is commonly employed for fitting this type of data. Remarkably, the DNN model even surpassed the performance of the strain hardening power law model after retraining. Additionally, the extrapolation capabilities of the DNN were evaluated, revealing a physically meaningful behavior beyond the fitting range. By integrating the DNN into an assemblage of rheological elements, the thermodynamic consistency of the model was easily ensured by constraining the DNN viscosity to be positive. This represents a significant advantage over other proposals in the literature.

It is important to note that for simplicity and clarity we initially limited our strain energy function to the Neo-Hookean type. However, our framework is designed with the flexibility to incorporate any constitutive model for the hyper-elastic springs, including NN-based model, e.g. Constitutive Artificial Neural Networks (CANN) (Linka et al., 2021; Linka and Kuhl, 2023), Input Convex Neural Networks (ICNN) (Klein et al., 2022; As'ad et al., 2022; Chen and Guilleminot, 2022), Physics-Augmented Neural Network (Linden et al., 2023) and Neural Ordinary Differential Equations (NODE) (Taç et al., 2022b).

CRediT authorship contribution statement

Federico Califano: Writing – review & editing, Writing – original draft, Visualization, Validation, Software, Methodology, Investigation, Formal analysis, Data curation, Conceptualization. **Jacopo Ciambella:** Writing – review & editing, Visualization, Validation, Supervision, Resources, Project administration, Methodology, Investigation, Funding acquisition, Formal analysis, Conceptualization.

Ethical approval

Not applicable.

Funding

The authors wish to express their gratitude for the financial support received for this research. FC is thankful for the support from the PON-MIUR "Ricerca e Innovazione" 2014–2020, as part of the project "Experimental and numerical characterization of rubber compounds at high frequency". Both JC and FC would like to extend their appreciation to the Italian National Group of Mathematical Physics (GNFM-INdAM) for their backing.

Declaration of competing interest

The authors declare the following financial interests/personal relationships which may be considered as potential competing interests: Federico Califano and Jacopo Ciambella have patent pending related to the subject matter of this work.

Appendix A. Theoretical foundations of the generalized maxwell model

In this appendix, we provide the complete theoretical derivation of the modeling framework presented in Section 2. Our derivation relies on two fundamental principles: (i) the Principle of Virtual Power, which enables the introduction of the balance laws of the theory; (ii) the Clausius–Duhem inequality, which allows the definition of constitutive prescriptions in a manner that ensures thermodynamic consistency and stability in a iso-thermal framework.

A.1. Kinematics

We define a point x in the present configuration Ω_t as the image of a point X in the reference configuration Ω_0 attained at a certain time t_0. We call χ the deformation map such that $x = \chi(X,t)$, and the deformation gradient is

$$\mathbf{F}(X,t) = \frac{\partial \chi(X,t)}{\partial X}, \tag{A.1}$$

in a way that $\det\mathbf{F}(X,t) > 0$. We further assume that at each material point, the deformation gradient can be multiplicatively decomposed into elastic and viscous components by

$$\mathbf{F}(X,t) = \mathbf{F}_{e_i}(X,t)\mathbf{F}_{v_i}(X,t), \qquad i = 1,\dots,n \tag{A.2}$$

which is referred to as *Bilby–Kröner–Lee* decomposition (Sadik and Yavari, 2017; Ciambella and Nardinocchi, 2023). The Eq. (A.2) extends the additive decomposition typically found in linear viscoelasticity, as observed in the Maxwell rheological element (Riande et al., 2000), and enables the introduction of n states. Each state may represent different structural components, such as polymer chains with different relaxation times. Accordingly, $\mathbf{F}_{e_i}$ is the elastic deformation that can instantaneously unload the current configuration into a stress-free state. For the sake of brevity, we will hereafter omit the subscript i, assuming the existence of only one Maxwell element as sketched in Fig. 1, yet all derivations can be extended to multiple rheological elements straightforwardly. Additionally, we will refrain from indicating dependence on position and time unless explicitly required.

With the multiplicative decomposition in (2.1) on hands, one can define the strain rate, the elastic strain rate and the viscous strain rate by

$$\mathbf{L} = \nabla\mathbf{v} = \dot{\mathbf{F}}\mathbf{F}^{-1}, \quad \mathbf{L}_e = \dot{\mathbf{F}}_e\mathbf{F}_e^{-1}, \quad \mathbf{L}_v = \dot{\mathbf{F}}_v\mathbf{F}_v^{-1}. \tag{A.3}$$

By push-forwarding $\mathbf{L}_v$ with $\mathbf{F}_e$, one obtains the *Eulerian* viscous rate defined by

$$\boldsymbol{\Lambda}_v = \mathbf{F}_e\,\mathbf{L}_v\,\mathbf{F}_e^{-1} \tag{A.4}$$

which, together with (A.3) leads to the following additive decomposition of the strain rate $\mathbf{L}$

$$\mathbf{L} = \mathbf{L}_e + \boldsymbol{\Lambda}_v\,. \tag{A.5}$$

Additionally, the rate of the left Cauchy–Green strain tensors $\mathbf{B}_e = \mathbf{F}_e\mathbf{F}_e^T$ is computed through Eq. (A.2) as

$$\dot{\mathbf{B}}_e = 2\,\mathrm{sym}\{\mathbf{L}\mathbf{B}_e\} - 2\,\mathbf{F}_e\mathbf{D}_v\mathbf{F}_e^T\,, \tag{A.6}$$

where $\mathrm{sym}\{\cdot\}$ indicates the symmetric operator, such that $\mathrm{sym}\{\mathbf{A}\} = \frac{1}{2}(\mathbf{A}^T + \mathbf{A})$, such that $\mathbf{D} = \mathrm{sym}\{\mathbf{L}\}$, $\mathbf{D}_v = \mathrm{sym}\{\mathbf{L}_v\}$.

A.2. Principle of virtual power

The principle of virtual power is employed to derive the balance equations and the stress tensor restrictions imposed by the requirement of frame invariance. We formulate an Eulerian version of the theory by assuming $\{\tilde{\mathbf{v}}, \tilde{\boldsymbol{\Lambda}}_v\}$ the set of virtual velocities such that the internal power expenditure over a part $\mathcal{P} \subset \Omega_t$ is

$$\mathcal{W}_{int} = \int_{\mathcal{P}} \mathbf{T}_\infty \cdot \tilde{\mathbf{L}} + \int_{\mathcal{P}} \mathbf{T}_e \cdot \tilde{\mathbf{L}}_e + \int_{\mathcal{P}} \mathbf{T}_v \cdot \tilde{\boldsymbol{\Lambda}}_v \tag{A.7}$$

where $\mathbf{T}_\infty$ denotes the stress associated with the long-term spring that expends power on the total stretch rate, $\mathbf{T}_e$ represents the Maxwell element spring stress conjugates of the elastic stretch rate, and $\mathbf{T}_v$ represents the Maxwell element dashpot viscous stress tensor. It is noted that, for notational conciseness, the dV in the integral is omitted, with the limits of integration allowing us to distinguish between volume and surface integrals.

By using the kinematic relationship (A.5), the inner power can be rewritten as

$$\mathcal{W}_{int} = \int_{\mathcal{P}} \mathbf{T} \cdot \tilde{\mathbf{L}} + \int_{\mathcal{P}} \left(\mathbf{T}_v - \mathbf{T}_e\right) \cdot \tilde{\boldsymbol{\Lambda}}_v \tag{A.8}$$

where $\mathbf{T} = \mathbf{T}_\infty + \mathbf{T}_e$. The external power is the power expended by body $\mathbf{b}$ and surface $\mathbf{t}$ forces (per unit of current volume and area, respectively) over the virtual velocity field $\tilde{\mathbf{v}}$,

$$\mathcal{W}_{ext} = \int_{\mathcal{P}} \mathbf{b} \cdot \tilde{\mathbf{v}} + \int_{\partial\mathcal{P}} \mathbf{t} \cdot \tilde{\mathbf{v}}\,. \tag{A.9}$$

The physics underlying (A.8) and (A.9) requires the power expenditures be invariant under a change of frame, i.e., *a change of observer*. This is turn requires them to be unchanged under the transformation (Gurtin et al., 2013)

$$\mathbf{F}^*(X,t) = \mathbf{Q}(t)\,\mathbf{F}(X,t) \qquad \mathbf{Q}(t) \in \mathbb{R}\mathrm{ot}. \tag{A.10}$$

where $\mathbb{R}\mathrm{ot}$ is the special orthogonal group, i.e., the set of orthogonal tensors with determinant $+1$. In addition the multiplicative decomposition of the deformation gradient (2.1) introduces in the theory an indeterminacy due to the fact that any rotation $\overline{\mathbf{Q}}(X,t)$ overimposed on the natural state is macroscopically invisible, in the sense that it keeps the deformation gradient $\mathbf{F}$ unchanged. In fact,

$$\begin{aligned}&\mathbf{F}(X,t) = \mathbf{F}_e(X,t)\overline{\mathbf{Q}}(X,t)\,\overline{\mathbf{Q}}^T(X,t)\mathbf{F}_v(X,t) = \mathbf{F}_e(X,t)\mathbf{F}_v(X,t),\\ &\quad\overline{\mathbf{Q}}(X,t) \in \mathbb{R}\mathrm{ot}\end{aligned} \tag{A.11}$$

This leads to the introduction of an additional requirement for the theory to be invariant under

$$\begin{aligned}&\mathbf{F}_e^\star(X,t) = \mathbf{F}_e(X,t)\overline{\mathbf{Q}}(X,t),\\ &\mathbf{F}_v^\star(X,t) = \overline{\mathbf{Q}}^T(X,t)\mathbf{F}_v(X,t)\,,\end{aligned} \tag{A.12}$$

usually called *Structural Frame Invariance* to distinguish it from *Frame Invariance* expressed by (A.10) (see Gurtin et al., 2013). This is indeed a constitutive prescription for the theory as pointed out in Ciambella and Nardinocchi (2023). In both (A.10) and (A.11), the explicit dependence on time t and position X is utilized to emphasize the distinction between frame invariance and structural-frame invariance. While frame invariance dictates invariance under a *global* rotation of the configuration (with the same $\mathbf{Q}$ applied to all material points), structural-frame invariance is a *local* requirement. This means it operates on the natural state, which is defined locally and does not necessarily represent an attainable configuration of the continuum.

By combining (A.10) with (A.12), the deformation rates in the transformed reference system are

$$\tilde{\mathbf{L}}^* = \dot{\mathbf{Q}}\mathbf{Q}^T + \mathbf{Q}\tilde{\mathbf{L}}\mathbf{Q}^T \tag{A.13}$$

$$\tilde{\mathbf{L}}_e^* = \dot{\mathbf{Q}}\mathbf{Q}^T + \mathbf{Q}\tilde{\mathbf{L}}_e\mathbf{Q}^T + \mathbf{Q}\mathbf{F}_e\dot{\overline{\mathbf{Q}}}\,\overline{\mathbf{Q}}^T\mathbf{F}_e^{-1}\mathbf{Q}^T \tag{A.14}$$

and

$$\tilde{\boldsymbol{\Lambda}}_v^* = \mathbf{Q}\tilde{\boldsymbol{\Lambda}}_v\mathbf{Q}^T + \mathbf{Q}\mathbf{F}_e\overline{\mathbf{Q}}\,\dot{\overline{\mathbf{Q}}}^T\mathbf{F}_e^{-1}\mathbf{Q}^T \tag{A.15}$$

where we made use of the definition (A.4) and of the fact that $\mathbf{Q}, \overline{\mathbf{Q}} \in \mathbb{R}\mathrm{ot}$.

In order to have the first term in (A.8) invariant under the transformations (A.13)–(A.15), it is required that

$$\int_{\mathcal{P}} \mathbf{T} \cdot \tilde{\mathbf{L}} = \int_{\mathcal{P}^*} \mathbf{T}^* \cdot \tilde{\mathbf{L}}^* = \int_{\mathcal{P}^*} \mathbf{T}^* \cdot \left(\dot{\mathbf{Q}}\mathbf{Q}^T + \mathbf{Q}\tilde{\mathbf{L}}\mathbf{Q}^T\right) \tag{A.16}$$

which is satisfied for each $\mathcal{P}$ if and only if

$$\mathbf{T}^* = \mathbf{Q}\mathbf{T}\mathbf{Q}^T \quad \text{and} \quad \mathbf{T} \in \mathbb{S}\mathrm{ym} \tag{A.17}$$

that expresses the well-known requirement for the stress tensor $\mathbf{T}$ to be symmetric and to be transformed under a rigid body motion overimposed on the current frame as (A.17). The other term in (A.8) gives

$$\begin{aligned}\int_{\mathcal{P}} (\mathbf{T}_v - \mathbf{T}_e) \cdot \tilde{\boldsymbol{\Lambda}}_v &= \int_{\mathcal{P}^*} (\mathbf{T}_v^* - \mathbf{T}_e^*) \cdot \tilde{\boldsymbol{\Lambda}}_v^* \\ &= \int_{\mathcal{P}^*} (\mathbf{T}_v^* - \mathbf{T}_e^*) \cdot (\mathbf{Q}\tilde{\boldsymbol{\Lambda}}_v\mathbf{Q}^T + \mathbf{Q}\mathbf{F}_e\overline{\mathbf{Q}}\dot{\overline{\mathbf{Q}}}^T\mathbf{F}_e^{-1}\mathbf{Q}^T)\,,\end{aligned} \tag{A.18}$$

which is satisfied if and only if

$$\mathbf{T}_v^* - \mathbf{T}_e^* = \mathbf{Q}(\mathbf{T}_v - \mathbf{T}_e)\mathbf{Q}^T \quad \text{and} \quad \mathbf{F}_e^T(\mathbf{T}_v - \mathbf{T}_e)\mathbf{F}_e^{-T} \in \mathbb{S}\mathrm{ym} \tag{A.19}$$

Finally, the request for the inner and external powers to be the same under each virtual velocity fields $\{\tilde{\mathbf{v}}, \tilde{\boldsymbol{\Lambda}}_v\}$ in each part $\mathcal{P} \subset \Omega_t$ allows us to derive the balance equations. On assuming $\tilde{\mathbf{L}} \neq \mathbf{0}$ and $\tilde{\boldsymbol{\Lambda}}_v = \mathbf{0}$, we get $\tilde{\mathbf{L}}_e = \tilde{\mathbf{L}}$ and

$$\int_{\mathcal{P}} \mathbf{T} \cdot \tilde{\mathbf{L}} = \int_{\mathcal{P}} \mathbf{b} \cdot \tilde{\mathbf{v}} + \int_{\partial\Pi} \mathbf{t} \cdot \tilde{\mathbf{v}}$$

from which we derive through integration by parts and localization due to the arbitrariness of $\tilde{\mathbf{v}}$, the standard force balance equation,

$$\begin{aligned}&\operatorname{div}\mathbf{T} + \mathbf{b} = \mathbf{0} && \text{in} \quad \Omega_t \\ &\mathbf{T}\mathbf{n} = \mathbf{t} && \text{on} \quad \partial\Omega_t\end{aligned} \tag{A.20}$$

that allows us to recognize $\mathbf{T}$ is indeed the Cauchy stress tensor.

Next we assume $\tilde{\mathbf{L}} = \mathbf{0}$ and $\tilde{\boldsymbol{\Lambda}}_v \neq \mathbf{0}$, such that $\tilde{\mathbf{L}}_e = -\tilde{\boldsymbol{\Lambda}}_v$ so to have

$$\int_{\mathcal{P}} (\mathbf{T}_v - \mathbf{T}_e) \cdot \tilde{\boldsymbol{\Lambda}}_v = \mathbf{0} \tag{A.21}$$

and from the arbitrariness of $\mathcal{P}$ we get

$$\mathbf{T}_v = \mathbf{T}_e \qquad \text{in} \quad \Omega_t \tag{A.22}$$

which is often called *micro-force* balance to highlight the fact that the difference between $\mathbf{T}_e$ and $\mathbf{T}_v$ acts on an inner degree of freedom $\boldsymbol{\Lambda}_v$.

A.3. Constitutive prescriptions based on thermodynamics

Under the constitutive assumptions outlined in Section 2, the internal rate of working over a part $\mathcal{P} \subset \Omega_t$ of the current space can be written as

$$\mathcal{F} = \frac{d}{dt}\int_{\mathcal{P}} \rho_t\,(\varphi_\infty(\mathbf{B}) + \varphi(\mathbf{B}_e)) = \int_{\mathcal{P}} \rho_t(\frac{\partial\varphi_\infty}{\partial\mathbf{B}} \cdot \dot{\mathbf{B}} + \frac{\partial\varphi}{\partial\mathbf{B}_e} \cdot \dot{\mathbf{B}}_e), \tag{A.23}$$

ρ_t being the density per unit of current volume, which is related to the reference density by $\rho_t = J^{-1}\rho_0$ due to the hypothesis of mass conservation.

Dissipation is defined as the difference between the external power, which equals the internal power, and the rate of internal energy. On using (A.8) and (A.9), we obtain

$$\begin{aligned}\mathcal{D} &= \mathcal{W}_{ext} - \frac{d}{dt}\int_{\mathcal{P}} \rho_t\,(\varphi_\infty(\mathbf{B}) + \varphi(\mathbf{B}_e)) \\ &= \int_{\mathcal{P}} (\mathbf{T}\cdot\mathbf{L} + (\mathbf{T}_v - \mathbf{T}_e)\cdot\boldsymbol{\Lambda}_v) - \int_{\mathcal{P}} \rho_t(\frac{\partial\varphi_\infty}{\partial\mathbf{B}} \cdot \dot{\mathbf{B}} + \frac{\partial\varphi}{\partial\mathbf{B}_e} \cdot \dot{\mathbf{B}}_e) \geq 0\end{aligned} \tag{A.24}$$

On noting that

$$\frac{\partial\varphi_\infty}{\partial\mathbf{B}} \cdot \dot{\mathbf{B}} = 2\,\frac{\partial\varphi_\infty}{\partial\mathbf{B}}\,\mathbf{B} \cdot \mathbf{D}$$

and

$$\frac{\partial\varphi}{\partial\mathbf{B}_e} \cdot \dot{\mathbf{B}}_e = 2\,\frac{\partial\varphi}{\partial\mathbf{B}_e}\,\mathbf{B}_e \cdot \operatorname{sym}\{\mathbf{L} - \boldsymbol{\Lambda}_v\}$$

Eq. (A.24) gives

$$\mathcal{D} = \int_{\mathcal{P}} \Big\{ (\mathbf{T} - 2\rho_t\frac{\partial\varphi}{\partial\mathbf{B}_e}\mathbf{B}_e - 2\rho_t\frac{\partial\varphi}{\partial\mathbf{B}}\mathbf{B}) \cdot \mathbf{D} + (\mathbf{T}_v - \mathbf{T}_e) \cdot \boldsymbol{\Lambda}_v + (2\rho_t\frac{\partial\varphi}{\partial\mathbf{B}_e}\mathbf{B}_e) \cdot \operatorname{sym}\boldsymbol{\Lambda}_v \Big\} \geq 0 \tag{A.25}$$

For what concerns the second term in (A.25), the invariance requirement $\mathbf{F}_e^T(\mathbf{T}_v - \mathbf{T}_e)\mathbf{F}_e^{-T} \in \mathbb{S}\mathrm{ym}$ (Eq. (A.19)) and the definition of $\boldsymbol{\Lambda}_v$ in (A.4) yields

$$(\mathbf{T}_v - \mathbf{T}_e) \cdot \boldsymbol{\Lambda}_v = \mathbf{F}_e^T(\mathbf{T}_v - \mathbf{T}_e)\mathbf{F}_e^{-T} \cdot \mathbf{D}_v = (\mathbf{T}_v - \mathbf{T}_e) \cdot \mathbf{F}_e\mathbf{D}_v\mathbf{F}_e^{-1}\,.$$

On the other hand, the first term in (A.25) leads to the usual constitutive assumption for the Cauchy stress $\mathbf{T} = \mathbf{T}_\infty + \mathbf{T}_e$, given in Eq. (2.3). For isotropic materials, $\frac{\partial\varphi_\infty}{\partial\mathbf{B}}$ and $\mathbf{B}$ ($\frac{\partial\varphi}{\partial\mathbf{B}_e}$ and $\mathbf{B}_e$) commute, and so $\mathbf{T}_\infty$ and $\mathbf{T}_e$ are symmetric tensors (and so $\mathbf{T}_v$ is). Therefore inequality (A.25) reduces to

$$\begin{aligned}\mathcal{D} &= \int_{\mathcal{P}} (\mathbf{T}_v - \mathbf{T}_e) \cdot \operatorname{sym}\{\mathbf{F}_e\mathbf{D}_v\mathbf{F}_e^{-1}\} + \int_{\mathcal{P}} \mathbf{T}_e \cdot \operatorname{sym}\{\mathbf{F}_e\mathbf{L}_v\mathbf{F}_e^{-1}\} \\ &= \int_{\mathcal{P}} \mathbf{T}_v \cdot \operatorname{sym}\{\mathbf{F}_e\mathbf{D}_v\mathbf{F}_e^{-1}\} + \int_{\mathcal{P}} \mathbf{T}_e \cdot \operatorname{sym}\{\mathbf{F}_e\mathbf{W}_v\mathbf{F}_e^{-1}\} \geq 0\end{aligned} \tag{A.26}$$

where $\mathbf{W}_v = \operatorname{skw}\{\mathbf{L}_v\}$. The latter term, and the fact that $\mathbf{T}_e$ is assumed symmetric by the constitutive Eq. (2.3), show that

$$\mathbf{T}_e \cdot \mathbf{F}_e\mathbf{W}_v\mathbf{F}_e^{-1} = \mathbf{T}_e\mathbf{B}_e^{-1} \cdot \mathbf{F}_e\mathbf{W}_v\mathbf{F}_e^T$$

If the material is isotropic $\mathbf{T}_e$ and $\mathbf{B}_e^{-1}$ have the same eigenvectors and so their product is symmetric. Since $\mathbf{F}_e\mathbf{W}_v\mathbf{F}_e^T \in \mathbb{S}\mathrm{kw}$, the inner product is zero. The arbitrarity of the part $\mathcal{P}$ allows us to rewrite the local form of the reduced dissipation inequality

$$\rho_0\,\delta = J\,\mathbf{T}_v \cdot \operatorname{sym}\{\mathbf{F}_e\mathbf{D}_v\mathbf{F}_e^{-1}\} \geq 0 \tag{A.27}$$

where δ is the specific dissipation density per unit of reference mass. Previous inequality is usually expressed in terms of the codeformational derivative of $\mathbf{B}_e$ defined by

$$\mathbf{B}_e^{\triangledown} = \dot{\mathbf{B}}_e - \mathbf{L}\mathbf{B}_e - \mathbf{B}_e\mathbf{L}^T\,, \tag{A.28}$$

that, by using (A.6) gives

$$\operatorname{sym}\{\mathbf{F}_e\mathbf{D}_v\mathbf{F}_e^{-1}\} = -\frac{1}{2}\operatorname{sym}\{\mathbf{B}_e^{\triangledown}\mathbf{B}_e^{-1}\}\,. \tag{A.29}$$

(A.27) can be used to introduce a constitutive equation of the viscous stress $\mathbf{T}_v$ that guarantees that the dissipation remains positive for every strain rate. Most of the models commonly employed in the literature assume the dissipation to be a quadratic function of the rate of deformation, i.e.

$$J\mathbf{T}_v = -\frac{1}{2}\mathbb{D}[\operatorname{sym}\{\mathbf{B}_e^{\triangledown}\mathbf{B}_e^{-1}\}] \tag{A.30}$$

where $\mathbb{D}$ the fourth-order positive definite viscosity tensor, that has the following representation formula for a isotropic material:

$$\mathbb{D} = 2\,\eta_D\big(\mathbb{I} - \frac{1}{3}\mathbf{I}\otimes\mathbf{I}\big) + 3\,\eta_V(\frac{1}{3}\mathbf{I}\otimes\mathbf{I}) \tag{A.31}$$

where $\mathbb{I}$ is the fourth-order identity tensor, i.e., $\mathbb{I}[\mathbf{A}] = \mathbf{A}$, and $(\mathbf{I}\otimes\mathbf{I})\mathbf{A} = (\mathbf{I}\cdot\mathbf{A})\mathbf{I}$. Here, η_D and η_V are the deviatoric and volumetric viscosities (per unit of reference volume), that can eventually depends on the elastic or total deformation. It is important to note that the reduced dissipation inequality (A.27) is automatically satisfied if the two viscosity functions are positive, namely

$$\eta_D \geq 0, \qquad \eta_V \geq 0\,. \tag{A.32}$$

By using the representation formula of the viscosity tensor, we obtain the following constitutive equation of the viscous stress

$$\begin{aligned}J\,\mathbf{T}_v^D &= -\,\eta_D\operatorname{dev}\operatorname{sym}\{\mathbf{B}_e^{\triangledown}\mathbf{B}_e^{-1}\} \\ J\,\mathbf{T}_v^V &= -\frac{3}{2}\,\eta_V\operatorname{vol}\operatorname{sym}\{\mathbf{B}_e^{\triangledown}\mathbf{B}_e^{-1}\}\end{aligned} \tag{A.33}$$

in terms of the deviatoric and volumetric rates, i.e., $\operatorname{dev}\mathbf{A} = \mathbf{A} - \frac{1}{3}(\mathbf{A}\cdot\mathbf{I})\mathbf{I}$ and $\operatorname{vol}\mathbf{A} = \frac{1}{3}(\mathbf{A}\cdot\mathbf{I})\mathbf{I}$, thus $\mathbf{T}_v = \mathbf{T}_v^V + \mathbf{T}_v^D$.

By combining the constitutive Eq. (A.33) with the micro-balance Eq. (A.22), we obtain the equation governing the evolution of the inelastic variable (Biot's equation):

$$-\frac{1}{2}\operatorname{sym}\{\mathbf{B}_e^{\triangledown}\mathbf{B}_e^{-1}\} = \frac{J}{2\eta_D}\mathbf{T}_e^D + \frac{J}{3\eta_V}\mathbf{T}_e^V\,. \tag{A.34}$$

Since elastomers and many other soft materials can be considered as incompressible, the only admissible deformations are isochoric, that is, $J = \det\mathbf{F} = 1$, and volumetric viscoelastic effects are usually neglected, i.e., $\det \mathbf{F}_e = 1$. In particular the latter yields $\operatorname{vol}\operatorname{sym}\{\mathbf{F}_e\mathbf{D}_v\mathbf{F}_e^{-1}\} = 0$ ($\eta_V \to \infty$) and only one of the evolution Eqs. (A.34) survives:

$$-\operatorname{sym}\{\mathbf{B}_e^{\triangledown}\mathbf{B}_e^{-1}\} = \frac{1}{\eta_D}\mathbf{T}_e^D \tag{A.35}$$

that is Eq. (2.7) (and Eq. 25 in Reese and Govindjee, 1998).

Appendix B. Analytical derivation for simple shear under finite strain

To compute the material's stress response in the re-training phase, we utilized the methodology outlined in Califano and Ciambella (2023). Looking at Fig. 1, the total Cauchy stress $\mathbf{T}$ is given by $\mathbf{T} = \mathbf{T}_\infty + \mathbf{T}_e$ as in Eq. (2.3). In the case of simple shear with deformation gradient $\mathbf{F} = \mathbf{I} + \gamma\, \mathbf{e}_1 \otimes \mathbf{e}_2$, the components of the equilibrium stress $\mathbf{T}_\infty$ under plane stress and incompressibility read:

$$T_{\infty_{11}} = 2\,\varphi_{\infty_1}\,\gamma^2, \quad T_{\infty_{22}} = -2\,\varphi_{\infty_2}\,\gamma^2, \quad T_{\infty_{12}} = 2\,(\varphi_{\infty_1} + \varphi_{\infty_2})\,\gamma,$$

where $\varphi_{\infty_1} = \rho_0\partial\varphi_\infty/\partial I_1$ and $\varphi_{\infty_2} = \rho_0\partial\varphi_\infty/\partial I_2$ are the partial derivatives of the strain energy density $\varphi_\infty(I_1, I_2)$. In the case of a neo-Hookean material, as defined in Eq. (2.10), one has $\varphi_{\infty_1} = \mu_\infty/2$ and $\varphi_{\infty_2} = 0$. Therefore, in the context of hyperelastic materials,[2] the simple shear scenario can be fully described by one parameter, namely the shear magnitude. Conversely, for viscoelastic materials, the elastic deformation emerges from the combination of a triaxial stretch and a simple shear. The evolution equation of the inelastic variable (2.7) is solved by assuming the following constitutive equation of the elastic stress

$$\mathbf{T}_e^D = 2\,\varphi_1\,\mathbf{B}_e^D - 2\,\varphi_2\,\left(\mathbf{B}_e^{-1}\right)^D, \tag{B.1}$$

where $\varphi_1 = \rho_0\,\partial\varphi/\partial I_1^e$ and $\varphi_2 = \rho_0\,\partial\varphi/\partial I_2^e$, s.t. $I_1^e = \mathbf{I}\cdot\mathbf{B}_e$ and $I_2^e = \mathbf{I}\cdot\mathbf{B}_e^{-1}$. In the case of a neo-Hookean material, as described in Eq. (2.10), one has $\varphi_1 = \mu/2$ and $\varphi_2 = 0$. For simple shear deformation in the $\mathbf{e}_1$-$\mathbf{e}_2$ plane, the deformation rate tensor takes the form:

$$\mathbf{L} = \dot{\gamma}(t)\,\mathbf{e}_1 \otimes \mathbf{e}_2, \tag{B.2}$$

where $\dot{\gamma}(t)$ denotes the shear rate, and $\mathbf{L} = \dot{\mathbf{F}}\mathbf{F}^{-1}$. The explicit form of $\mathbf{B}_e$ is given by:

$$\mathbf{B}_e = b_1\mathbf{e}_1 \otimes \mathbf{e}_1 + b_2\mathbf{e}_2 \otimes \mathbf{e}_2 + b_3\mathbf{e}_3 \otimes \mathbf{e}_3 + b_{12}\left(\mathbf{e}_1 \otimes \mathbf{e}_2 + \mathbf{e}_2 \otimes \mathbf{e}_1\right), \tag{B.3}$$

To ensure the isochoric nature of this deformation, we mandate that:

$$b_3 = \left(b_1 b_2 - b_{12}^2\right)^{-1} \tag{B.4}$$

The conditions $b_1 > 0, b_2 > 0$, and $b_{12}^2 < b_1 b_2$ are necessary conditions for the positive definiteness of $\mathbf{B}_e$. The determination of the three unknown functions $b_1(t)$, $b_2(t)$, and $b_{12}(t)$ requires the numerical integration of the three ordinary differential equations derived by adapting (2.7) and (B.1) to the assumed form of $\mathbf{B}_e$, i.e.,

$$\left.\begin{aligned} \dot{b}_1 &= \tfrac{4A_1}{3\eta_D} + 2\dot{\gamma}b_{12} \\ \dot{b}_2 &= \tfrac{4A_2}{3\eta_D} \\ \text{and}\quad \dot{b}_{12} &= \tfrac{4A_{12}}{3\eta_D}b_{12} + \dot{\gamma}b_2 \end{aligned}\right\} \tag{B.5}$$

where η_D represents the material's viscosity coefficient defined in (2.14) and A_1, A_2, and A_{12} are material parameters derived in Califano and Ciambella (2023), that for the specific of the energy densities (2.10) are given by

$$\begin{aligned} A_1 &= \frac{\mu}{2}\,(b_1(-2\,b_1 + b_2 + b_3) - 3\,b_{12}^2) \\ A_2 &= \frac{\mu}{2}\,(b_2(b_1 - 2\,b_2 + b_3) - 3\,b_{12}^2) \\ A_{12} &= \frac{\mu}{2}\,(b_3 - 2\,b_1 - 2\,b_2) \end{aligned} \tag{B.6}$$

Under the assumption of plane stress and material incompressibility, in the specific case of a neo-Hookean material, the non-vanishing components of the elastic Cauchy stress tensor $\mathbf{T}_e$ can be expressed as

$$\begin{aligned} T_{11} &= 2\,\mu\,(b_1 - b_3) \\ T_{22} &= 2\,\mu\,(b_2 - b_3) \\ T_{12} &= 2\,\mu\,b_{12} \end{aligned} \tag{B.7}$$

Data availability

The datasets generated and/or analyzed during the current study are available from the corresponding author on reasonable request.

[2] That is, when considered on its own, the long-term branch $\mathbf{T}_\infty$ of the generalized Maxwell model behaves as a purely hyperelastic material, represented by an ideal spring governed by a strain energy function.

References

Amos, B., Xu, L., Kolter, J.Z., 2017. Input convex neural networks. In: D. Precup, Y. W. Teh (Ed.), Proceedings of the 34th International Conference on Machine Learning. In: Proceedings of Machine Learning Research, vol. 70, PMLR, pp. 146–155.
As'ad, F., Avery, P., Farhat, C., 2022. A mechanics-informed artificial neural network approach in data-driven constitutive modeling. Internat. J. Numer. Methods Engrg. 123 (12), 2738–2759. http://dx.doi.org/10.1002/nme.6957.
As'ad, F., Farhat, C., 2023. A mechanics-informed deep learning framework for data-driven nonlinear viscoelasticity. Comput. Methods Appl. Mech. Engrg. 417, 116463. http://dx.doi.org/10.1016/j.cma.2023.116463.
Bergström, J., Boyce, M., 1998. Constitutive modeling of the large strain time-dependent behavior of elastomers. J. Mech. Phys. Solids 46 (5), 931–954. http://dx.doi.org/10.1016/S0022-5096(97)00075-6.
Bergstrom, J.S., Boyce, M.C., 2000. Large strain time-dependent behavior of filled elastomers. Mech. Mater. 32 (11), 627–644. http://dx.doi.org/10.1016/S0167-6636(00)00028-4.
Bradbury, J., Frostig, R., Hawkins, P., Johnson, M.J., Leary, C., Maclaurin, D., Necula, G., Paszke, A., VanderPlas, J., Wanderman-Milne, S., Zhang, Q., 2018. JAX: composable transformations of python+numpy programs.
Brancati, R., Di Massa, G., Pagano, S., Santini, S., 2019. A magneto-rheological elastomer vibration isolator for lightweight structures. Meccanica 54 (1–2), 333–349. http://dx.doi.org/10.1007/s11012-019-00951-2.
Califano, F., Ciambella, J., 2023. Viscoplastic simple shear at finite strains. Proc. R. Soc. A 479 (2263), 20230603. http://dx.doi.org/10.1098/rspa.2023.0603.
Chazeau, L., Brown, J.D., Yanyo, L.C., Sternstein, S.S., 2000. Modulus recovery kinetics and other insights into the Payne effect for filled elastomers. Polym. Compos. 21 (2), 202–222. http://dx.doi.org/10.1002/pc.10178.
Chen, G., 2021. Recurrent neural networks (rnns) learn the constitutive law of viscoelasticity. Comput. Mech. 67 (3), 1009–1019. http://dx.doi.org/10.1007/s00466-021-01981-y.
Chen, P., Guilleminot, J., 2022. Polyconvex neural networks for hyperelastic constitutive models: A rectification approach. Mech. Res. Commun. 125, 103993. http://dx.doi.org/10.1016/j.mechrescom.2022.103993.
Ciambella, J., Nardinocchi, P., 2021. A structurally frame-indifferent model for anisotropic visco-hyperelastic materials. J. Mech. Phys. Solids 147, 104247. http://dx.doi.org/10.1016/j.jmps.2020.104247.
Ciambella, J., Nardinocchi, P., 2023. Non-affine Fiber Reorientation in Finite Inelasticity. J. Elasticity 153 (4–5), 735–753. http://dx.doi.org/10.1007/s10659-022-09945-w.
Ciambella, J., Paolone, A., Vidoli, S., 2011. Memory decay rates of viscoelastic solids: not too slow, but not too fast either. Rheol. Acta 80 (7–8), 661–674. http://dx.doi.org/10.1007/s00397-011-0549-y.
Cleément, F., Bokobza, L., Monnerie, L., 2005. Investigation of the Payne Effect and its Temperature Dependence on Silica-Filled Polydimethylsiloxane Networks, journal=Part I: Exp. Results, rubber chem. Technol.. 78 (2), 211–231. http://dx.doi.org/10.5254/1.3547879.
Conti, S., Müller, S., Ortiz, M., 2018. Data-driven problems in elasticity. Arch. Ration. Mech. Anal. 229 (1), 79–123. http://dx.doi.org/10.1007/s00205-017-1214-0.
DeepMind I. Babuschkin, Baumli, K., Bell, A., Bhupatiraju, S., Bruce, J., Buchlovsky, P., Budden, D., Cai, T., Clark, A., Danihelka, I., Dedieu, A., Fantacci, C., Godwin, J., Jones, C., Hemsley, R., Hennigan, T., Hessel, M., Hou, S., Kapturowski, S., Keck, T., Kemaev, I., King, M., Kunesch, M., Martens, L., Merzic, H., Mikulik, V., Norman, T., Papamakarios, G., Quan, J., Ring, R., Ruiz, F., Sanchez, A., Sartran, L., Schneider, R., Sezener, E., Spencer, S., Srinivasan, S., Stanojević, M., Stokowiec, W., Wang, L., Zhou, G., Viola, F., 2020. The DeepMind JAX ecosystem.

DiCarlo, A., Quiligotti, S., 2002. Growth and balance. Mech. Res. Commun. 29 (6), 449–456. http://dx.doi.org/10.1016/s0093-6413(02)00297-5.

Drozdov, A.D., 2004. Finite viscoplasticity of non-affine networks: Stress overshoot under shear. Contin. Mech. Thermodyn. 16 (1–2), 73–95. http://dx.doi.org/10.1007/s00161-003-0142-4.

Eggersmann, R., Kirchdoerfer, T., Reese, S., Stainier, L., Ortiz, M., 2019. Model-free data-driven inelasticity. Comput. Methods Appl. Mech. Engrg. 350, 81–99. http://dx.doi.org/10.1016/j.cma.2019.02.016.

Faroughi, S.A., Pawar, N.M., Fernandes, C., Raissi, M., Das, S., Kalantari, N.K., Kourosh Mahjour, S., Physics-Guided, . and Physics-Informed, ., title=and Physics-Encoded Neural Networks and Operators in Scientific Computing: Fluid and Solid Mechanics, 2024. J. Comput. Inf. Sci. Eng. 24 (4), 040802. http://dx.doi.org/10.1115/1.4064449.

Furukawa, T., Yagawa, G., 1998. Implicit constitutive modelling for viscoplasticity using neural networks. Internat. J. Numer. Methods Engrg. 43 (2), 195–219. http://dx.doi.org/10.1002/(SICI)1097-0207(19980930)43:2<195::AID-NME418>3.0.CO;2-6.

Gent, A.N., 2012. Engineering with Rubber: How To Design Rubber Components Edition: 3rd Revised Edition. vol. 17.

Ghaboussi, J., Sidarta, D., 1998. New nested adaptive neural networks (nann) for constitutive modeling. Comput. Geotech. 22 (1), 29–52. http://dx.doi.org/10.1016/S0266-352X(97)00034-7.

Ghaderi, A., Morovati, V., Dargazany, R., 2020. A physics-informed assembly of feed-forward neural network engines to predict inelasticity in cross-linked polymers. Polymers 12 (11), 2628. http://dx.doi.org/10.3390/polym12112628.

Ghosh, K., Lopez-Pamies, O., 2021. On the two-potential constitutive modeling of dielectric elastomers. Meccanica 56 (6), 1505–1521. http://dx.doi.org/10.1007/s11012-020-01179-1.

Goodfellow, I., Bengio, Y., Courville, A., 2016. Deep Learning. MIT Press.

Gouhier, F., Diani, J., 2024. A comparison of finite strain viscoelastic models based on the multiplicative decomposition. Eur. J. Mech. A Solids 108, 105424. http://dx.doi.org/10.1016/j.euromechsol.2024.105424.

Gurtin, M.E., Fried, E., Anand, L., 2013. The Mechanics and Thermodynamics of Continua, vol. 53, http://dx.doi.org/10.1017/cbo9781107415324.004.

Holthusen, H., Lamm, L., Brepols, T., Reese, S., Kuhl, E., 2024. Theory and implementation of inelastic constitutive artificial neural networks. Comput. Methods Appl. Mech. Engrg. 428, 117063. http://dx.doi.org/10.1016/j.cma.2024.117063.

Hurtado, J., Lapczyk, I., Govindarajan, S., 2013. Parallel rheological framework to model nonlinear viscoelasticity, permanent set, and mullins effect in elastomers. In: N. Gil-Negrete, A. Alonso (Ed.), Constitutive Models for Rubber VIII. CRC Press, London, pp. 6–10.

Jones, R., Frankel, A., Johnson, K., 2022. A neural ordinary differential equation framework for modeling inelastic stress response via internal state variables. J. Mach. Learn. Model. Comput. 3 (3), 1–35. http://dx.doi.org/10.1615/JMachLearnModelComput.2022042917.

Jung, S., Ghaboussi, J., 2006. Neural network constitutive model for rate-dependent materials. Comput. Struct. 84 (15), 955–963. http://dx.doi.org/10.1016/j.compstruc.2006.02.015.

Kestin, J., Rice, J.R., 1970. Paradoxes in the application of thermodynamics to strained solids. In: E. B. Stuart, A. J. Brainard (Ed.), A Critical Review of Thermodynamics, Mono Book Corp. Baltimore, pp. 275–298.

Kidger, P., 2021. On Neural Differential Equations. (Ph.D. thesis). University of Oxford, http://dx.doi.org/10.48550/arXiv.2202.02435.

Kim, S., Berger, T., Kaliske, M., Piao, M., 2022. A strain rate dependent constitutive model for uncured rubber. In: J. Diani, N. Alves (Ed.), Constitutive Models for Rubber XII, first ed. CRC Press, p. 7.

Kingma, D.P., Ba, J., 2014. Adam: A method for stochastic optimization. arXiv preprint arXiv:1412.6980.

Klein, D.K., Fernández, M., Martin, R.J., Neff, P., Weeger, O., 2022. Polyconvex anisotropic hyperelasticity with neural networks. J. Mech. Phys. Solids 159 (2021), http://dx.doi.org/10.1016/j.jmps.2021.104703.

Klüppel, M., 2008. Evaluation of viscoelastic master curves of filled elastomers and applications to fracture mechanics. J. Phys.: Condens. Matter. 21 (3), 035104. http://dx.doi.org/10.1088/0953-8984/21/3/035104.

Kumar, A., Lopez-Pamies, O., 2016. On the two-potential constitutive modeling of rubber viscoelastic materials. C. R. Mec. 344 (2), 102–112. http://dx.doi.org/10.1016/j.crme.2015.11.004.

LeCun, Y., Bengio, Y., Hinton, G., learning, Deep., 2015. Nature 521 (7553), 436–444. http://dx.doi.org/10.1038/nature14539, published on 2015/05/01.

LeCun, Y., Bottou, L., Orr, G.B., Müller, K.R., 1998. Efficient BackProp. Springer Berlin Heidelberg, Berlin, Heidelberg, pp. 9–50. http://dx.doi.org/10.1007/3-540-49430-8_2.

Linden, L., Klein, D.K., Kalina, K.A., Brummund, J., Weeger, O., Kästner, M., 2023. Neural networks meet hyperelasticity: A guide to enforcing physics. J. Mech. Phys. Solids 179, 105363. http://dx.doi.org/10.1016/j.jmps.2023.105363.

Linka, K., Hillgärtner, M., Abdolazizi, K.P., Aydin, R.C., Itskov, M., Cyron, C.J., 2021. Constitutive artificial neural networks: A fast and general approach to predictive data-driven constitutive modeling by deep learning. J. Comput. Phys. 429, 110010. http://dx.doi.org/10.1016/j.jcp.2020.110010.

Linka, K., Kuhl, E., 2021. Constitutive artificial neural networks: A general anisotropic constitutive modeling framework utilizing machine learning. J. Appl. Mech. 88 (5), 051003. http://dx.doi.org/10.1115/1.4049845.

Linka, K., Kuhl, E., 2023. A new family of constitutive artificial neural networks towards automated model discovery. Comput. Methods Appl. Mech. Engrg. 403, 115731. http://dx.doi.org/10.1016/j.cma.2022.115731.

McLeish, T.C.B., Larson, R.G., 1998. Molecular constitutive equations for a class of branched polymers: The pom-pom polymer. J. Rheol. 42 (1), 81–110. http://dx.doi.org/10.1122/1.550933.

Pan, S.J., Yang, Q., 2010. A survey on transfer learning. IEEE Trans. Knowl. Data Eng. 22 (10), 1345–1359.

Prechelt, L., 2012. Early Stopping — But when?. Springer Berlin Heidelberg, Berlin, Heidelberg, pp. 53–67. http://dx.doi.org/10.1007/3-540-49430-8_3.

Reese, S., Govindjee, S., 1998. A theory of finite viscoelasticity and numerical aspects. Int. J. Solids Struct. 35 (26–27), 3455–3480. http://dx.doi.org/10.1016/S0020-7683(97)00217-5.

Riande, E., Diaz-Calleja, R., Prolongo, M.G., Msegosa, R.M., Salom, C., 2000. Polymer Viscoelasticity: Stress and Strain in Practice. Marcel Dekker, Inc..

Rubin, M., Papes, O., 2011. Advantages of formulating evolution equations for elastic-viscoplastic materials in terms of the velocity gradient instead of the spin tensor. J. Mech. Mater. Struct. 6 (1–4), 529–543. http://dx.doi.org/10.2140/jomms.2011.6.529.

Sadik, S., Yavari, A., 2017. On the origins of the idea of the multiplicative decomposition of the deformation gradient. Math. Mech. Solids 22 (4), 771–772. http://dx.doi.org/10.1177/1081286515612280.

Salahshoor, H., Ortiz, M., 2023. Model-free data-driven viscoelasticity in the frequency domain. Comput. Methods Appl. Mech. Engrg. 403, 115657. http://dx.doi.org/10.1016/j.cma.2022.115657.

Schraudolph, N.N., 1998. Centering Neural Network Gradient Factors. Springer Berlin Heidelberg, Berlin, Heidelberg, pp. 207–226. http://dx.doi.org/10.1007/3-540-49430-8_11.

Sherstinsky, A., 2020. Fundamentals of recurrent neural network (rnn) and long short-term memory (lstm) network. Phys. D: Nonlinear Phenom. 404, 132306. http://dx.doi.org/10.1016/j.physd.2019.132306.

Si, X., Luo, M., Li, M., Ma, Y., Huang, Y., Ge, Y., 2023. Magnetic properties of a soft magnetic elastomer based on antioxidant magnetic composite particles and a water-soluble polymer matrix. Soft Matter 19, 1008–1016. http://dx.doi.org/10.1039/d2sm01426b.

Taç, V., Linka, K., Sahli-Costabal, F., Kuhl, E., Tepole, A.B., 2024a. Benchmarking physics-informed frameworks for data-driven hyperelasticity. Comput. Mech. 73 (1), 49–65. http://dx.doi.org/10.1007/s00466-023-02355-2.

Taç, V., Rausch, M.K., Bilionis, I., Sahli Costabal, F., Tepole, A.B., 2024b. Generative hyperelasticity with physics-informed probabilistic diffusion fields. Eng. Comput. http://dx.doi.org/10.1007/s00366-024-01984-2.

Taç, V., Rausch, M., Sahli Costabal, F., Buganza Tepole, A., 2023. Data-driven anisotropic finite viscoelasticity using neural ordinary differential equations. Comput. Methods Appl. Mech. Engrg. 411 (3), 116046. http://dx.doi.org/10.1016/j.cma.2023.116046.

Taç, V., Sahli Costabal, F., Tepole, A.B., 2022a. Data-driven tissue mechanics with polyconvex neural ordinary differential equations. Comput. Methods Appl. Mech. Engrg. 398, 115248. http://dx.doi.org/10.1016/j.cma.2022.115248.

Taç, V., Sree, V.D., Rausch, M.K., Tepole, A.B., 2022b. Data-driven Modeling of the Mechanical Behavior of Anisotropic Soft Biological Tissue. Eng. Comput. 38, 4167–4182. http://dx.doi.org/10.1007/s00366-022-01733-3.

Wang, L.M., Linka, K., Kuhl, E., 2023. Automated model discovery for muscle using constitutive recurrent neural networks. J. Mech. Behav. Biomed. Mater. 145, 106021. http://dx.doi.org/10.1016/j.jmbbm.2023.106021.

Yoshida, J., Sugiyama, T., 2015. A hyperelastic visco-elasto-plastic damage model for rubber-like solids including strain-dependency of hysteretic loops. J. Jpn. Soc. Civ. Eng. Ser. A2 (Appl. Mech.) 71, 14–33. http://dx.doi.org/10.2208/jscejam.71.14.

Zhang, P., Yin, Z.Y., Jin, Y.F., Sheil, B., 2022. Physics-constrained hierarchical data-driven modelling framework for complex path-dependent behaviour of soils. Int. J. Numer. Anal. Methods Geomech. 46 (10), 1831–1850. http://dx.doi.org/10.1002/nag.3370.