Engineering Applications of Artificial Intelligence 186 (2026) 116481

Contents lists available at ScienceDirect

Engineering Applications of Artificial Intelligence

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

Local evolution decoupling paradigm for neural network constitutive models of history-dependent materials: Theory, architecture, and applications to shape memory polymers

Chen Wang [a], Fengfeng Li [a,*], Dou Zhang [a,**], Hanxing Zhao [b,***], Xianyi Li [a], Yanju Liu [a,c], Jinsong Leng [b]

[a] Department of Astronautical Science and Mechanics, Harbin Institute of Technology (HIT), P.O. Box 301, No. 92 West Dazhi Street, Harbin, 150001, People's Republic of China
[b] Center for Composite Materials and Structures, Harbin Institute of Technology (HIT), P.O. Box 3011, No. 2 Yikuang Street, Harbin, 150080, People's Republic of China
[c] Suzhou Research Institute, Harbin Institute of Technology, Suzhou, 215100, People's Republic of China

ARTICLE INFO

Keywords:
History-dependent materials
Neural network
Constitutive modeling
Shape memory polymers

ABSTRACT

The mechanical response of history-dependent materials depends on loading paths and thermomechanical histories, rendering neural network-based constitutive modeling challenging. Existing methods typically establish global memory based on historical sequences, requiring large model scales and extensive training data. Targeting materials whose states can be fully described by finite-dimensional internal variables, this paper proposes a Local Decoupling Evolution Operator Network (LDEONet) grounded in the Markovian property. The proposed framework explicitly decomposes the internal variable evolution model into evolution direction and evolution intensity. Using the thermomechanical cycle of shape memory polymers as a case study, eight models are trained, and a series of experiments are designed, including path generalization tests, discretization generalization tests, and decoupling ablation studies. Experimental results demonstrate that the model contains 242 trainable parameters and requires only about 1/175 of the training curves compared to sequence-based models, while making predictions without needing historical memory. The model maintains accurate predictions across various unseen loading paths and unseen step sizes, and the decoupling architecture is shown to provide structural guarantees for generalization. The Local Decoupling Evolution Operator Network thus offers a lightweight modeling approach with both path generalization and discretization generalization capabilities for history-dependent materials that possess explicit internal variable models.

1. Introduction

Shape memory polymers (SMPs), as a representative class of history-dependent materials, are capable of recovering from a deformed temporary shape to a pre-programmed permanent shape under external stimuli (Chandaka et al., 2025; Teng et al., 2025). Their high recoverable strain, low density, tunable glass transition temperature, favorable biocompatibility, and excellent processability render them promising candidates for cutting-edge applications in soft robotics, biomedical devices, space-deployable structures, and four-dimensional (4D) printing (Chandaka et al., 2025; Fetisova et al., 2026). With ongoing technological advances, these application domains impose increasingly stringent requirements on constitutive models capable of describing material behavior under complex loading paths and thermomechanical histories (Dutta et al., 2025; Zhao et al., 2025).

The intrinsic shape memory mechanism of SMPs originates from the temperature-driven reversible transition between rubbery and glassy states. At elevated temperatures, the material exhibits entropic elasticity, characterized by high mobility of flexible chain segments. As the temperature decreases below the glass transition temperature, molecular segmental motion becomes frozen, and the material enters a glassy state dominated by energetic elasticity (Cooper et al., 2021). This

* Corresponding author.
** Corresponding author.
*** Corresponding author.
E-mail addresses: lifengfeng@hit.edu.cn (F. Li), dou_zhang@hit.edu.cn (D. Zhang), zhaohxhit@126.com (H. Zhao).

https://doi.org/10.1016/j.engappai.2026.116481
Received 2 May 2026; Received in revised form 22 August 2026; Accepted 6 October 2026
Available online 30 October 2026
0952-1976/© 2026 Elsevier Ltd. All rights are reserved, including those for text and data mining, AI training, and similar technologies.

transition process manifests as the evolution of internal state variables: the stored strain records and freezes the macroscopic deformation experienced during cooling and is gradually released upon reheating above the transition temperature, thereby enabling shape recovery (Chen and Lagoudas, 2008; Xu et al., 2023).

Extensive research efforts have been devoted to developing constitutive models that describe such thermomechanical behavior. Tobushi et al. first proposed a linear viscoelastic theory to characterize the shape memory effect of SMPs and investigated the thermomechanical response of polyurethane SMPs (Tobushi et al., 2001). Subsequently, linear viscoelastic theory has been progressively refined and extended. Established constitutive models for SMPs can be broadly categorized into viscoelastic models (Tobushi et al., 2001), phase transition models (Zhao et al., 2020), and hybrid formulations (Park et al., 2016). Chen and Lagoudas (2008) introduced a transformation gradient to establish the relationship between the deformation gradients of the rubbery and glassy phases. Volk et al. (Volk et al., 2010, 2011) experimentally validated and further advanced the finite deformation constitutive theory for SMPs based on comprehensive shape memory characterization of polyurethane. Although these physics-based models offer theoretical rigor and predictive capability, they often rely on cumbersome derivations and involve a large number of material parameters requiring extensive calibration (Park et al., 2016). With the rapid development of SMPs and their associated composites, an increasing variety of physical stimuli and material hybridization strategies are being employed in practical engineering applications (F. Li, 2020), further amplifying the complexity of constitutive modeling.

In recent years, SMP research has gradually expanded from conventional thermally responsive polymers to composites, multi-stimuli-responsive materials, and 4D-printed structures. Pan et al. established thermomechanical models for particle-reinforced SMP composites under cold and hot programming conditions, and discussed the effects of particle reinforcement, relaxation processes, and temperature-dependent evolution on shape recovery behavior (Pan et al., 2026). In addition to material composition, programming and service histories also significantly alter the shape memory performance and mechanical behavior of polymers (Huang et al., 2026). For example, high-temperature prestretching can tailor the brittle–ductile transition of PMMA and polycarbonate, while long-term storage affects the shape retention, recovery performance, and mechanical response of pre-stretched PMMA (Wang et al., 2019, 2022). Concurrently, the development of materials such as solvent-responsive porous semicrystalline thermoplastic polyurethanes has further expanded the stimulus modes and application environments of SMPs (Liu et al., 2022; Wu et al., 2022). These studies indicate that material composition, programming conditions, and service history collectively determine the macroscopic response of SMPs, thereby heightening the necessity of establishing reliable thermomechanical models.

4D printing has further propelled the application of SMPs in magnetically actuated structures, soft robotics, wearable devices, and biomedical instruments. Mirasadi et al. systematically reviewed 4D printing methods, material systems, and applications of magnetically responsive SMPs in soft robotics and wearables (Mirasadi et al., 2026); Yousefi et al. discussed the development of 4D-printed smart materials in minimally invasive deployable devices, drug delivery, and tissue regeneration (Yousefi et al., 2026). These applications involve different stimulation modalities, boundary conditions, and operational stages, indicating that a single evolution mode may be insufficient to describe the complete material working process.

Consequently, the thermomechanical behavior of SMPs, which inherently depends on the entire loading history, has emerged as both a focal point and a significant challenge in constitutive modeling (Fuhg et al., 2025). Unlike conventional materials, the stress response of history-dependent materials depends not only on the instantaneous strain but also on the entire preceding loading history (Abueidda et al., 2021). Constitutive descriptions of such materials often necessitate the introduction of numerous internal variables to capture their complex evolution laws, which not only complicates the solution procedure but also leads to low computational efficiency, particularly when confronting nonlinearities such as finite deformations and multi-field coupling (Collins-Craft et al., 2020; Feyel, 2003; Masi et al., 2018, 2020).

To overcome the limitations inherent in traditional modeling paradigms, neural network-based constitutive models have emerged as a promising alternative for history-dependent materials (Graves and Schmidhuber, 2005; Hochreiter and Schmidhuber, 1997; Schuster and Paliwal, 1997). Among these, recurrent neural networks (RNNs) and their variants, by virtue of their innate sequence-processing capabilities, can automatically encode historical information through hidden states, thereby accurately capturing material memory effects without explicitly defining internal variables (Graves et al., 2013a, 2013b). This characteristic has enabled RNN-based models to demonstrate potential exceeding that of exact solutions to traditional differential equations in modeling time-dependent phenomena such as the long-term creep of synthetic concrete (Freitag et al., 2011; Graf et al., 2010, 2012; Oeser and Freitag, 2009). Currently, modeling frameworks based on RNNs (Dornheim et al., 2024), long short-term memory networks (LSTM) (Zhang et al., 2020), and gated recurrent units (GRUs) (Gorji et al., 2020) have become mainstream approaches for developing neural network constitutive models for history-dependent materials. Despite their notable successes, these approaches exhibit certain limitations. Recurrent neural networks and their variants are capable of extracting latent state information of materials from historical sequences through hidden states, making them particularly suitable for cases where physical internal variables are unknown, difficult to measure directly, or inconvenient to prescribe in advance (Anumasa et al., 2023). However, the inputs and hidden states of such models are typically tied to the historical sequences and discretization schemes employed during training. Without specialized retraining or state reconstruction, initiating predictions from a new intermediate state or directly altering the discretization of the input sequence may still affect their predictive performance (Sun et al., 2025).

To address the limited generalization capabilities of RNNs in material constitutive modeling, researchers have begun exploring paradigms with enhanced physical interpretability. A prominent direction involves incorporating neural ordinary differential equations (NODEs) into the theoretical framework of material evolution. This approach parameterizes the instantaneous terms of dynamical systems and reconstructs material evolution trajectories through numerical integration (Tac et al., 2022). Jones et al. proposed the internal state variable neural ODE (ISV-NODE), which implicitly embeds internal variables within the NODE framework, achieving discretization-invariant predictions for viscoelastic and elastoplastic materials (Jones and Fuhg, 2025). More recently, He and Semnani introduced the incremental neural controlled differential equation (INCDE), which further incorporates stabilization mechanisms and a piecewise formulation to enhance robustness under complex loading paths (He and Semnani, 2024).

In parallel with the aforementioned paradigms, another line of research focuses on directly embedding thermodynamic principles into network architectures by introducing physics-informed constraints to enhance both generalization capability and interpretability. Examples include thermodynamics-based artificial neural networks (TANNs), which enforce thermodynamic consistency (Masi et al., 2021; Masi and Stefanou, 2022), and input-convex neural networks, which constrain the convexity of the free energy to ensure the physical validity and stability of the constitutive model (Klein et al., 2022; Thakolkaran et al., 2022).

Deep operator networks (DeepONets) and Fourier neural operators (FNOs), as emerging techniques, have achieved substantial progress in scientific artificial intelligence and operator approximation (Lu et al., 2021), offering new avenues for constitutive modeling of history-dependent materials. Guo et al. proposed the history-aware neural operator (HANO) based on the FNO framework, which constructs an autoregressive formulation to infer future stress directly from

stress-strain data within a short-term historical window (Guo et al., 2025). By eliminating hidden state propagation and conditioning predictions on recent historical segments, HANO simultaneously mitigates discretization dependence and initial state sensitivity (Guo et al., 2025). As an emerging computational paradigm, HANO exhibits considerable potential for practical applications, yet several issues warrant further investigation. For instance, enhancing the physical interpretability of the model with respect to the underlying thermomechanical mechanisms and improving its generalization to unseen loading conditions remain important directions for future research.

These successful cases provide a foundation for developing neural network constitutive models for SMPs. The defining characteristic distinguishing SMPs from conventional materials is the history dependence introduced by the shape memory effect, implying that a constitutive model must be capable of understanding and representing the complex evolution laws embedded within long-range thermomechanical loading histories. Recognizing such laws typically necessitates relatively complex models, which partly accounts for the comparatively slow progress of neural network approaches in SMP constitutive modeling. To date, early exploratory studies have achieved significant advances in this direction. Deng et al. employed an LSTM network to construct a neural network constitutive model for SMPs, demonstrating the capability of LSTM to capture the nonlinear thermomechanical evolution paths of the material (Deng and Lu, 2026). That model contained approximately 1.4×105 parameters and achieved a mean squared error (MSE) on the order of 1.4×10-5 (Deng and Lu, 2026). Diverging from the aforementioned paradigm, Yan et al. developed a physics-driven constitutive model based on physics-informed neural networks (PINNs) combined with experimental data, enabling direct mapping from strain and temperature fields to the stress response (Yan et al., 2025). This model contained only 273 parameters and, after 2.2×10^6 iterations, converged to an MSE on the order of 3.0×10^{-4} (Yan et al., 2025). Nevertheless, both the models trained by Deng et al. and Yan et al. are not explicitly designed to support path generalization; consequently, initiating predictions from arbitrary initial path conditions remains challenging, which may limit their direct applicability in certain practical computational scenarios.

Data-driven methods have also been employed for modeling other stimuli-responsive soft materials. Askari-Sedeh et al. incorporated governing equations, incompressibility conditions, and boundary conditions into the neural network loss function, establishing a physics-constrained neural network for the large-deformation electromechanical coupling response of dielectric elastomer tubes (Askari-Sedeh et al., 2025). Faraji et al. combined analytical models with augmented neural networks to predict the transient swelling and stress distribution of pH-responsive hydrogels (Faraji et al., 2025). These studies demonstrate that integrating physical relations with neural networks facilitates the handling of multiphysics coupling problems in soft materials. It should be noted, however, that the aforementioned studies primarily focus on multiphysics boundary value problems or surrogate computation of analytical models, whereas the present work investigates the local incremental evolution of material states under known internal variable conditions.

A crucial theoretical observation provides a new perspective for this work and prompts a re-examination of the limitations inherent in existing methods. A review of the aforementioned studies reveals that prevailing approaches predominantly focus on employing neural networks to approximate a complex memory system (Guo et al., 2025; Kapturowski et al., 2019; Sun et al., 2025), to fit global physical laws (Masi et al., 2021; Masi and Stefanou, 2022), or to simulate global recurrent states (He and Semnani, 2024; Jones and Fuhg, 2025; Tac et al., 2022). Such models rely on latent variables to store path sequence information and are inherently dependent on global path data. It is worth noting that although SMPs exhibit pronounced history dependence during macroscopic evolution, the evolution of the material's internal state variables obeys localized equations. This implies that the complex loading history is already fully encoded within the current internal state. This locality property suggests that the instantaneous evolution of the system depends solely on the current state, rather than on explicit reliance on the historical path. Consequently, the modeling emphasis can be shifted from fitting complex long-range memory effects toward learning local evolution laws, thereby mitigating model dependence on specific paths and discretization schemes while simultaneously reducing model complexity.

Building upon this theoretical foundation, a Local Decoupling Evolution Operator Network (LDEONet) is proposed, which theoretically achieves both path generalization and discretization generalization capabilities. By emulating the local evolution laws of internal variables, LDEONet explicitly decouples the evolution direction from the evolution intensity, thereby reducing both sample requirements and model complexity. Subsequently, the performance of the proposed model is trained and evaluated using shape memory polymers as a representative case study. Finally, a comprehensive suite of generalization experiments is designed to validate the generalization capability of LDEONet. Using the cooling and fixing process of SMPs under 1D uniaxial conditions as a numerical example, this paper conducts numerical validations on model scale, sample requirements, path generalization, and discretization generalization. This paper does not consider sequence models themselves to be redundant; rather, it investigates whether, under conditions where physical internal variables are already defined and accessible, it is possible to avoid feeding the complete historical sequence into the network as additional input.

2. Internal variable evolution and its Markovian property

2.1. Internal variable evolution model

Shape memory polymers, as a representative class of history-dependent materials, exhibit macroscopic mechanical responses that depend not only on instantaneous loading but also on the entire history of coupled thermomechanical loading. The macroscopic shape memory effect originates microscopically from the cumulative manifestation of the time-dependent evolution of the material's internal structure, i.e., the hindered motion and thermally activated processes of polymer chain segments across various energy barriers. To describe the deformation process, the macroscopic total strain tensor of SMPs, denoted by $\boldsymbol{\varepsilon}$, can be decomposed into the sum of the elastic strain tensor $\boldsymbol{\varepsilon}_e$ induced by the instantaneous stress, the thermal expansion strain tensor $\boldsymbol{\varepsilon}_T$ induced by temperature variation, and the stored strain tensor $\boldsymbol{\varepsilon}_s$ characterizing the shape memory effect:

$$\boldsymbol{\varepsilon} = \boldsymbol{\varepsilon}_e + \boldsymbol{\varepsilon}_T + \boldsymbol{\varepsilon}_s \tag{1}$$

The shape memory process of SMPs studied in this work starts from an initial stress-free heating stage. Here, the thermal strain $\boldsymbol{\varepsilon}_T$ is obtained by integrating the coefficient of thermal expansion α with respect to temperature T from a stress-free initial state at $T_0 = 393$ K:

$$\boldsymbol{\varepsilon}_T = \int_{T_0}^{T} \alpha \, \mathrm{d}T \tag{2}$$

Based on the series model derived from phenomenological theory, the deformation of SMPs can be regarded as the combined response of the glassy phase and the rubbery phase subjected to an identical stress state (Liu et al., 2006). Accordingly, the equivalent compliance tensor of the material $\boldsymbol{S}(T)$, can be expressed in terms of the compliance tensor of the glassy phase $\boldsymbol{S}_g$, and that of the rubbery phase $\boldsymbol{S}_r$, as follows:

$$\boldsymbol{S}(T) = \phi_g \boldsymbol{S}_g + \left(1 - \phi_g\right)\boldsymbol{S}_r \tag{3}$$

where ϕ_g denotes the volume fraction of the glassy phase. The equivalent stiffness tensor $\boldsymbol{C}$ and the corresponding constitutive relation can be formulated as:

$$\boldsymbol{\sigma} = \boldsymbol{C} : \boldsymbol{\varepsilon}_e = \left(\phi_g \boldsymbol{S}_g + (1-\phi_g)\boldsymbol{S}_r\right)^{-1} : \boldsymbol{\varepsilon}_e \tag{4}$$

The physical admissibility of the foregoing kinematic description and constitutive model must be rigorously constrained by the laws of thermodynamics. Based on the free energy function and the Clausius-Duhem inequality, Liu et al. derived a thermodynamically consistent local evolution law for the stored strain $\boldsymbol{\varepsilon}_s$ (Liu et al., 2006).

$$\frac{\mathrm{d}\boldsymbol{\varepsilon}_s}{\mathrm{d}T} = \boldsymbol{S}_e(T) : \boldsymbol{\sigma}(T)\frac{\mathrm{d}\phi_g}{\mathrm{d}T}, \tag{5}$$

where $\boldsymbol{S}_e(T) = 1/3NkT\boldsymbol{I}$ denotes the entropic compliance tensor, with N being the crosslink density, k the Boltzmann constant, and $\boldsymbol{I}$ the fourth-order identity tensor. To effectively utilize the experimental data reported by Li (F. Li, 2020), the present study approximates the entropic compliance tensor by the experimentally measured compliance tensor of the rubbery phase. Consequently, the evolution law for the internal variable $\boldsymbol{\varepsilon}_s$ adopted herein is expressed as:

$$\frac{\mathrm{d}\boldsymbol{\varepsilon}_s}{\mathrm{d}T} = \boldsymbol{S}_r(T) : \boldsymbol{\sigma}(T)\frac{\mathrm{d}\phi_g}{\mathrm{d}T} \tag{6}$$

2.2. Markovian property of the SMP evolution equations

Although the coupled thermomechanical evolution of SMPs is history-dependent, the evolution law governing the internal variables is inherently history-independent. Substituting equation (4) into equation (6) to eliminate the stress tensor $\boldsymbol{\sigma}$ yields:

$$\frac{\mathrm{d}\boldsymbol{\varepsilon}_s}{\mathrm{d}T} = \boldsymbol{S}_r : \left(\phi_g \boldsymbol{S}_g + (1-\phi_g)\boldsymbol{S}_r\right)^{-1} : \boldsymbol{\varepsilon}_e\frac{\mathrm{d}\phi_g}{\mathrm{d}T} \tag{7}$$

Treating the glassy phase compliance $\boldsymbol{S}_g$ and the rubbery phase compliance $\boldsymbol{S}_r$ as constants, the mapping relation $\mathscr{F}$ governing the internal variable evolution can be summarized as follows:

$$\frac{\mathrm{d}\boldsymbol{\varepsilon}_s}{\mathrm{d}T} = \mathscr{F}\left(\boldsymbol{\varepsilon}_e, \phi_g, \frac{\mathrm{d}\phi_g}{\mathrm{d}T}\right) \tag{8}$$

In practical incremental computations, the current state variables are obtained from the previous increment step and stored at the corresponding material points or finite element integration points. Therefore, the method still follows a time-recurrence relation and relies on the selected internal variables being sufficient to fully describe the current material state. What this paper emphasizes is that, once the current internal variables are available, predicting the next state does not require additional input of the complete historical sequence from the initial time or a history window of fixed length.

Assuming that the evolution of the internal variable constitutes a dynamical system $\mathscr{X}(T)$ defined over time, and in accordance with the mapping relation for the internal variable, the dynamical system can be expressed as:

$$\mathscr{X}(T) = \mathscr{F}(T) \tag{9}$$

According to the fundamental theorem of calculus, the future state of the internal variable can be obtained by integrating the current dynamical system:

$$\boldsymbol{\varepsilon}_s(T+\Delta T) = \boldsymbol{\varepsilon}_s(T) + \int_T^{T+\Delta T} \mathscr{X}(\tau)\,\mathrm{d}\tau \tag{10}$$

If the dynamical system possesses the Markovian property, the evolution mapping relation $\mathscr{F}$ must constitute a purely local algebraic function, devoid of any internal memory integral kernels or hereditary functions that depend on the historical time $t \in (-\infty, \tau)$. From the perspective of dynamical systems theory, this implies that the evolved state of the system at time $t+\Delta t$, denoted by $\mathscr{X}_{t+\Delta t}$, can be uniquely determined by the instantaneous characteristics $\mathscr{X}_t$ at time t and the variation occurring within that time increment. The future evolution trajectory of the system is independent of the specific historical loading path $\{\mathscr{X}_{t-\Delta t}, \ldots, \mathscr{X}_0\}$ traversed prior to reaching the current state $\mathscr{X}_t$. Evidently, the mapping relation $\mathscr{F}$ governing the internal variable evolution of shape memory polymers satisfies this property and therefore exhibits the Markovian property.

The aforementioned local state description has well-defined conditions under which it holds. Suppose the material state is represented by a finite-dimensional vector or tensor internal variable $\boldsymbol{\xi}$. If these variables can adequately describe the current material state and satisfy a closed first-order evolution relation:

$$\dot{\boldsymbol{\xi}} = \mathscr{G}(\boldsymbol{\xi}, \boldsymbol{F}, T, \ldots) \tag{11}$$

Then, given the current state and the current external inputs, the evolution of internal variables at the next time step does not require additional invocation of the complete historical sequence. Here, $\boldsymbol{F}$ denotes the deformation gradient (or other kinematic quantities), and T denotes temperature. In this case, the local evolution operator can be written in a general vector or tensor form. Classical elastoplasticity, SMPs described by a finite number of internal variables, and viscoelastic models approximated by a finite number of Prony terms typically satisfy this condition. For materials with a continuous relaxation spectrum, spatially nonlocal damage, fractional-order memory, or those whose history cannot be adequately described by the current finite internal variables, the finite-dimensional local state assumption of this paper does not hold directly. Therefore, the subsequent conclusions of this paper are only applicable to materials that can be described by a finite-dimensional closed state formulation.

Within the framework of material constitutive theory, many path-dependent materials are often described using integral-type memory models, the general form of which is given by:

$$\boldsymbol{\sigma}(t) = \int_0^t K(t-s)\,\mathscr{F}(\boldsymbol{\varepsilon}(s), T(s))\,\mathrm{d}s \tag{12}$$

where $K(t)$ denotes the memory kernel function, which characterizes the influence of the historical loading on the current response. Models of this form exhibit an explicit dependence on the complete historical path and thus constitute a typical non-local temporal description of material evolution.

However, for a class of memory kernel functions possessing a finite spectral expansion. For instance, those representable as a superposition of exponential functions:

$$K(t) \approx \sum_{i=1}^{N} a_i e^{-\frac{t}{\tau_i}} \tag{13}$$

the corresponding hereditary integral representation can be equivalently transformed into a set of evolution equations for a finite number of internal variables:

$$\dot{\boldsymbol{z}}_i = \frac{1}{\tau_i}(\mathscr{F}(\varepsilon, T) - \boldsymbol{z}_i), i = 1, \ldots, N. \tag{14}$$

This transformation recasts the model, which originally depended on the complete loading history, into a local dynamical system that is closed in the state space. Under these circumstances, the evolution of the material system satisfies the Markovian property: its future response depends solely on the current state variables and no longer requires explicit reference to the historical loading path.

The Markovian property of the internal variable model for SMPs inspires the pursuit of generalization to unseen loading paths. Path-dependent, data-driven models, represented by RNNs, LSTMs, and HANO, possess inherent advantages in capturing the history dependence

of materials by virtue of their specialized memory mechanisms. Through deep learning on large volumes of data, these models can circumvent cumbersome constitutive derivations and implicitly fit the nonlinear coupled thermomechanical constitutive mapping relations. However, models trained on sequential data tend to be influenced, from an architectural standpoint, by the path dependence and discretization dependence present in the training data. Consequently, achieving robust path and discretization generalization with such architectures remains challenging. HANO systematically summarized the limitations of models such as RNNs and LSTMs, establishing a predictive model that relies on sequential windows. This approach alleviates the dependence on global sequence data and achieves a certain degree of path generalization capability (Guo et al., 2025).

To overcome the intrinsic deficiencies of conventional models, Neural Ordinary Differential Equations (NODEs) and their variant, Internal State Variable Neural ODEs (ISV-NODEs), embed neural networks into the right-hand side of differential equations, thereby incorporating the instantaneous evolution of internal variables within a numerical integration framework (He and Semnani, 2024; Jones and Fuhg, 2025; Tac et al., 2022). This formulation approximates the entire thermomechanical loading history and thus attains discretization generalization capability. Thermodynamics-based Artificial Neural Networks (TANNs) integrate thermodynamic scalar potentials and constraint equations into the model architecture, significantly enhancing the thermodynamic consistency of the implicit internal variables (Masi et al., 2021). Physics-Informed Neural Networks (PINNs) employ evolution equations that satisfy thermodynamic consistency as loss functions to directly train constitutive relations for SMPs (Yan et al., 2025). Neural network models such as PINNs, NODEs, and TANNs perform notably well in scenarios characterized by relatively sparse data availability. By leveraging physical constraints—including the laws of thermodynamics, free energy dissipation mechanisms, or internal variable evolution equations—these models are capable of rigorously preserving the physical consistency of their predictions. Nevertheless, despite adopting local evolution laws and explicit physical models in their mathematical formulation, the training process of these models typically depends on full historical path data. This implicit reliance on global training data implies that the models remain fundamentally constrained by the loading history.

3. Model, sample generation, and training

3.1. Local Decoupling Evolution Operator Network

In the field of operator learning, DeepONet achieves universal approximation of nonlinear operators through a decoupled dual-network architecture comprising a branch network and a trunk network (Lu et al., 2021). This architectural philosophy of decomposing complex laws into two independent information streams provides important inspiration for the present work. By abstracting the complex evolution law of internal variables as an operator mapping problem on the tangent bundle of the state space, and by explicitly decoupling the evolution direction D from the evolution intensity M, the 1D internal variable evolution equation for SMPs can be formulated as:

$$\frac{\mathrm{d}\varepsilon_s}{\mathrm{d}T} = \mathscr{T}\left(\varepsilon_e, \phi_g, \frac{\mathrm{d}\phi_g}{\mathrm{d}T}\right) = D(\boldsymbol{p}) \bowtie M(\boldsymbol{q}) = D\left(\varepsilon_e, \phi_g\right) \bowtie M\left(\frac{\mathrm{d}\phi_g}{\mathrm{d}T}\right) \tag{15}$$

where $\bowtie$ denotes a coupling operation between direction and intensity. As illustrated in Fig. 1, from the perspective of differential geometry, if the ensemble of all possible physical states of the material (e.g., strain, phase fraction) is defined as a high-dimensional state manifold $\mathscr{L}$, then the loading history of the material can be represented as a continuous curve on the manifold $\mathscr{L}$. Based on the local evolution nature elucidated previously, the evolution direction $\mathrm{d}\boldsymbol{\Gamma}/\mathrm{d}T$ of this curve at any point should be entirely determined by the location of that point on the manifold, the instantaneous driving intensity applied at that point, and the coupling mode. Accordingly, the Local Decoupling Evolution Operator Network (LDEONet) is proposed and expressed as:

$$\frac{\mathrm{d}\boldsymbol{\Gamma}}{\mathrm{d}T} = \mathrm{LDEONet}(\boldsymbol{p}, \boldsymbol{q}) = D(\boldsymbol{p}) \bowtie M(\boldsymbol{q}) \tag{16}$$

where $\boldsymbol{p}$ and $\boldsymbol{q}$ denote the model prediction parameter vectors. Based on the obtained instantaneous evolution rate, LDEONet is embedded within a temporal integration scheme, enabling the determination of the stored strain at the subsequent temperature increment ΔT as follows:

$$\varepsilon_s(T + \Delta T) = \varepsilon_s(T) + \int_T^{T+\Delta T} \frac{\mathrm{d}\boldsymbol{\Gamma}(\tau)}{\mathrm{d}\tau}\, \mathrm{d}\tau \tag{17}$$

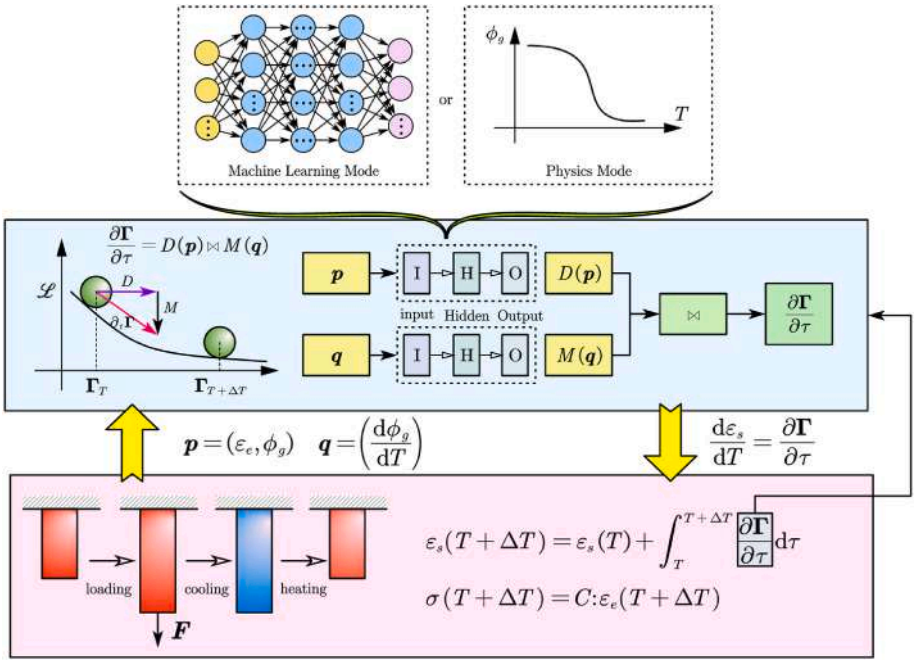


Fig. 1. Schematic illustration of the LDEONet model architecture.

Subsequently, based on the stored strain, elastic strain, and thermal strain at the next step, the corresponding stress can be determined as:

$$\sigma(T+\Delta T) = C{\cdot}\varepsilon_e(T+\Delta T) \tag{18}$$

Fig. 1 illustrates the model architecture and the underlying modeling philosophy of LDEONet, namely, utilizing a machine learning model to replace the core evolution model within a complex physical process to achieve more generalizable and fundamental predictions. LDEONet endows the purely data-driven model for the internal variable evolution law of shape memory polymers with a distinct differential geometric interpretation. Through the decoupling of evolution direction and evolution intensity, the physical interpretability of the model is substantially enhanced. Consequently, LDEONet constitutes a neural network model with an inherently physics-structured architecture, rather than merely embedding physical models into the loss function. Furthermore, this architectural-level decoupling imparts modular characteristics to the model: both the evolution direction and the evolution intensity can be directly replaced or augmented by specific physical models according to research requirements. This implies that LDEONet can transform established knowledge into pluggable model modules, thereby unlocking the generalization potential of data-driven approaches while preserving physical determinism.

In the explicit finite element analysis framework, LDEONet can serve as a surrogate model for the internal variable evolution law and be directly embedded into the constitutive update procedure at each integration point. Specifically, within each increment step, given the elastic strain ε_e at the current temperature T, the glass transition fraction ϕ_g, and its derivative with respect to temperature $d\phi_g/dT$, the evolution rate $d\varepsilon_s/dT$ of the internal variable ε_s is predicted by LDEONet as expressed in Eq. (15). Subsequently, the temperature integration scheme is employed to update ε_s to the next temperature $T+\Delta T$ according to Eq. (17). Finally, the stress tensor at this integration point is computed from the updated elastic strain using Eq. (18).

Throughout this process, each integration point only needs to store the current internal variable values, without recording the global loading history, thereby significantly reducing storage overhead and naturally aligning with the explicit finite element solution procedure. Meanwhile, the decoupled architecture of LDEONet allows flexible incorporation of physical prior knowledge into both the evolution direction and evolution intensity modules, providing a constitutive update strategy for finite element simulations that combines interpretability with generalization capability.

3.2. Multi-modal limitations and improvements

Under typical service conditions, shape memory polymers usually undergo a complete thermomechanical coupling cycle, including high-temperature loading, cooling fixation, low-temperature unloading, and heating recovery. In the study by Liu et al., cooling fixation and heating recovery obey different physical laws (Liu et al., 2006). Under thermomechanical coupling, the evolution direction of internal variables follows the law:

$$D_{\text{cooling}} = \boldsymbol{S}_r : \boldsymbol{\sigma} \tag{19}$$

$$D_{\text{heating}} = \frac{\boldsymbol{\varepsilon}_s}{\phi_g} \tag{20}$$

Where D_{cooling} is the evolution direction under cooling fixation, and D_{heating} is the evolution direction under heating recovery. The evolution intensity of the thermomechanical coupling cycle is consistent:

$$M_{\text{cooling}} = M_{\text{heating}} = \frac{\mathrm{d}\phi_g}{\mathrm{d}T} \tag{21}$$

LDEONet directly learns the differential laws and is a single-physical-mode model; therefore, it does not inherently possess the capability to learn the complete thermomechanical coupling cycle. In contrast, sequence-learning model frameworks such as LSTM, which directly learn from sequences, are capable of capturing the entire thermomechanical coupling cycle. To endow LDEONet with the ability to handle multiple physical modes, a physics-mode-based gating inference strategy is adopted. Specifically, the evolution direction D is trained separately for the two physical modes, and the model switches between these modes depending on the current physical regime. More concretely, a gated thermodynamic phase discriminator $G(S)$ is employed:

$$G(S) = \begin{cases} 1, S \in \textit{cooling} \\ 0, S \in \textit{heating} \end{cases} \tag{22}$$

Here, $S = \{\dot{T}\}$ denotes the current thermal state of the material, which is discriminated by the current temperature direction $\dot{T}$. Therefore, the physics-mode-based LDEONet is transformed into:

$$\text{LDEONet}(\boldsymbol{p},\boldsymbol{q}) = G(S){\cdot}\text{LDEONet}(\boldsymbol{p},\boldsymbol{q})_{cooling} + (1-G(S)){\cdot}\text{LDEONet}(\boldsymbol{p},\boldsymbol{q})_{heating} \tag{23}$$

Clearly, in mechanical stages that do not involve temperature variations, there is no phase transition; therefore, LDEONet does not participate in the evolution computation.

3.3. Sample generation

As a representative class of compressible history-dependent materials, SMPs exhibit complex spatiotemporal evolution characteristics, making them an ideal testbed for evaluating the path generalization and discretization generalization capabilities of LDEONet. Within the shape memory cycle, the cooling and fixing stage essentially constitutes a nonlinear evolution and storage process of the material's internal state variables, thereby epitomizing the history-dependent nature of the material. In this study, the thermomechanical cycle of epoxy-based shape memory polymers is selected as the research subject to validate the comprehensive capability of the proposed model. The finite difference method is employed to numerically simulate the stored strain evolution equation of SMPs, thereby generating a one-dimensional sample database. The numerical scheme is expressed as:

$$\frac{\varepsilon_s^{n+1} - \varepsilon_s^n}{\Delta T} = \mathscr{F}^n \tag{24}$$

where ε_s^{n+1} denotes the stored strain tensor at the subsequent temperature increment. As a purely data-driven framework, LDEONet can achieve higher predictive accuracy upon acquiring more precise data. In the subsequent training and validation, all data were obtained from numerical simulations. Fig. 2 presents the temperature–stress experimental and numerical simulation data for the cooling–fixation process, while Fig. 3 shows the temperature–stress, temperature–strain, and strain–-stress experimental and numerical simulation data during the thermomechanical cycle of the material. The experimental results are taken from the study by Li (2020). The comparison results demonstrate that the simulated data used for the samples are close to the experimental data and can therefore be employed for training the neural network models.

Benefiting from the explicit decoupling of the local evolution law, LDEONet can leverage physical models to reduce model complexity. Specifically, the relationship between the glassy phase fraction and temperature $\phi_g = \phi_g(T)$, has been experimentally determined, from which $\mathrm{d}\phi_g/\mathrm{d}T$ can be derived:

$$\phi_g(T) = \frac{1}{1+\exp\left(a\left(T-T_g\right)\right)} \tag{25}$$

Where T_g is the glass transition temperature of the material, and $a = 0.1797$ is the model coefficient. Consequently, the evolution intensity M

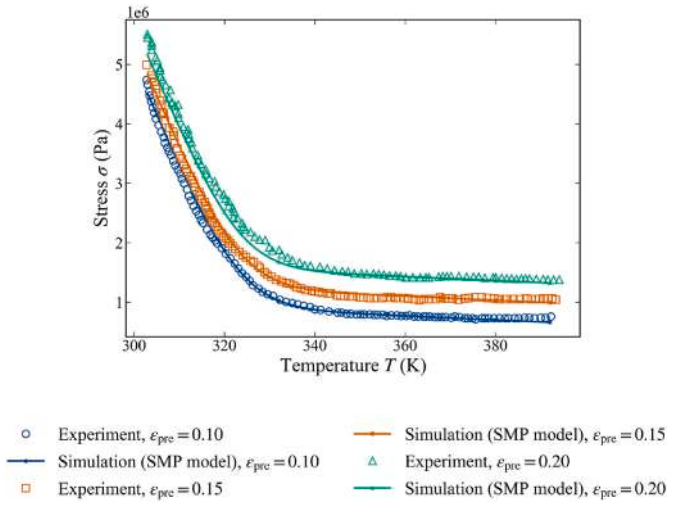


Fig. 2. Experimental (Li, 2020) and numerical results (this paper) for the cooling-fixation process.

can be replaced by a physical model, and LDEONet is specialized as:

$$\frac{\mathrm{d}\varepsilon_s}{\mathrm{d}T} = \mathrm{LDEONet}(\boldsymbol{p},\boldsymbol{q}) = D(\boldsymbol{p}) \bowtie \boldsymbol{M}(\boldsymbol{q}) = D(\varepsilon_e,\phi_g)\ \frac{\mathrm{d}\phi_g}{\mathrm{d}T} \tag{26}$$

Therefore, in practical implementation, LDEONet only needs to train the evolution direction. Considering the differences in magnitude among various data quantities, the sample data are normalized by scaling. The sample database includes both the cooling–fixation process and the heating–recovery process; isothermal loading and isothermal unloading processes are not included in the database. The range of pre-strain in the sample database is $\varepsilon_{pre} \in [0.01, 0.2]$ with a pre-strain increment of $\Delta\varepsilon_{pre} = 0.01$, resulting in a total of 20 thermomechanical cycle samples. The temperature ranges from $T \in [303, 393]$ K, with a temperature step of $\Delta T = -1K$ for the cooling–fixation process and $\Delta T = 1K$ for the heating–recovery process. The sample database contains a total of 3600 state samples, comprising 1800 state points from the cooling–fixation process and 1800 state points from the heating–recovery process, with each thermomechanical cycle consisting of 90 temperature steps. The sample database is randomly divided into training, validation, and test sets in a ratio of 8:1:1. Finally, each dataset is randomly shuffled to prevent sequential bias, encouraging the model to learn mapping mechanisms independent of sequence order. The heating–recovery model and the cooling–fixation model are trained separately on their respective data, with no cross-training between them.

3.4. Training

To train and test LDEONet, an internally developed materials artificial intelligence platform was utilized. This platform was developed based on reverse-mode automatic differentiation and tape-recording techniques. The Mean Squared Error (*MSE*) function is adopted as the loss function to quantify the discrepancy between predicted and actual values:

$$MSE = \frac{1}{N}\sum_{i=1}^{N}\left(y_i^{pred} - y_i^{ref}\right)^2 \tag{27}$$

Here, y_i^{pred} denotes the predicted data, and y_i^{ref} denotes the reference data. The model parameters are optimized using the Adam algorithm with a learning rate of 10^{-3}, trained for 10^5 epochs without weight decay. Considering the drastic numerical variations in the glassy phase volume fraction during the phase transition, conventional saturating activation functions tend to induce the vanishing gradient problem. Therefore, the Rectified Linear Unit (ReLU) is adopted as the hidden-layer activation function, leveraging its unbounded positive interval to prevent gradient vanishing. Combined with a linear output layer, this ensures convergence of the model in learning the internal variable evolution law.

To test different neuron configurations and examine the demand of LDEONet on model parameter count, four configurations with hidden-layer neuron numbers of 5, 10, 20, and 30 are designed. To accommodate the two physical modes, a total of eight models are trained, namely LDEONet-C1, LDEONet-C2, LDEONet-C3, LDEONet-C4 and LDEONet-H1, LDEONet-H2, LDEONet-H3, LDEONet-H4. Among them, LDEONet-C1 to LDEONet-C4 are cooling–fixation models, and LDEONet-H1 to LDEONet-H4 are heating–recovery models. Fig. 4 presents the loss function convergence curves for all models, showing both training and validation sets. Table 1 lists the mean squared errors (MSE) of all models on the test set. From Fig. 4 and Tables 1 and it can be observed that the MSEs of LDEONet-C4 and LDEONet-H4 reach 2.83×10^{-4} and 3.49×10^{-3}, respectively. Therefore, these two models are selected as the baseline models for subsequent discussion. LDEONet-C4 and LDEONet-H4 each have 121 trainable parameters, amounting to a total of 242 trainable parameters for the complete thermomechanical cycle prediction.

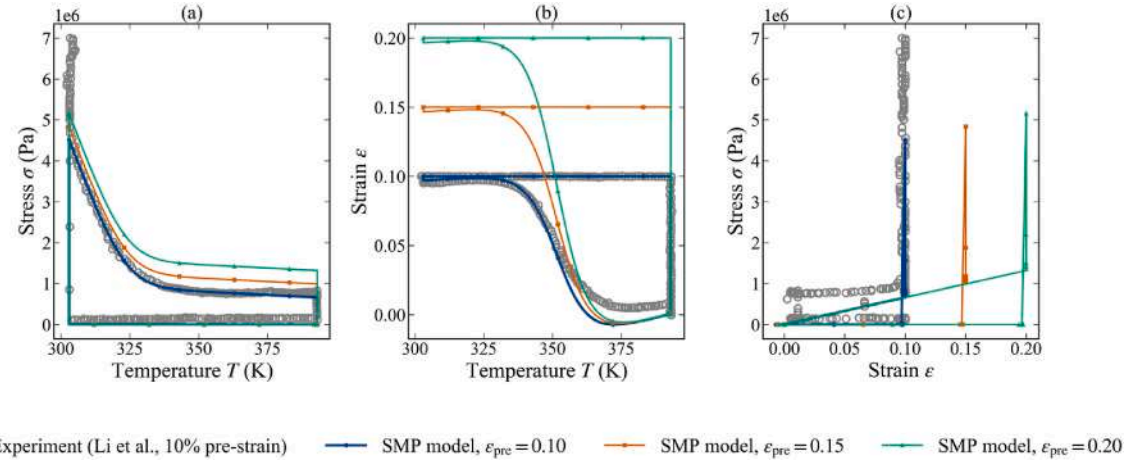


Fig. 3. Experimental (Li, 2020) and numerical results (this paper) for the thermomechanical cycle: (a) temperature–stress results; (b) temperature–strain results; (c) strain–stress results.

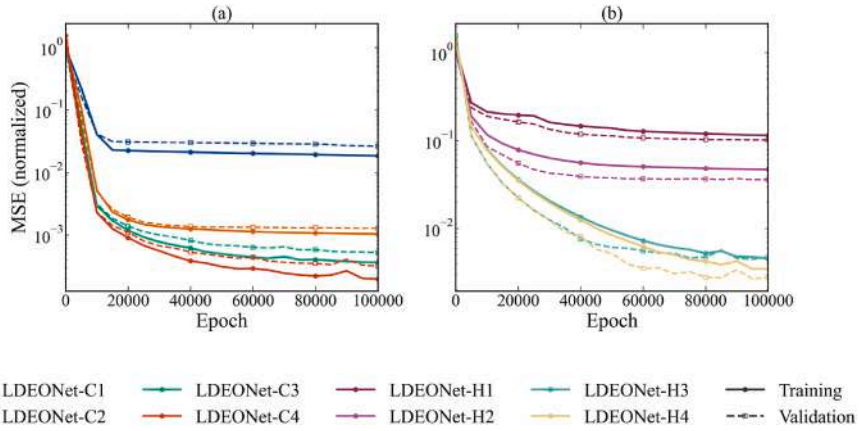


Fig. 4. Loss function variation during training: (a) LDEONet-C1, LDEONet-C2, LDEONet-C3, LDEONet-C4; (b) LDEONet-H1, LDEONet-H2, LDEONet-H3, LDEONet-H4.

Table 1
Test set errors of the models.

Model	Configuration	*MSE*
LDEONet-C1	1 hidden layer, 5 neurons	2.20×10^{-2}
LDEONet-C2	1 hidden layer, 10 neurons	1.23×10^{-3}
LDEONet-C3	1 hidden layer, 20 neurons	4.01×10^{-4}
LDEONet-C4	1 hidden layer, 30 neurons	2.83×10^{-4}
LDEONet-H1	1 hidden layer, 5 neurons	1.15×10^{-1}
LDEONet-H2	1 hidden layer, 10 neurons	4.70×10^{-2}
LDEONet-H3	1 hidden layer, 20 neurons	4.53×10^{-3}
LDEONet-H4	1 hidden layer, 30 neurons	3.49×10^{-3}

4. Discussion on generalization performance

In this section, a series of experiments are designed to evaluate the generalization prediction capability of LDEONet-C4/H4, including tests on unseen loading paths, unseen step sizes, and decoupling ablation studies, which respectively validate the path generalization, discretization generalization, and the effect of model decoupling on generalization. After validation, LDEONet-C4/H4 is compared horizontally with existing similar models to clarify its generalization, limitations, and applicable scope. The prediction performance of the physical quantities for SMP materials is uniformly evaluated by the relative error, defined as:

$$RE = \frac{\frac{1}{N}\sum_{i=1}^{N}\left|y_i^{pred} - y_i^{ref}\right|}{\frac{1}{N}\sum_{i=1}^{N}\left|y_i^{ref}\right|} \times 100\% \tag{28}$$

where y_i^{pred} denotes the predicted data, and y_i^{ref} denotes the reference data.

4.1. Prediction capability for unseen paths

To demonstrate the prediction capability of LDEONet for unseen paths, path generalization experiments are designed using LDEONet-C4/H4 as the predictive model, with the thermomechanical cycle process as the experimental variable. Without retraining the model, the predictive accuracy is examined when the thermomechanical path undergoes abrupt changes. The training sample database contains only two monotonic temperature processes, in which the temperature always changes monotonically and never undergoes alternating cooling and heating sequences. Therefore, any path that involves temperature reversals possesses internal state sequences that do not exist in the training data, and thus constitutes a strictly unseen path. The specific experimental scheme is as follows:

(1) Pre-strains $\varepsilon_{\text{pre}} = 0.05, 0.10, 0.15$.
(2) Predictions are performed using the LDEONet-C4/H4 mode-switching strategy (LDEONet-C4 for the cooling stage, LDEONet-H4 for the heating stage).
(3) The experimental content consists of three reversal paths, with progressively increasing reversal depth and frequency in the thermomechanical cycles of the material. Path A is:393→363→378→303 K single reversal: cooling for 30 steps, heating for 15 steps, cooling for 75 steps. Path B is:393→333→363→303 K deep reversal: cooling for 60 steps, heating for 30 steps, cooling for 60 steps. Path C is:393→353→368→338→353→303 K double reversal: cooling for 40 steps, heating for 15 steps, cooling for 30 steps, heating for 15 steps, cooling for 50 steps. The reference data are the numerical solutions obtained by the finite difference method.
(4) For the prediction step sizes, the same $|\Delta T| = 1$ K sequence is used for both cooling and heating.

In Fig. 5, the path generalization experiments demonstrate that LDEONet-C4/H4 can accurately predict unseen paths. In Fig. 5(a)–(f), the predicted curves maintain good agreement with the reference curves in terms of the stress recovery during cooling, the stress returning to zero during heating, and the transitions at the two reversal points; the strain curves are continuous throughout the entire temperature history and closely follow the reference curves. Fig. 6 summarizes the global relative errors for all cases. The relative errors of stress during the cooling stage range from 0.674% to 2.674%, while those of strain during the heating stage range from 0.943% to 1.841%. As the reversal depth and frequency increase (from Path A to C), the errors increase slightly but remain within 3%. These error results imply that even when the temperature history undergoes multiple reversals and the historical path experienced by the material is completely altered, LDEONet can still achieve accurate subsequent predictions solely based on the current instantaneous state, without requiring any historical sequence or history window. The above results experimentally confirm the theoretical analysis in Section 2: the evolution law of internal variables in SMPs possesses the Markovian property, and the complex loading history is fully encoded in the current finite-dimensional internal variable state. Unlike RNN/LSTM that rely on global sequence matching and HANO that relies on history window matching, LDEONet's learning target is only the local evolution direction D, and its path generalization capability arises from the architecture itself rather than from memorizing training paths. Notably, the reversal paths cause the cooling and heating physical modes to alternate within the same temperature range, requiring the model to rely on gating switching to correctly transition

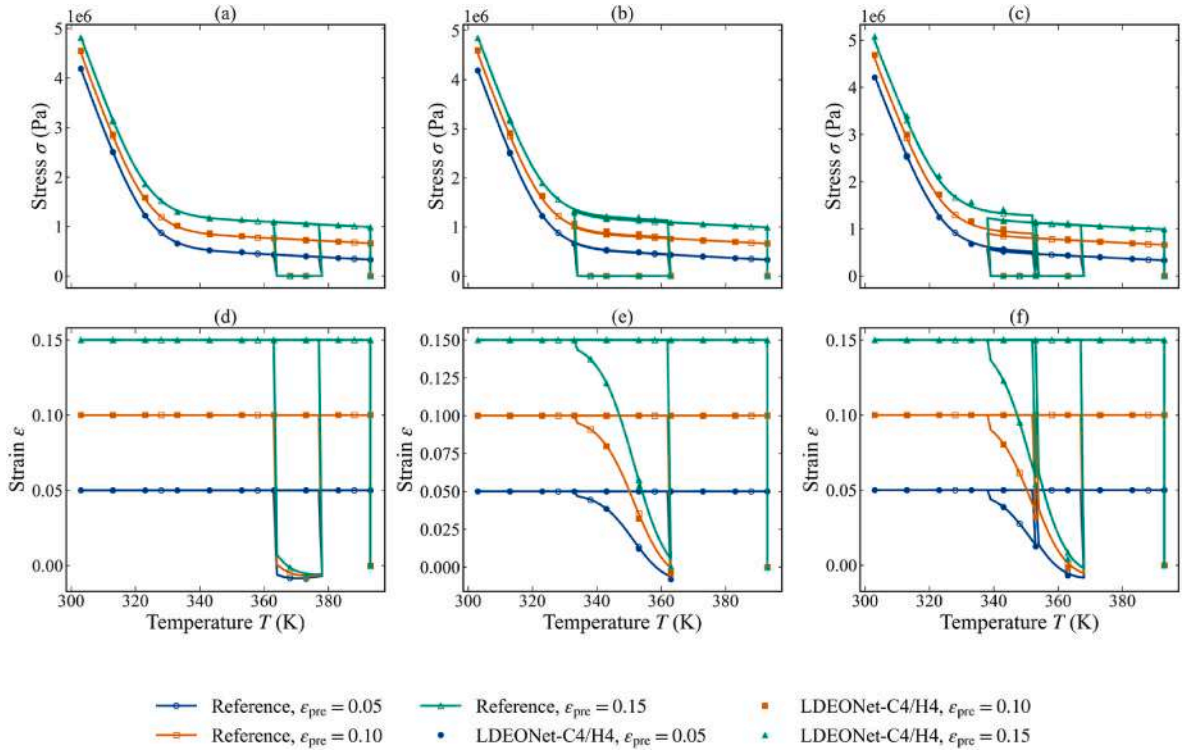


Fig. 5. Path generalization experiments: (a)-(c) temperature-stress results; (d)-(f) temperature-strain results; the columns correspond to the three reversal paths (Path A, Path B, and Path C), respectively.

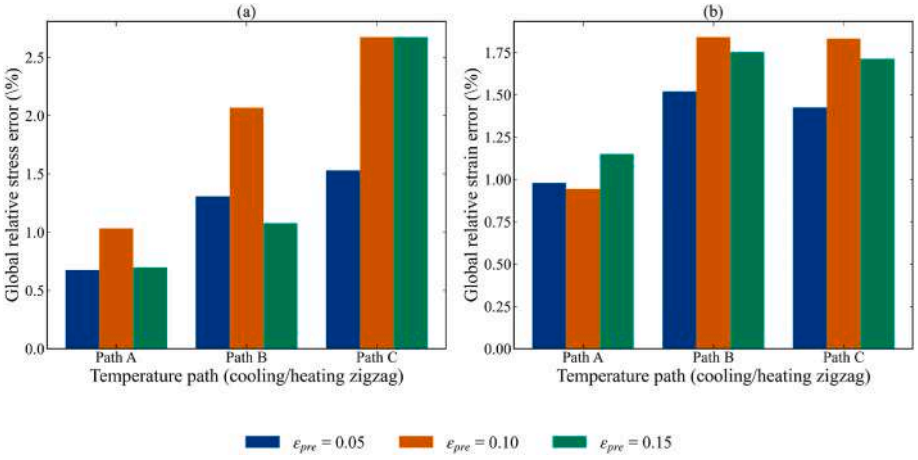


Fig. 6. Path generalization relative errors: (a) global relative stress error; (b) global relative strain error.

between the two branches. The experiments demonstrate that the mode-switching strategy can operate stably on unseen paths, further illustrating the adaptability of local evolution modeling to complex temperature histories.

4.2. Prediction capability for unseen step sizes

To evaluate the prediction generalization capability of LDEONet under unseen temperature step sizes, this paper designs four unseen temperature increment schemes: constant, random, sinusoidal, and stepwise. It should be emphasized that the discretization generalization capability discussed here refers to the performance of the model within the test range and generalization experiments of this paper, rather than a rigorous proof that it remains invariant to arbitrary temperature increments or arbitrary numerical discretization schemes. The specific experimental scheme is as follows:

(1) Pre-strains $\varepsilon_{\text{pre}} = 0.05, 0.10, 0.15$;
(2) Predictions are performed using the LDEONet-C4/H4 mode-switching strategy (LDEONet-C4 for the cooling stage, LDEONet-H4 for the heating stage);
(3) The experimental content consists of a conventional thermo-mechanical cycle of the material (high-temperature loading at 393 K → cooling fixation to 303 K → isothermal unloading at 303 K → heating recovery to 393 K), with the reference data being the numerical solutions obtained by the finite difference method.

(4) In the prediction step sizes, the same $|\Delta T|$ sequence is used for both cooling and heating, with the number of steps $N = 90$. The temperature span for each scheme is fixed at 90 K to ensure the same final temperature. Among them, the constant scheme serves as the control group, with $|\Delta T|_n = 1$ K, consistent with the sample database. The random scheme is the first experimental group, where the step sizes follow a uniform distribution:

$$|\Delta T|_n = 90 \cdot \frac{w_n}{\sum_{k=1}^{90} w_k} \tag{29}$$

Where the weights $w_n \sim U(0.1, 2.0)$ follow a uniform distribution, to examine the model's robustness to large random perturbations. The sinusoidal scheme is the second experimental group, in which the step size varies continuously as a sinusoidal function of the integration step number.

$$|\Delta T|_n = 1 - 0.8 \cdot \sin(2\pi n/90) \tag{30}$$

This scheme is designed to examine the numerical stability of the model under continuously varying step sizes. The stepwise scheme is the third experimental group, in which the step size changes in a piecewise stepwise manner, to evaluate the model's transient response to abrupt step-size changes.

$$|\Delta T|_n = \begin{cases} 1.5, \mathrm{n} \le 30 \\ 0.2, 30 < \mathrm{n} \le 60 \\ 1.3, \mathrm{n} > 60 \end{cases} \tag{31}$$

Fig. 7 presents the complete cycle prediction results of LDEONet-C4/H4 under the four schemes. In the figure, the first column shows the step-size sequences for each scheme, the second to fourth columns present the temperature–stress, temperature–strain, and strain–stress results, respectively, and the first to fourth rows correspond to the constant, random, sinusoidal, and stepwise schemes. The data show that, although the perturbation magnitudes of the step sizes in the random, sinusoidal, and stepwise schemes are far larger than the training step size, the predicted curves of LDEONet-C4/H4 for unseen step sizes remain highly consistent with the reference solutions, correctly predicting the complete shape memory cycle with no abnormal jumps at the reversal transitions.

Fig. 8 gives the global relative errors of stress during the cooling stage and strain during the heating stage for each scheme. After averaging the results over the three pre-strain cases, the relative stress errors for the constant, random, sinusoidal, and stepwise schemes are 0.746%, 0.741%, 1.025%, and 0.825%, respectively, which are almost on par with the training step-size control group (0.746%). The relative strain errors are 0.887%, 1.064%, 1.070%, and 1.754%, respectively; only the

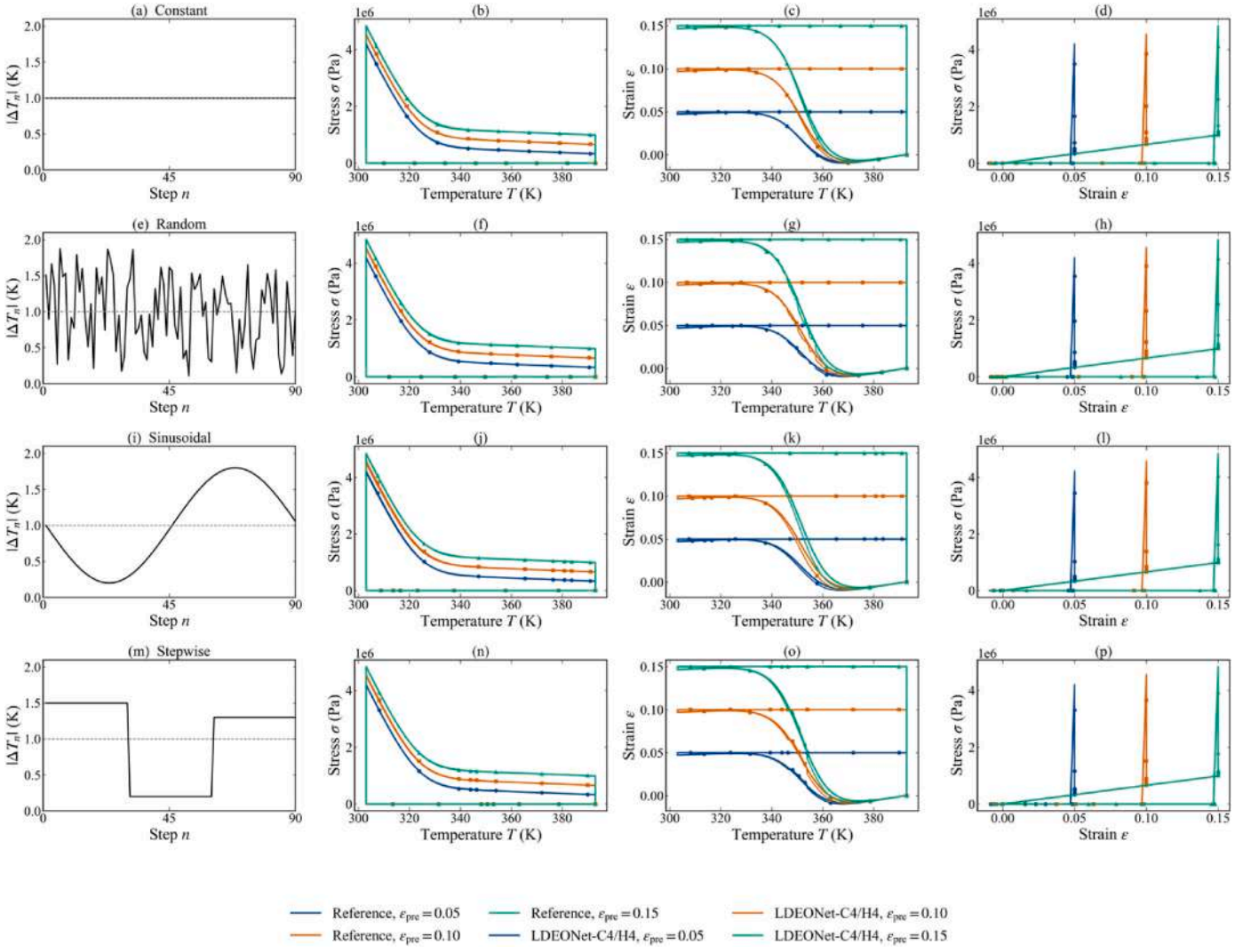


Fig. 7. Discretization generalization experiments: (a, e, i, m) step size sequences; (b, f, j, n) temperature–stress results; (c, g, k, o) temperature–strain results; (d, h, l, p) strain–stress results.

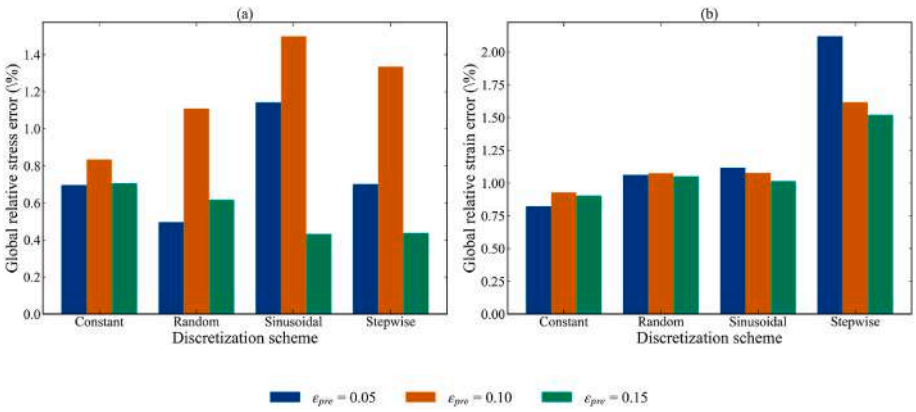


Fig. 8. Discretization generalization relative errors: (a) global relative stress error; (b) global relative strain error.

stepwise scheme shows a slightly higher value due to the dense sampling points in the heating stage caused by the extremely small step size of 0.2 K in the middle segment, but it still remains within 2%. This implies that, within the range of temperature increments tested in this paper, even when the step size undergoes large random fluctuations, continuous sinusoidal variations, or discontinuous step changes, the prediction accuracy of LDEONet-C4/H4 is almost unaffected.

4.3. Decoupling ablation study

In the LDEONet framework proposed in this paper, the evolution direction D and evolution intensity M serve as the key structural guarantees for achieving generalization capability. LDEONet models the direction and intensity of the thermomechanical evolution separately, thereby reducing model complexity and sample requirements, enabling the model to learn more fundamental thermomechanical evolution laws of the material while isolating the coupling effects between direction and intensity. To demonstrate that the decoupling design contributes to enhancing the generalization capability of the LDEONet framework, a decoupling ablation study is designed in this section, as detailed below:

(1). Pre-strain $\varepsilon_{\text{pre}} = 0.10$.
(2). The prediction using the LDEONet-C4/H4 mode-switching strategy (LDEONet-C4 for the cooling stage, LDEONet-H4 for the heating stage) serves as the control group model, while the experimental group model is the coupled model.

$$\frac{\mathrm{d}\boldsymbol{\epsilon}_s}{\mathrm{d}T} = Net(\boldsymbol{p}, \boldsymbol{q}) = Net\left(\boldsymbol{\epsilon}_e, \phi_g, \frac{\mathrm{d}\phi_g}{\mathrm{d}T}\right) \tag{32}$$

The same mode-switching strategy is adopted. In the experimental group, the evolution direction and evolution intensity are combined into a single mapping to be learned directly, without explicit decoupling. To ensure a fair comparison, the control model adopts a mode-switching architecture identical to that of LDEONet. Each branch is a 3-input single-hidden-layer network with 24 neurons in the hidden layer, containing 121 trainable parameters, which is comparable to the parameter count of a single branch of LDEONet-C4/H4. Therefore, any differences in generalization capability between the two can be attributed to the decoupling design itself.

(3). The experimental content consists of a conventional thermomechanical cycle of the material (high-temperature loading at 393 K → cooling fixation to 303 K → isothermal unloading at 303 K → heating recovery to 393 K). The reference solutions are obtained by advancing the finite difference model within the same integration framework to ensure point-by-point comparability.
(4). For the prediction step sizes, constant sequences of $|\Delta T| = 1$ K, 2 K, 3 K, 4 K, 5 K are used for both cooling and heating to test the prediction capability of the two models for unseen step sizes, where $|\Delta T| = 1$ K serves as the control group and $|\Delta T| = 2$ K, 3 K, 4 K, 5 K constitute the unseen step-size experimental groups. The same step size is used for both cooling and heating, with the number of steps $N = 90/\Delta T$.

Fig. 9 presents the complete cycle prediction results of the three groups of models under the five step sizes (first row: temperature–stress; second row: temperature–strain; third row: strain–stress). Fig. 10 gives the global relative errors of stress during the cooling stage and strain during the heating stage. The experimental results indicate that, at the training step size, the prediction accuracies of the decoupled and coupled models are comparable, with relative stress errors of 0.834% and 0.445%, respectively, with the coupled model being slightly better, indicating that decoupling itself does not introduce additional accuracy loss. However, as the step size increases to unseen ranges, the differences between the two become rapidly apparent. The relative stress error of the decoupled model remains nearly constant with increasing step size, reaching only 1.191% at $|\Delta T| = 5$ K. In contrast, the relative stress error of the coupled model grows relatively quickly, reaching 3.981%, 8.037%, 11.240%, and 14.927% at $|\Delta T| = 2$ K,3 K,4 K,5 K, respectively; at $|\Delta T| = 5$ K, the stress error of the coupled model is 12.5 times that of the decoupled model. Although the relative strain errors during the heating stage remain at relatively low levels for both models, the error of the coupled model still increases monotonically with step size (from 0.861% to 1.214%), while the decoupled model remains approximately stable (from 0.930% to 1.004%).

The comparison demonstrates that the decoupled structure enhances the generalization capability of LDEONet. By decoupling the evolution direction from the evolution intensity, the network only needs to learn the state-dependent mapping D, which is independent of step size, thereby enabling the model to remain stable under unseen step sizes. In contrast, the coupled model incorporates $d\phi_g/dT$ as a network input, which expands the sample demand of the network; when the input distribution deviates from the training distribution, extrapolation failure occurs, increasing prediction errors. This result confirms from an ablation perspective that the generalization capability of LDEONet does not arise from network capacity or training techniques, but rather from its architecture design, which is consistent with the local evolution laws of internal variables.

4.4. Model comparison

Under the premise that the internal variables are already defined and accessible, this paper directly learns the local mapping from the current state to the next state based on the Markovian property, thereby

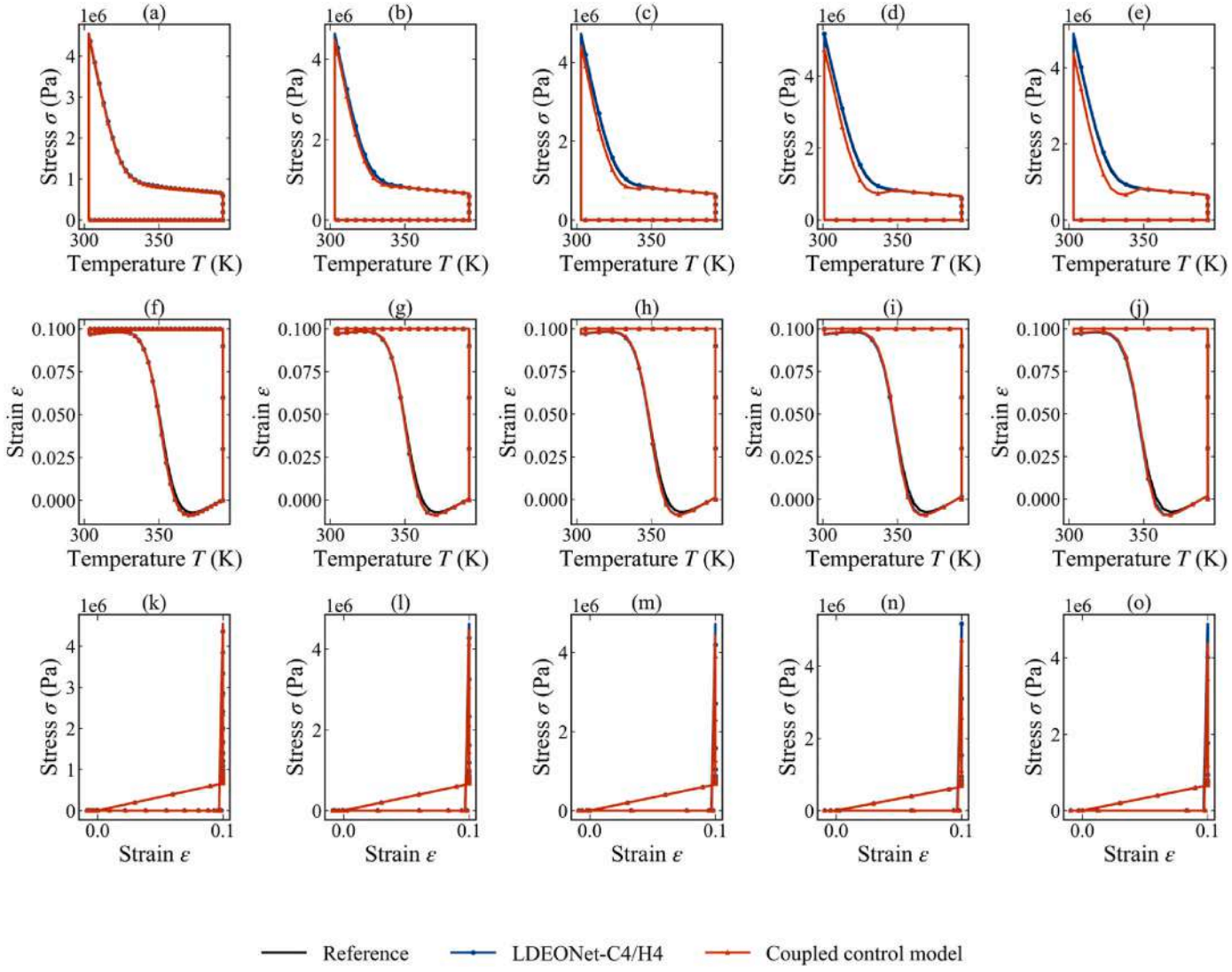


Fig. 9. Decoupling ablation experiments: (a)-(e) temperature-stress results; (f)-(j) temperature-strain results; (k)-(o) strain-stress results.

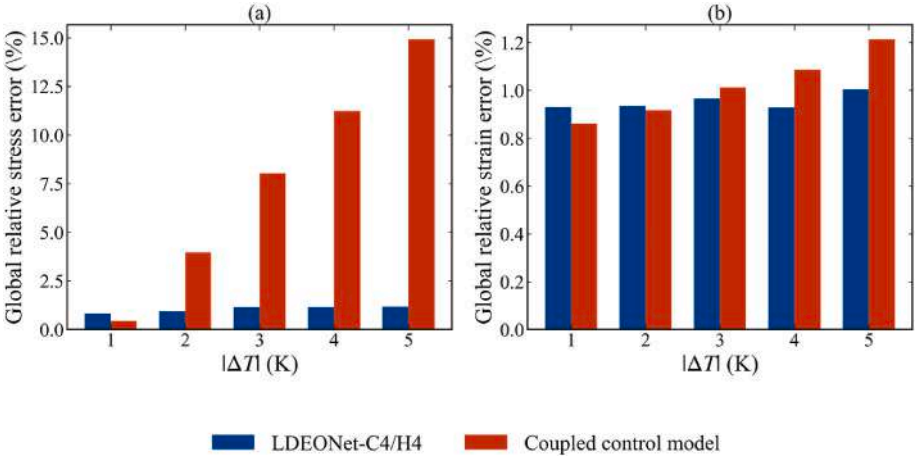


Fig. 10. Decoupling ablation experiment relative errors: (a) global relative stress error; (b) global relative strain error.

achieving path generalization and discretization generalization capabilities. However, the Markovian property does not imply that relevant sequence models are redundant. When an explicit internal variable evolution model is unavailable, sequence models such as LSTM can still learn the evolution laws of history-dependent materials through latent variable modeling.

To systematically evaluate the performance of the LDEONet architecture in practical cycle prediction, this section uses the complete

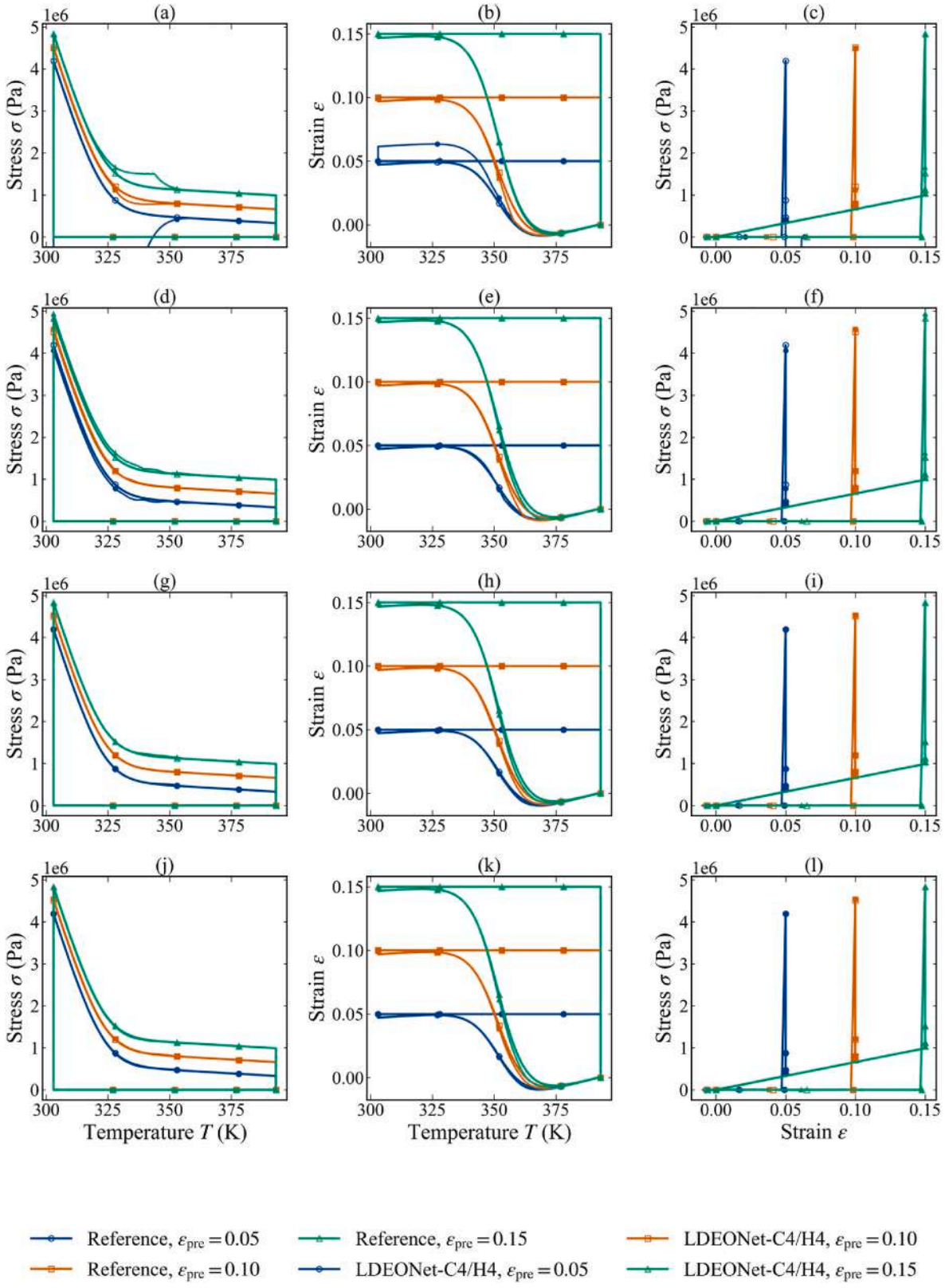


Fig. 11. LDEONet cycle prediction experiments: (a)-(c) LDEONet-C1/H1, (d)-(f) LDEONet-C2/H2, (g)-(i) LDEONet-C3/H3, (j)-(l) LDEONet-C4/H4, showing temperature-stress, temperature-strain, and strain-stress results for each model (three columns per model), respectively.

thermomechanical cycle as the experimental vehicle and compares the cycle prediction results of four groups of decoupled models (LDEONet-C1/H1, LDEONet-C2/H2, LDEONet-C3/H3, LDEONet-C4/H4), providing a comprehensive assessment of LDEONet's predictive capability and offering data for comparison with other similar studies. The experimental scheme is as follows:

(1). Pre-strains $\varepsilon_{\text{pre}} = 0.05, 0.10, 0.15$;
(2). Predictions are performed using the mode-switching strategy of LDEONet-C1/H1, LDEONet-C2/H2, LDEONet-C3/H3, and LDEONet-C4/H4;
(3). The experimental content consists of a conventional thermomechanical cycle (high-temperature loading at 393 K → cooling fixation to 303 K → isothermal unloading at 303 K → heating recovery to 393 K), with the reference data being the numerical solutions obtained by the finite difference method;
(4). In the prediction step sizes, both cooling and heating adopt a constant step size of $|\Delta T| = 1$ K, with the number of steps $N = 90$.

Fig. 11 presents the thermomechanical cycle prediction results of the four groups of models, including temperature-stress, temperature-strain, and strain-stress data in the first, second, and third columns, respectively. The results show that as the hidden-layer scale of the models increases, the agreement between predictions and the reference improves progressively. LDEONet-C1/H1 (5 neurons, 21 parameters) exhibits severe divergence under the $\varepsilon_{\text{pre}} = 0.05$ condition, with a relative stress error as high as 491.7%, indicating that the model capacity is insufficient to learn the complete evolution direction mapping, and the error accumulates and amplifies along the path during incremental integration. LDEONet-C2/H2 (10 neurons, 41 parameters) still shows noticeable deviations under $\varepsilon_{\text{pre}} = 0.05$, with a relative stress error of 5.06%, but has already converged to 1.26% under the $\varepsilon_{\text{pre}} = 0.10$ condition. LDEONet-C3/H3 (20 neurons, 81 parameters) and LDEONet-C4/H4 (30 neurons, 121 parameters) achieve high consistency with the reference data across all conditions, with relative stress errors not exceeding 0.50% and 0.81%, respectively, and relative strain errors all below 1.0%.

Fig. 12 summarizes the global relative errors of the four groups of models for the three pre-strain cases. The error bar charts of LDEONet-C3/H3 and LDEONet-C4/H4 remain at the same low level, indicating that the decoupled models with more than 20 neurons already possess sufficient capacity, and further parameter increases yield only marginal improvements. In addition, the closure of the complete cycle is satisfactory: taking LDEONet-C4/H4 as an example, under the $\varepsilon_{\text{pre}} = 0.10$ cycle, the strain at the end of heating recovery is 4×10^{-4}, compared with the reference value of 4.6×10^{-4}, giving a relative residual strain error of about 0.06%. The residual strain error after unloading does not exceed 1.4×10^{-4} in all cases, indicating that the model can maintain consistency in energy and state throughout the entire cycle.

It should be noted that research on neural network constitutive models for SMPs is still relatively limited, and the results presented in this table are derived from different studies, which employ distinct material parameters, datasets, loading conditions, training methodologies, and evaluation metrics. Therefore, this table is intended solely to describe the basic characteristics of different modeling approaches and does not constitute a controlled baseline comparison under identical data and training conditions. Even though the training conditions vary, the training curves still represent a set of reasonably independent data, describing the thermomechanical cycle conditions included in the sample database for each neural network model. Consequently, the training curve serves as the unit of horizontal comparison in this section.

Table 2 summarizes the model scales, training data volumes, and relevant predictive metrics of LSTM, PINN, and LDEONet reported in the existing literature. The figure "2.7 curves" is taken from the original reference (Yan et al., 2025), indicating the data partitioning ratio in the study by Yan et al., rather than representing the actual collection of 2.7 independent physical curves. In modeling SMP materials, the number of training curves used by LDEONet in this work is approximately 1/175 of that used by LSTM, and the parameter scale is compressed to the order of 242. Meanwhile, LDEONet achieves path generalization and discretization generalization capabilities with a small-sample cost, and the decoupling design ensures generalization accuracy—capabilities that have not been reported in related works.

5. Conclusion

This paper analyzes and discusses the generalization difficulties of neural network constitutive models for history-dependent materials.

Table 2
Neural network constitutive models for shape memory polymers (the MSE of LDEONet is taken as the larger value between LDEONet-C4 and LDEONet-H4).

Metric	LSTM	PINN	LDEONet
Parameter scale	1.33×10^5	273	242
Training sample size	3680 *curves*	2.7 *curves*	20 *curves*
Convergence accuracy (*MSE*)	1.01×10^{-7}	2.3×10^{-4}	3.49×10^{-3}
Optimization iterations	500	1.2×10^7	1×10^6
Path generalization	Not reported	Not reported	Reported
Discretization generalization	Not reported	Not reported	Reported
Cycle description capability	Reported	Not reported	Reported
Ref.	Deng and Lu (2026)	Yan et al. (2025)	(This Work)

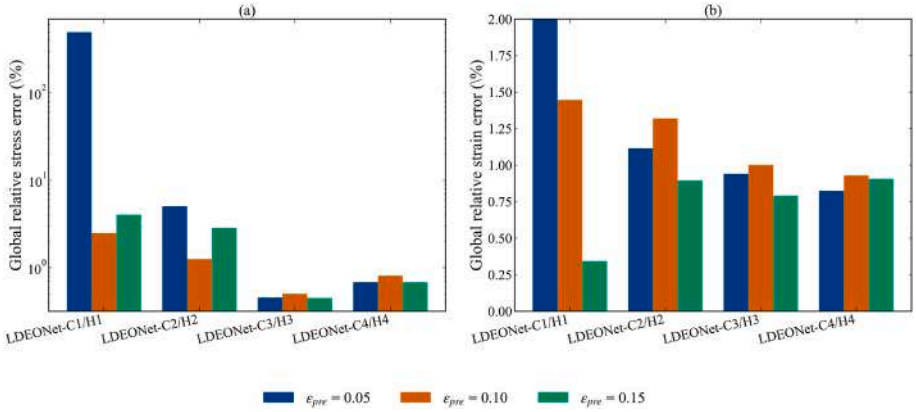


Fig. 12. Relative errors of LDEONet cycle prediction: (a) global relative stress error; (b) global relative strain error.

Inspired by the Markovian property of the internal variable evolution differential equations of shape memory polymers, a Local Decoupling Evolution Operator Network is proposed. The generalization performance and the advantages of decoupling are examined through designed path generalization experiments, discretization generalization experiments, and decoupling ablation studies. The main conclusions are as follows:

(a) Revelation of the Markovian property of the evolution process.

The internal variable evolution differential equations of shape memory polymers exhibit the Markovian property. When finite-dimensional internal variables are defined, accessible, and capable of forming a closed state description, a local update network that depends only on the current state can be used, without the need to feed explicit historical sequences into the model.

(b) Proposal of the LDEONet architecture based on localization and decoupling.

Based on the internal variable evolution differential equations of shape memory polymers, this paper proposes the Local Decoupling Evolution Operator Network, which decouples the complex material evolution law into evolution direction and evolution intensity, thereby achieving in principle generalization to unseen paths and unseen prediction step sizes. The demand for thermomechanical cycle samples is approximately 1/175 of that of LSTM, and the number of trainable parameters is reduced to 242.

(c) Design of generalization experiments and validation of model generalization and decoupling advantages.

This paper designs generalization experiments for neural network models of history-dependent materials, including path generalization tests and discretization generalization tests. The path generalization experiments demonstrate that LDEONet-C4/H4 possesses predictive capability for various unseen reversal paths. The discretization generalization experiments prove that LDEONet-C4/H4 can predict thermomechanical cycles under unseen random, sinusoidal, and stepwise step sizes. The decoupling ablation study confirms the structural advantages of decoupling, effectively suppressing generalization errors.

In summary, this research not only provides a constitutive prediction tool for history-dependent materials that combines a lightweight architecture with high generalization capability, but more importantly, establishes a new paradigm for neural network constitutive modeling of history-dependent materials—namely, leveraging the underlying local evolution mechanisms of materials to build models with stronger generalization. It should be noted that the research subject of this paper is the thermomechanical cycle of one-dimensional shape memory polymers. Future work will focus on extending LDEONet to a three-dimensional finite-deformation framework and deepening its integration with finite element solution frameworks, thereby advancing the application of neural network constitutive models in high-performance nonlinear numerical simulations.

CRediT authorship contribution statement

Chen Wang: Writing – original draft, Validation, Software, Methodology, Investigation, Formal analysis, Data curation, Conceptualization. **Fengfeng Li:** Writing – review & editing, Validation, Supervision, Methodology. **Dou Zhang:** Validation, Formal analysis. **Hanxing Zhao:** Resources, Investigation. **Xianyi Li:** Investigation, Data curation. **Yanju Liu:** Writing – review & editing, Supervision, Resources, Project administration, Funding acquisition. **Jinsong Leng:** Writing – review & editing, Supervision, Project administration, Funding acquisition, Conceptualization.

Declaration of competing interest

The authors declare the following financial interests/personal relationships which may be considered as potential competing interests: Fengfeng Li reports financial support was provided by National Natural Science Foundation of China. [grant numbers 12102107]Fengfeng Li reports financial support was provided by China Postdoctoral Science Foundation. (No. 2022M710967)Fengfeng Li reports financial support was provided by China National Postdoctoral Program for Innovative Talents. [grant number BX2021090]Fengfeng Li reports financial support was provided by Postdoctoral Scientific Research Development Fund of Heilongjiang Province. (No. LBH-Z21155)Fengfeng Li reports financial support was provided by the Fundamental research fund of Laboratory for Space Environment and Physical Sciences of the Harbin Institute of Technology, the Open research fund of Suzhou Laboratory. (No.SZLAB-1508-2024-ZD016)Fengfeng Li reports financial support was provided by the Open Project of National Key Laboratory of Aerospace Mechanism. (No.2024ASM-KF03)Dou Zhang reports financial support was provided by National Natural Science Foundation of China. [grant numbers 12402403]If there are other authors, they declare that they have no known competing financial interests or personal relationships that could have appeared to influence the work reported in this paper.

Acknowledgements

This work was supported by the China Postdoctoral Science Foundation (No. 2022M710967), the China National Postdoctoral Program for Innovative Talents (No. BX2021090), the National Natural Science Foundation of China (No. 12102107, No. 12402403), Heilongjiang Postdoctoral Fund (No. LBH-Z21155), the Fundamental research fund of Laboratory for Space Environment and Physical Sciences of the Harbin Institute of Technology, the Open research fund of Suzhou Laboratory (No. SZLAB-1508-2024-ZD016), the Open Project of National Key Laboratory of Aerospace Mechanism (No. 2024ASM-KF03).

Data availability

Data will be made available on request.

References

Abueidda, D.W., Koric, S., Sobh, N.A., Sehitoglu, H., 2021. Deep learning for plasticity and thermo-viscoplasticity. Int. J. Plast. 136, 102852. https://doi.org/10.1016/j.ijplas.2020.102852.

Anumasa, S., Gunapati, G., Srijith, P.K., 2023. Continuous depth recurrent neural differential equations. In: Koutra, D., Plant, C., Gomez Rodriguez, M., Baralis, E., Bonchi, F. (Eds.), Machine Learning and Knowledge Discovery in Databases: Research Track. Springer Nature, Switzerland, Cham, pp. 223–238.

Askari-Sedeh, et al., 2025. Polymers 17 (17), 2329. https://doi.org/10.3390/polym17172329.

Chandaka, S.K., Das, A., Laskar, P., 2025. Emergence of shape memory polymers as a new material for diverse applications. RSC Adv. 15, 31210–31229.

Chen, Y.-C., Lagoudas, D.C., 2008. A constitutive theory for shape memory polymers. Part I. J. Mech. Phys. Solid. 56, 1752–1765. https://doi.org/10.1016/j.jmps.2007.12.005.

Collins-Craft, N.A., Stefanou, I., Sulem, J., Einav, I., 2020. A Cosserat breakage mechanics model for brittle granular media. J. Mech. Phys. Solid. 141, 103975. https://doi.org/10.1016/j.jmps.2020.103975.

Cooper, C.B., Nikzad, S., Yan, H., Ochiai, Y., Lai, J.-C., Yu, Z., Chen, G., Kang, J., Bao, Z., 2021. High energy density shape memory polymers using strain-induced supramolecular nanostructures. ACS Cent. Sci. 7, 1657–1667. https://doi.org/10.1021/acscentsci.1c00829.

Deng, J., Lu, H., 2026. Deep learning of long short-term memory neural networks in shape memory polymers towards shape memory behaviors. Polymer 346, 129651. https://doi.org/10.1016/j.polymer.2026.129651.

Dornheim, J., Morand, L., Nallani, H.J., Helm, D., 2024. Neural networks for constitutive modeling: from universal function approximators to advanced models and the integration of physics. Arch. Comput. Methods Eng. 31, 1097–1127. https://doi.org/10.1007/s11831-023-10009-y.

Dutta, S., Havanagi, S., Mishra, P., Sharma, D., Singh, B.N., 2025. From inert to intelligent: smart biomaterials reshaping the landscape of bioengineering. Int. J. Soc. Netw. Min. 16, 845–877. https://doi.org/10.1080/19475411.2025.2570535.

Faraji, 2025. Appl. Math. Mech. (English Edition) 46 (9), 1787. https://doi.org/10.1007/s10483-025-3290-9.
Fetisova, A.A., Surmeneva, M.A., Surmenev, R.A., 2026. Advanced design concepts for shape-memory polymers in biomedical applications and soft robotics. Polymers 18, 214.
Feyel, F., 2003. A multilevel finite element method (FE2) to describe the response of highly non-linear structures using generalized continua. Comput. Methods Appl. Mech. Eng. 192, 3233–3244. https://doi.org/10.1016/S0045-7825(03)00348-7.
Freitag, S., Graf, W., Kaliske, M., 2011. Recurrent neural networks for fuzzy data. Integrated Comput. Aided Eng. 18, 265–280. https://doi.org/10.3233/ICA-2011-0373.
Fuhg, J.N., Anantha Padmanabha, G., Bouklas, N., Bahmani, B., Sun, W., Vlassis, N.N., Flaschel, M., Carrara, P., De Lorenzis, L., 2025. A review on data-driven constitutive laws for solids. Arch. Comput. Methods Eng. 32, 1841–1883. https://doi.org/10.1007/s11831-024-10196-2.
Gorji, M.B., Mozaffar, M., Heidenreich, J.N., Cao, J., Mohr, D., 2020. On the potential of recurrent neural networks for modeling path dependent plasticity. J. Mech. Phys. Solid. 143, 103972. https://doi.org/10.1016/j.jmps.2020.103972.
Graf, W., Freitag, S., Kaliske, M., Sickert, J.-U., 2010. Recurrent neural networks for uncertain time-dependent structural behavior. Comput. Aided Civ. Infrastruct. Eng. 25, 322–323. https://doi.org/10.1111/j.1467-8667.2009.00645.x.
Graf, W., Freitag, S., Sickert, J., Kaliske, M., 2012. Structural analysis with fuzzy data and neural network based material description. Comput. Aided Civ. Infrastruct. Eng. 27, 640–654. https://doi.org/10.1111/j.1467-8667.2012.00779.x.
Graves, A., Jaitly, N., Mohamed, A. -r., 2013a. Hybrid speech recognition with deep bidirectional LSTM. In: 2013 IEEE Workshop on Automatic Speech Recognition and Understanding. IEEE, Olomouc, Czech Republic, pp. 273–278. https://doi.org/10.1109/ASRU.2013.6707742.
Graves, A., Mohamed, A. -r., Hinton, G., 2013b. Speech recognition with deep recurrent neural networks. In: 2013 IEEE International Conference on Acoustics, Speech and Signal Processing. IEEE, Vancouver, BC, Canada, pp. 6645–6649. https://doi.org/10.1109/ICASSP.2013.6638947.
Graves, A., Schmidhuber, J., 2005. Framewise phoneme classification with bidirectional LSTM and other neural network architectures. Neural Netw. 18, 602–610. https://doi.org/10.1016/j.neunet.2005.06.042.
Guo, B., Lin, Z., He, Q., 2025. History-aware neural operator: robust data-driven constitutive modeling of path-dependent materials. Comput. Methods Appl. Mech. Eng. 447, 118358. https://doi.org/10.1016/j.cma.2025.118358.
He, Y., Semnani, S.J., 2024. Incremental neural controlled differential equations for modeling of path-dependent material behavior. Comput. Methods Appl. Mech. Eng. 422, 116789. https://doi.org/10.1016/j.cma.2024.116789.
Hochreiter, S., Schmidhuber, J., 1997. Long short-term memory. Neural Comput. 9, 1735–1780. https://doi.org/10.1162/neco.1997.9.8.1735.
Huang, J., Wang, M., Gu, J., Wang, C., Ying, S., 2026. Solvent-responsive shape memory porous semicrystalline thermoplastic polyurethane. J. Appl. Polym. Sci. 143, e70825. https://doi.org/10.1002/app.70825.
Jones, R.E., Fuhg, J.N., 2025. An attention-based neural ordinary differential equation framework for modeling inelastic processes. Comput. Methods Appl. Mech. Eng., 118186 https://doi.org/10.1016/j.cma.2025.118186.
Kapturowski, S., Ostrovski, G., Dabney, W., Quan, J., Munos, R., 2019. Recurrent experience replay in distributed reinforcement learning. In: International Conference on Learning Representations.
Klein, D.K., Fernández, M., Martin, R.J., Neff, P., Weeger, O., 2022. Polyconvex anisotropic hyperelasticity with neural networks. J. Mech. Phys. Solid. 159, 104703. https://doi.org/10.1016/j.jmps.2021.104703.
Li, F., 2020. Constitutive Model and Space Application Verification of Epoxy-based Shape Memory Polymer Composite (PhD Thesis). Harbin Institute of Technology, Harbin. https://doi.org/10.27061/d.cnki.ghgdu.2020.004918.
Liu, B., Jin, K., Tao, J., Wang, H., He, D., Li, H., 2022. Performance optimization of shape memory epoxy polymers based on machine learning. Polym. Adv. Technol. 33, 1222–1232. https://doi.org/10.1002/pat.5595.
Liu, Y., Gall, K., Dunn, M.L., Greenberg, A.R., Diani, J., 2006. Thermomechanics of shape memory polymers: uniaxial experiments and constitutive modeling. Int. J. Plast. 22, 279–313. https://doi.org/10.1016/j.ijplas.2005.03.004.
Lu, L., Jin, P., Pang, G., Zhang, Z., Karniadakis, G.E., 2021. Learning nonlinear operators via DeepONet based on the universal approximation theorem of operators. Nat. Mach. Intell. 3, 218–229. https://doi.org/10.1038/s42256-021-00302-5.
Masi, F., Stefanou, I., 2022. Multiscale modeling of inelastic materials with Thermodynamics-based artificial neural networks (TANN). Comput. Methods Appl. Mech. Eng. 398, 115190. https://doi.org/10.1016/j.cma.2022.115190.
Masi, F., Stefanou, I., Maffi-Berthier, V., Vannucci, P., 2020. A discrete element method based-approach for arched masonry structures under blast loads. Eng. Struct. 216, 110721. https://doi.org/10.1016/j.engstruct.2020.110721.
Masi, F., Stefanou, I., Vannucci, P., 2018. A study on the effects of an explosion in the pantheon of rome. Eng. Struct. 164, 259–273. https://doi.org/10.1016/j.engstruct.2018.02.082.
Masi, F., Stefanou, I., Vannucci, P., Maffi-Berthier, V., 2021. Thermodynamics-based artificial neural networks for constitutive modeling. J. Mech. Phys. Solid. 147, 104277. https://doi.org/10.1016/j.jmps.2020.104277.
Mirasadi, K., Yousefi, M.A., Jin, L., Rahmatabadi, D., Baniassadi, M., Liao, W.-H., Bodaghi, M., Baghani, M., 2026. 4D printing of magnetically responsive shape memory polymers: toward sustainable solutions in soft robotics, wearables, and biomedical devices. Adv. Sci. 13, e13091. https://doi.org/10.1002/advs.202513091.
Oeser, M., Freitag, S., 2009. Modeling of materials with fading memory using neural networks. Int. J. Numer. Methods Eng. 78, 843–862. https://doi.org/10.1002/nme.2518.
Pan, W., Wan, L., Duan, H., Gu, J., Sun, H., 2026. Modelling the thermo-mechanical responses and shape recovery performance of particle-reinforced shape memory polymer composites in cold and hot programming. Polymer 360, 130370. https://doi.org/10.1016/j.polymer.2026.130370.
Park, H., Harrison, P., Guo, Z., Lee, M.-G., Yu, W.-R., 2016. Three-dimensional constitutive model for shape memory polymers using multiplicative decomposition of the deformation gradient and shape memory strains. Mech. Mater. 93, 43–62. https://doi.org/10.1016/j.mechmat.2015.10.014.
Schuster, M., Paliwal, K., 1997. Bidirectional recurrent neural networks. IEEE Trans. Signal Process. 45, 2673–2681. https://doi.org/10.1109/78.650093.
Sun, P., Wu, J., Zhang, M., Devos, P., Botteldooren, D., 2025. Delayed memory unit: modeling temporal dependency through delay gate. IEEE Transact. Neural Networks Learn. Syst. 36, 10808–10818.
Tac, V., Sahli Costabal, F., Tepole, A.B., 2022. Data-driven tissue mechanics with Polyconvex neural ordinary differential equations. Comput. Methods Appl. Mech. Eng. 398, 115248. https://doi.org/10.1016/j.cma.2022.115248.
Teng, Y., Shi, Z., Hou, S., Lv, Q., Shi, J., Fan, H., Ramakrishna, S., 2025. Intelligent/smart materials for wearables used in fire emergency medical services. Int. J. Soc. Netw. Min. 16, 726–741. https://doi.org/10.1080/19475411.2025.2559388.
Thakolkaran, P., Joshi, A., Zheng, Y., Flaschel, M., De Lorenzis, L., Kumar, S., 2022. NN-EUCLID: deep-learning hyperelasticity without stress data. J. Mech. Phys. Solid. 169, 105076. https://doi.org/10.1016/j.jmps.2022.105076.
Tobushi, H., Okumura, K., Hayashi, S., Ito, N., 2001. Thermomechanical constitutive model of shape memory polymer. Mech. Mater. 33, 545–554. https://doi.org/10.1016/S0167-6636(01)00075-8.
Volk, B.L., Lagoudas, D.C., Chen, Y.-C., 2010. Analysis of the finite deformation response of shape memory polymers: II. 1D calibration and numerical implementation of a finite deformation, thermoelastic model. Smart Mater. Struct. 19, 075006. https://doi.org/10.1088/0964-1726/19/7/075006.
Volk, B.L., Lagoudas, D.C., Maitland, D.J., 2011. Characterizing and modeling the free recovery and constrained recovery behavior of a polyurethane shape memory polymer. Smart Mater. Struct. 20, 094004. https://doi.org/10.1088/0964-1726/20/9/094004.
Wang, C., Dai, Y., Kou, B., Huang, W.M., 2019. Influence of long-term storage on shape memory performance and mechanical behavior of pre-stretched commercial poly (methyl methacrylate). Polymers 11, 1978. https://doi.org/10.3390/polym11121978.
Wang, C., Pek, J.X., Chen, H.M., Huang, W.M., 2022. On-demand tailoring between brittle and ductile of poly(methyl methacrylate) via high temperature stretching. Polymers 14, 985. https://doi.org/10.3390/polym14050985.
Wu, X., Wang, C., Pek, J.X., Huang, W.M., 2022. Brittle-to-ductile transition in high temperature pre-stretched polycarbonate. J. Polym. Res. 29, 435. https://doi.org/10.1007/s10965-022-03282-1.
Xu, Z., Meng, S., Wei, D.-W., Bao, R.-Y., Wang, Y., Ke, K., Yang, W., 2023. Hierarchical network relaxation of a dynamic cross-linked polyolefin elastomer for advanced reversible shape memory effect. Nanoscale 15, 5458–5468. https://doi.org/10.1039/D2NR06902D.
Yan, C., Feng, X., Mensah, P., Li, G., 2025. Toward a general physics-informed neural network for amorphous shape memory polymer modelling. Proc. R. Soc. A 481, 20240702. https://doi.org/10.1098/rspa.2024.0702.
Yousefi, M.A., Rahmatabadi, D., Ahmadi, M., Bayati, A., Baniassadi, M., Laachachi, A., Silberschmidt, V.V., George, D., Baghani, M., 2026. 4D printing for minimally invasive biomedical applications: programmable smart materials for deployable devices, drug delivery, and tissue regeneration. Mater. Des. 263, 115555. https://doi.org/10.1016/j.matdes.2026.115555.
Zhang, P., Yin, Z., Jin, Y., Ye, G., 2020. An AI-based model for describing cyclic characteristics of granular materials. Int. J. Numer. Anal. Methods GeoMech. 44, 1315–1335. https://doi.org/10.1002/nag.3063.
Zhao, W., Liu, L., Leng, J., Liu, Y., 2020. Thermo-mechanical behavior prediction of shape memory polymer based on the multiplicative decomposition of the deformation gradient. Mech. Mater. 143, 103263. https://doi.org/10.1016/j.mechmat.2019.103263.
Zhao, X., Chang, C., Zhu, W., Ablali, A., Li, M., Xie, F., Li, Y., 2025. Energy harvesting dynamic analysis of a piezoelectric beam under axially moving and spinning motions based on the Green's functions. Int. J. Soc. Netw. Min. 16, 878–907. https://doi.org/10.1080/19475411.2025.2574076.