



The mechanics and physics of tofu: Understanding hydrated soft solids through feature networks

Birte Boes ^{a,*}, Jaan-Willem Simon ^a, Hagen Holthausen ^b, Ellen Kuhl ^{b,c}

^a Computational Applied Mechanics, University of Wuppertal, Pauluskirchstraße 7, 42285, Wuppertal, Germany

^b Institute of Applied Mechanics, University of Erlangen-Nuremberg, Egerlandstraße 5, 91058, Erlangen, Germany

^c Department of Mechanical Engineering, Stanford University, United States

ARTICLE INFO

Keywords:

Inelasticity
Viscoelasticity
Poroelectricity
Automated model discovery
Feature network
Food
Tofu

ABSTRACT

Tofu remains one of the world's most important plant-based foods, famous for its cultural legacy, nutritional benefits, and environmental sustainability. Made up of only two ingredient—soy beans and water—it has nourished people for centuries; yet, its rheology remains incompletely understood. Here, we use tofu as a minimal model system to study how composition governs the mechanics and physics of hydrated soft solids. We perform over a hundred controlled compression tests on silken, firm, and extrafirm tofu, and observe a strong nonlinearity, a pronounced viscoelasticity, and a more than ten-fold increase in stress with only six percent decrease in water content. We demonstrate that automated model discovery can identify inelastic constitutive laws that accurately capture these unique characteristics across multiple loading magnitudes and rates. The discovered models autonomously separate elastic and inelastic mechanisms and reveal that tofu elasticity depends primarily on the second isochoric invariant, while tofu inelasticity depends on a combination of the second and third deviatoric invariants. This functional form is universal across all three tofu types, but scales in magnitude with water content. Our new water-content-based feature network reveals that this correlation is not linear, as assumed by traditional bi-phasic theories, but rather highly non-linear and sensitive to the individual model terms. These results position tofu as a quantitative benchmark for nonlinear poroviscoelastic solids and demonstrate how physics-informed machine learning can uncover the constitutive structure of hydrated soft materials. Our source code and examples are available at <https://doi.org/10.5281/zenodo.16993236>.

“Where soy meets water, tofu becomes matter.”

1. Introduction

Tofu feeds more than a billion people worldwide and has done so for over two thousand years (Shurtleff and Aoyagi, 1975). It is one of the simplest and most influential plant-based foods in human history (Watanabe, 1997). Beyond its cultural role, tofu promotes human health (Qin et al., 2022), contributes to planetary sustainability (Poore and Nemecek, 2018), and remains the cornerstone of a modern plant-based diet (Rizzo and Baroni, 2018). Yet, the material itself is remarkably simple (Kohyama et al., 1995): Soymilk and a mineral coagulant form a hydrated protein gel with only two essential ingredients, soybeans and water (Zheng et al., 2020). Whether we make our own tofu at home or buy it in supermarkets, Fig. 1 illustrates that its texture can range

* Corresponding author.

E-mail addresses: birte.boes@uni-wuppertal.de (B. Boes), jaan.simon@uni-wuppertal.de (J.-W. Simon), hagen.holthausen@fau.de (H. Holthausen), ekuhl@stanford.edu (E. Kuhl).

<https://doi.org/10.1016/j.jmps.2026.106632>

Received 10 December 2025; Received in revised form 9 April 2026; Accepted 11 April 2026

Available online 15 April 2026

0022-5096/© 2026 The Author(s). Published by Elsevier Ltd. This is an open access article under the CC BY license (<http://creativecommons.org/licenses/by/4.0/>).

from silken to extrafirm, determined almost entirely by its water content (Chen et al., 2023) and by the specifics of its coagulation process (Wang et al., 2018). This simplicity makes tofu an ideal prototype material for testing, validating, and revealing the power of automated model-discovery for soft matter systems.

For more than half a century, texture profile analysis has defined the gold standard in food science to quantify the textural and sensory attributes of food (Friedman et al., 1963). In texture profile analysis, mechanical double-compression tests simulate the behavior of chewing (Bourne, 1968), we can easily extract its parameters such as stiffness, hardness, cohesiveness, springiness, resilience, and chewiness from the resulting force–time curves (Dunne et al., 2025). Although texture profile analysis continues to be widely used (Paredes et al., 2022), the lack of standardized testing protocols limits comparability across the different studies and labs (Schreuders et al., 2021). Moreover, the interpretation of texture profile analysis curves remains largely rooted in food science rather than in materials science, even though these tests inherently capture valuable mechanical information.

Mechanical testing allows us to interpret texture through the lens of mechanics and to extract fundamental material properties directly from experimental data (St Pierre et al., 2024). We can quantify material parameters such as stiffness, viscosity, and relaxation time—commonly used in mechanics—from tension, compression, and shear tests to reveal the underlying macrostructural behavior and correlate it to the microstructural architecture (Budday et al., 2020). Grounding texture evaluation in the principles of material modeling (Vervenne et al., 2025a) enables a more rigorous and interpretable mechanistic analysis (Tac et al., 2026). This raises the question: Can we characterize the behavior of food through robust constitutive models that reliably reproduce experimental data and confidently predict the effects of different loading conditions?

Food materials form an emerging class of complex, heterogeneous systems that challenge conventional modeling techniques (Datta et al., 2025). Tofu, as a hydrated soy protein gel, contains water in three distinct states—tightly bound, immobilized, and freely mobile—each interacting differently with the protein network (Li et al., 2014). While bound and immobilized water contribute to network stability and elastic stiffness, freely mobile water governs fluid redistribution, rate dependence, and syneresis and gives tofu a fundamentally biphasic, poroelastic character (Lucey, 2002). This nonlinear, often rate-dependent response calls for new approaches that are capable of identifying the governing equations directly from experimental data (Kuhl, 2025). Recent advances in data-driven mechanics have introduced frameworks that automatically discover constitutive relationships (Linka and Kuhl, 2023) without pre-selecting the specific functional form of the model (Holthusen et al., 2024b). Instead of repeatedly fitting and refining candidate models, automated model discovery infers the governing equations in a self-consistent and physics-informed manner (Flaschel et al., 2021). This motivates us to ask: Can we adopt the concept of automated model discovery to discover the models and parameters that best characterize tofu as a minimal benchmark system for soft, structurally engineered foods?

A wide range of data-driven frameworks has emerged in recent years, including unsupervised learning (Flaschel et al., 2021) as well as mechanics-informed (As'ad et al., 2022), physics-constrained (Kalina et al., 2023), physics-augmented (Kalina et al., 2024; Klein et al., 2023), thermodynamics-based (Masi et al., 2021), input convex neural networks (Amos et al., 2017), and input monotonic neural networks (Holthusen and Kuhl, 2026). Among these, the framework of constitutive artificial neural networks has proven particularly effective because it hardwires continuum mechanical principles directly into the network architecture (Linka et al., 2021). This design ensures that the resulting models remain both physically sound and interpretable (Linka and Kuhl, 2023). The sparse network structure balances accuracy and interpretability and allows the simultaneous discovery of concise yet descriptive constitutive models and their model parameters.

Automated model discovery has successfully identified constitutive models across diverse applications, including soft biological tissues such as skin (Linka et al., 2023b), arteries (Vervenne et al., 2025a), the heart (Martonová et al., 2024), and the brain (Hou et al., 2024; Linka et al., 2023a), and, most recently, plant-based and animal meats (St Pierre et al., 2023). Most studies have focused on discovering models for the elastic behavior; yet, an accurate representation of realistic material responses—especially at larger loading magnitudes and rates—often requires the inclusion of inelastic effects. Recent extensions of the framework have incorporated viscoelasticity through Prony-series representations (Abdolazizi et al., 2024; Wang et al., 2023). Building on this idea, inelastic constitutive artificial neural networks, extend the formulation by discovering both, the Helmholtz free energy for the elastic behavior and a dual potential for the inelastic behavior (Holthusen et al., 2024b). This approach preserves thermodynamic consistency and interpretability, while expanding the range of representable material responses. The general framework of inelastic constitutive neural networks has since been applied to viscoelasticity (Holthusen et al., 2024a), growth modeling (Holthusen et al., 2025a), and plasticity (Boes et al., 2024; van der Velden et al., 2025). Importantly, inelastic constitutive neural networks autonomously discover whether inelastic effects are necessary, and reduce unnecessary model complexity when the data do not justify it (Holthusen et al., 2025b).

Motivated by the growing interest in meat alternatives, this study investigates the mechanics and physics of tofu by replacing the traditional texture profile analysis with classical mechanical testing. We perform compression experiments on multiple tofu types with varying water content and analyze the resulting force–time curves to explore the correlation between structure and function. First, we adopt an inelastic constitutive neural network to discover individual constitutive models for each tofu type and identify the dominant mechanisms that govern the material behavior. Then, we postulate that the water content is the single distinguishing feature across all individual models. We thus design a second neural network that takes the water content as input and outputs a feature vector that modulates the inelastic constitutive neural network. This architecture enables a flexible and unbiased mapping between composition and mechanical behavior. It results in a single unified model that provides predictive insight into how changes in formulation—such as water or protein content—affect the texture of tofu, or more generally, of any alternative protein product.

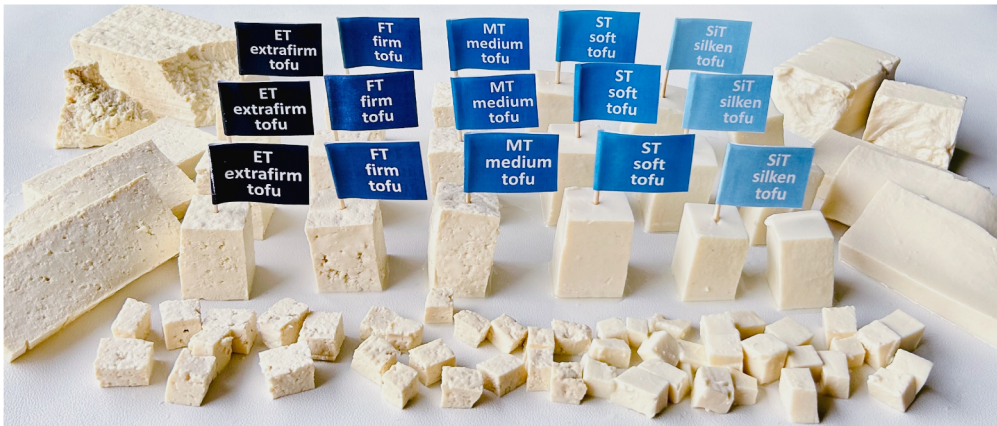


Fig. 1. Five different types of tofu. Extra-firm, firm, medium, soft, and silken tofu illustrate how small changes in composition and processing create large differences in mechanical behavior. Tofu combines soy beans with water to form a biphasic protein gel whose stiffness, porosity, and fracture characteristics arise directly from its water content and its coagulation pathway. Whether we make our own tofu at home or obtain it through commercial retailer, we encounter the same continuum of textures, from a dense cohesive protein matrix to a nearly fluid-like gel.

Table 1

Nutrition facts of silken, firm, and extrafirm tofu. Serving size, calories per serving, macronutrient composition per serving, and fluid mass and volume fractions for the three different tofu types.

tofu type	serving size	calories per serving	composition per serving					fluid mass fraction	fluid volume fraction
			fat	carbs	protein	ash	water		
silken	85g	40cal	2g	2g	4g	1g	76g	0.89	0.92
firm	85g	70cal	4g	2g	7g	1g	71g	0.84	0.87
extrafirm	85g	80cal	4g	2g	8g	1g	70g	0.82	0.86

Section 2 describes the mechanical testing. Sections 3 and 4 introduce the physics-based model and the neural network model. Section 5 summarizes our results and Section 6 discusses our main findings and outlines future directions.

2. Mechanical testing

We test a total of $n = 108$ samples with three different water contents across three different loading magnitudes at five different rates. Specifically, we test three types of tofu, silken, firm, and extrafirm (Azumaya, Nasoya Foods, Ayer, Massachusetts). Table 2 summarizes their serving sizes, calories, and macronutrient compositions per serving according to their individual nutrition fact labels. We estimate the water mass as the difference between the serving size and the listed masses of fat, carbs, protein, and ash.

We assume that tofu fat mainly presents as lipid droplets that do not move freely under mechanical loading, and conclude that the fluid phase of tofu is made up exclusively of the freely moving water. To calculate the mass fraction of the fluid w_f , we divide the water mass, $m_f = \{76g, 71g, 70g\}$, by the total mass per serving of 85g, resulting in $w_f = \{0.89, 0.84, 0.82\}$ for the three tofu types. To calculate the volume fraction of the fluid $v_f = V_f / [V_f + V_s]$, we divide the volume of the fluid, $V_f = w_f / \rho_f$, by the total volume of fluid and solid, $V_f + V_s$, with $V_s = [1 - w_f] / \rho_s$, assuming fluid and solid densities of $\rho_f = 1.0 \text{ g/ml}$ and $\rho_s = 1.3 \text{ g/ml}$, resulting in $v_f = \{0.92, 0.87, 0.86\}$ for the three tofu types. For each tofu type, we perform four sets of tests, three triple-compression tests at different rates, and one double-compression test. For each set, we test at least $n = 8$ samples.

Sample preparation. To prepare our samples, we first cut the tofu into 10 mm thick blocks and then extract cylindrical specimens of 8 mm diameter using biopsy punches. Fig. 2, left, illustrates our sample preparation. To prevent the samples from dehydration, we store them between wetted tissues and test them immediately after preparation.

Sample testing. We perform all mechanical tests using an HR20 Discovery Hybrid Rheometer (TA Instruments, New Castle, Delaware). Fig. 2, right, illustrates a representative sample inside a cylindrical chamber, mounted in the rheometer, to perform compression tests with controlled axial displacements and constrained lateral deformation.

Before testing, we calibrate the rheometer, tare the force sensor, and adjust the initial gap between the loading nod and the specimen to ensure contact without exceeding a force of 0.05 N. All experiments follow a displacement-controlled protocol with four distinct loading programs: three triple-compression tests that serve as training data and one double-compression test that serves as test data. During triple compression in Fig. 2 middle left, we compress each sample three times to a stretch of $\lambda = 0.7$, hold for 0.5 s, unload to $\lambda = 0.9$, and hold for another 0.5 s, at slow, medium, and fast stretch rates of $\dot{\lambda} = \{0.2/s, 0.4/s, 0.6/s\}$. During double compression in Fig. 2 middle right, we compress each specimen first to a stretch of $\lambda = 0.8$ at a rate of $\dot{\lambda} = 0.3/s$, hold for 1.0 s, further compress to $\lambda = 0.7$ at a rate of $\dot{\lambda} = 0.5/s$, and hold for another 1.0 s.

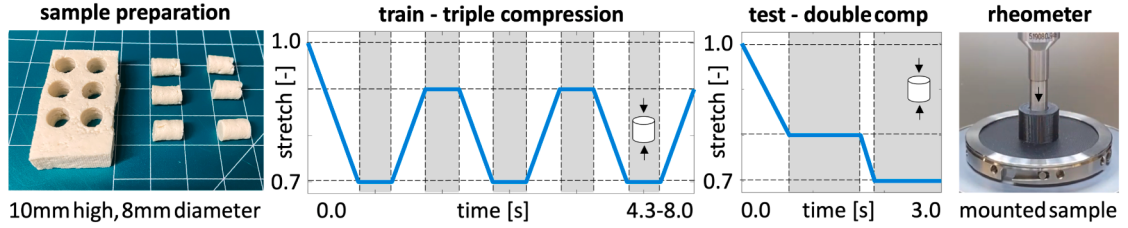


Fig. 2. Mechanical testing. Preparation of cylindrical samples; triple compression test for training; double compression test for testing; and rheometer with mounted sample. During triple compression, we load and unload between $\lambda = 0.7$ and $\lambda = 0.9$ at slow, medium, and fast stretch rates of $\dot{\lambda} = \{0.2/s, 0.4/s, 0.6/s\}$ and hold for 0.5s in between. During double compression, we load to $\lambda = 0.8$ at $\dot{\lambda} = 0.3/s$, hold for 1.0s, load to $\lambda = 0.7$ at $\dot{\lambda} = 0.5/s$, and hold for 1.0s. White regions represent loading and unloading, grey regions represent holding.

Data processing. We record the gap l and the force F throughout the loading history and smoothen both with a Gaussian filter using the `imgaussfilt` function in Matlab R2022b (MATLAB, 2022) with a filter width of $\sigma_{\text{gauss}} = \{8, 3, 3\}$ for silken, firm, and extrafirm tofu. We calculate and report the axial stretch $\lambda = l/L$ as the ratio between the current gap l and the initial gap L and the axial second Piola Kirchhoff stress $S = F/(A\lambda)$ as the recorded force F divided by the initial cross section area $A = \pi \cdot (4 \text{ mm})^2$ and the stretch λ .

3. Physics-based model

Discovering a physically sound constitutive model that explains the experimental data requires a thermodynamically consistent framework. In this section, we outline the continuum equations that motivate the design of our neural network architecture.

Kinematics. To characterize the experimentally observed rate dependence and nonlinear relaxation, we adopt a multiplicative decomposition of the deformation gradient $\mathbf{F}$ into elastic and inelastic parts, $\mathbf{F}_e$ and $\mathbf{F}_i$,

$$\mathbf{F} = \mathbf{F}_e \cdot \mathbf{F}_i. \quad (1)$$

As deformation measures, we introduce the right Cauchy Green deformation tensor $\mathbf{C}$ and its elastic and inelastic counterparts, $\mathbf{C}_e$ and $\mathbf{C}_i$,

$$\mathbf{C} = \mathbf{F}^t \cdot \mathbf{F} \quad \text{and} \quad \mathbf{C}_e = \mathbf{F}_e^t \cdot \mathbf{F}_e = \mathbf{F}_i^{-t} \cdot \mathbf{C} \cdot \mathbf{F}_i^{-1} \quad \text{and} \quad \mathbf{C}_i = \mathbf{F}_i^t \cdot \mathbf{F}_i. \quad (2)$$

To characterize the isotropic volume-preserving and volume-changing behavior of tofu, we introduce the isochoric invariants $\tilde{I}_1^C = I_1^C / (I_3^C)^{1/3}$ and $\tilde{I}_2^C = I_2^C / (I_3^C)^{2/3}$ in terms of the classical invariants $I_1^C = \text{tr}(\mathbf{C})$ and $I_2^C = \text{tr}(\text{cof}(\mathbf{C})) = \frac{1}{2} [(\text{tr}(\mathbf{C}))^2 - \text{tr}(\mathbf{C}^2)]$, and the volumetric invariant $I_3^C = \det(\mathbf{C})$ of the right Cauchy Green tensor $\mathbf{C}$,

$$\tilde{I}_1^C = \text{tr}(\mathbf{C}) / (I_3^C)^{1/3} \quad \text{and} \quad \tilde{I}_2^C = \text{tr}(\text{cof}(\mathbf{C})) / (I_3^C)^{2/3} \quad \text{and} \quad I_3^C = \det(\mathbf{C}), \quad (3)$$

where $\text{tr}(\circ) = (\circ) : \mathbf{I}$ denotes the trace, $\text{cof}(\circ) = I_3^{(\circ)} (\circ)^{-t}$ the cofactor, $\text{tr}(\text{cof}(\circ)) = \frac{1}{2} [(\text{tr}(\mathbf{C}_e))^2 - \text{tr}(\mathbf{C}_e^2)]$ the trace of the cofactor, and $\det(\circ)$ the determinant of the second-order tensor $(\circ)$. The first invariant $\tilde{I}_1^C$ measures the isochoric stretch, the second invariant $\tilde{I}_2^C$ measures the isochoric area change, and the third invariant I_3^C measures the volume change. To characterize the elastic behavior only, we introduce the elastic isochoric invariants $\tilde{I}_1^{C_e} = I_1^{C_e} / (I_3^{C_e})^{1/3}$ and $\tilde{I}_2^{C_e} = I_2^{C_e} / (I_3^{C_e})^{2/3}$ in terms of the classical elastic invariants $I_1^{C_e} = \text{tr}(\mathbf{C}_e)$ and $I_2^{C_e} = \text{tr}(\text{cof}(\mathbf{C}_e))$, and the volumetric invariant $I_3^{C_e} = \det(\mathbf{C}_e)$, of the elastic right Cauchy Green tensor $\mathbf{C}_e$,

$$\tilde{I}_1^{C_e} = \text{tr}(\mathbf{C}_e) / (I_3^{C_e})^{1/3} \quad \text{and} \quad \tilde{I}_2^{C_e} = \text{tr}(\text{cof}(\mathbf{C}_e)) / (I_3^{C_e})^{2/3} \quad \text{and} \quad I_3^{C_e} = \det(\mathbf{C}_e). \quad (4)$$

Elastic potential. We model tofu as an isotropic, inelastic solid and express its total Helmholtz free energy ψ as the sum of an equilibrium contribution ψ^{eq} for the elastic part that depends on the right Cauchy Green tensor $\mathbf{C}$, or, alternatively, on its invariants $\tilde{I}_1^C$, $\tilde{I}_2^C$, I_3^C , and a non-equilibrium contribution ψ^{neq} for the inelastic part that depends on the elastic right Cauchy Green tensor $\mathbf{C}_e$, or on its invariants $\tilde{I}_1^{C_e}$, $\tilde{I}_2^{C_e}$, $I_3^{C_e}$,

$$\psi = \psi^{\text{eq}}(\mathbf{C}) + \psi^{\text{neq}}(\mathbf{C}_e) = \psi^{\text{eq}}(\tilde{I}_1^C, \tilde{I}_2^C, I_3^C) + \psi^{\text{neq}}(\tilde{I}_1^{C_e}, \tilde{I}_2^{C_e}, I_3^{C_e}). \quad (5)$$

To ensure thermodynamic consistency, we evaluate the Clausius–Planck inequality, $D := \mathbf{S} : \frac{1}{2} \dot{\mathbf{C}} - \dot{\psi} \geq 0$, where $\mathbf{S}$ is the second Piola Kirchhoff stress and the dot denotes the material time derivative,

$$D := \left[\mathbf{S} - 2 \frac{\partial \psi^{\text{eq}}}{\partial \mathbf{C}} - 2 \mathbf{F}_i^{-1} \cdot \frac{\partial \psi^{\text{neq}}}{\partial \mathbf{C}_e} \cdot \mathbf{F}_i^{-t} \right] : \frac{1}{2} \dot{\mathbf{C}} + 2 \mathbf{C}_e \cdot \frac{\partial \psi^{\text{neq}}}{\partial \mathbf{C}_e} : \mathbf{D}_i \geq 0, \quad (6)$$

where $\mathbf{D}_i = \text{sym}(\dot{\mathbf{F}}_i \cdot \mathbf{F}_i^{-1})$ is the inelastic rate of deformation tensor. From the Clausius–Planck inequality (6), we obtain the definitions of the second Piola Kirchhoff stress $\mathbf{S} := 2 \partial \psi / \partial \mathbf{C}$,

$$\mathbf{S} = 2 \frac{\partial \psi^{\text{eq}}}{\partial \mathbf{C}} + 2 \mathbf{F}_i^{-1} \cdot \frac{\partial \psi^{\text{neq}}}{\partial \mathbf{C}_e} \cdot \mathbf{F}_i^{-t} = \mathbf{S}^{\text{eq}} + \mathbf{S}^{\text{neq}}, \quad (7)$$

its equilibrium part $\mathbf{S}^{\text{eq}} := 2 \partial \psi^{\text{eq}} / \partial \mathbf{C}$,

$$\mathbf{S}^{\text{eq}} = 2 \frac{\partial \psi^{\text{eq}}}{\partial \tilde{\mathbf{I}}^{\text{C}}} \left[\mathbf{I} - \frac{1}{3} I_1^{\text{C}} \mathbf{C}^{-1} \right] / (I_3^{\text{C}})^{1/3} + 2 \frac{\partial \psi^{\text{eq}}}{\partial \tilde{\mathbf{I}}_2^{\text{C}}} \left[I_1^{\text{C}} \mathbf{I} - \mathbf{C} - \frac{2}{3} I_2^{\text{C}} \mathbf{C}^{-1} \right] / (I_3^{\text{C}})^{2/3} + 2 \frac{\partial \psi^{\text{eq}}}{\partial I_3^{\text{C}}} I_3^{\text{C}} \mathbf{C}^{-1}, \quad (8)$$

and its non-equilibrium part $\mathbf{S}^{\text{neq}} := 2 \mathbf{F}_i^{-1} \cdot \partial \psi^{\text{neq}} / \partial \mathbf{C}_e \cdot \mathbf{F}_i^{-t}$,

$$\mathbf{S}^{\text{neq}} = 2 \frac{\partial \psi^{\text{neq}}}{\partial \tilde{\mathbf{I}}_1^{\text{C}_e}} \left[\mathbf{C}_i^{-1} - \frac{1}{3} I_1^{\text{C}_e} \mathbf{C}_i^{-1} \right] / (I_3^{\text{C}_e})^{1/3} + 2 \frac{\partial \psi^{\text{neq}}}{\partial \tilde{\mathbf{I}}_2^{\text{C}_e}} \left[I_1^{\text{C}_e} \mathbf{C}_i^{-1} - \mathbf{C}_i^{-1} \cdot \mathbf{C} \cdot \mathbf{C}_i^{-1} - \frac{2}{3} I_2^{\text{C}_e} \mathbf{C}_i^{-1} \right] / (I_3^{\text{C}_e})^{2/3} + 2 \frac{\partial \psi^{\text{neq}}}{\partial I_3^{\text{C}_e}} I_3^{\text{C}_e} \mathbf{C}_i^{-1} \quad (9)$$

In Section 4, we reverse-engineer two constitutive neural networks, one for the equilibrium potential ψ^{eq} and one for the non-equilibrium potential ψ^{neq} , from which we derive the equilibrium and non-equilibrium stresses, $\mathbf{S}^{\text{eq}}$ and $\mathbf{S}^{\text{neq}}$, according to Eqs. (8) and (9).

Dissipative potential. From the Clausius-Planck inequality (6), we motivate the Mandel stress Σ as the driving force for dissipative processes, with $\Sigma := 2 \mathbf{C}_e \cdot \partial \psi^{\text{neq}} / \partial \mathbf{C}_e$,

$$\Sigma = 2 \frac{\partial \psi^{\text{neq}}}{\partial \tilde{\mathbf{I}}_1^{\text{C}_e}} \left[\mathbf{C}_e - \frac{1}{3} I_1^{\text{C}_e} \mathbf{I} \right] / (I_3^{\text{C}_e})^{1/3} + 2 \frac{\partial \psi^{\text{neq}}}{\partial \tilde{\mathbf{I}}_2^{\text{C}_e}} \left[I_1^{\text{C}_e} \mathbf{C}_e - \mathbf{C}_e^2 - \frac{2}{3} I_2^{\text{C}_e} \mathbf{I} \right] / (I_3^{\text{C}_e})^{2/3} + 2 \frac{\partial \psi^{\text{neq}}}{\partial I_3^{\text{C}_e}} I_3^{\text{C}_e} \mathbf{I}. \quad (10)$$

To characterize the dissipative volume-changing and volume-preserving behavior of tofu, we introduce the volumetric invariant I_1^Σ and the deviatoric invariants J_2^Σ and J_3^Σ of the Mandel stress Σ ,

$$I_1^\Sigma = \text{tr}(\Sigma) \quad \text{and} \quad J_2^\Sigma = \frac{1}{2} \text{tr}(\text{dev}(\Sigma)^2) \quad \text{and} \quad J_3^\Sigma = \frac{1}{3} \text{tr}(\text{dev}(\Sigma)^3), \quad (11)$$

with derivatives $\partial I_1^\Sigma / \partial \Sigma = \mathbf{I}$ and $\partial J_2^\Sigma / \partial \Sigma = \text{dev}(\Sigma)$ and $\partial J_3^\Sigma / \partial \Sigma = \text{dev}(\text{dev}(\Sigma)^2)$, where $\text{tr}(\circ) = (\circ) : \mathbf{I}$ denotes the trace and $\text{dev}(\circ) = (\circ) - \frac{1}{3} I_1^\circ (\circ)$ denotes the deviator of a second-order tensor $(\circ)$. The first invariant I_1^Σ measures the hydrostatic stress, the second invariant J_2^Σ measures the magnitude of shear, and the third invariant J_3^Σ measures the magnitude and mode of shear. To characterize the evolution of the inelastic deformation, we introduce the dissipative potential ϕ that depends on the Mandel stress Σ or, alternatively, on its invariants $I_1^\Sigma, J_2^\Sigma, J_3^\Sigma$,

$$\phi = \phi(\Sigma) = \phi(I_1^\Sigma, J_2^\Sigma, J_3^\Sigma). \quad (12)$$

To a priori satisfy the reduced Clausius-Planck inequality,

$$D_{\text{red}} = \Sigma : \mathbf{D}_i \geq 0 \quad (13)$$

we postulate that the dissipative potential ϕ is convex, non-negative, and zero at the origin and define the inelastic rate of deformation tensor as $\mathbf{D}_i := \partial \phi / \partial \Sigma$,

$$\mathbf{D}_i = \frac{\partial \phi}{\partial I_1^\Sigma} \mathbf{I} + \frac{\partial \phi}{\partial J_2^\Sigma} \text{dev}(\Sigma) + \frac{\partial \phi}{\partial J_3^\Sigma} \text{dev}(\text{dev}(\Sigma)^2). \quad (14)$$

The inelastic rate of deformation tensor $\mathbf{D}_i$ introduces the evolution of the inelastic right Cauchy Green tensor, $\dot{\mathbf{C}}_i = 2 \mathbf{F}_i^t \cdot \mathbf{D}_i \cdot \mathbf{F}_i$, which we now discretize in time to calculate the inelastic right Cauchy Green tensor $\mathbf{C}_i$ and its elastic counterpart $\mathbf{C}_e$.

Time discretization. To advance the governing equations in time, from the previous time point t^n with known deformation gradient $\mathbf{F}^n$, inelastic tensor $\mathbf{F}_i^n$, right Cauchy Green tensor $\mathbf{C}_i^n$, and rate of deformation tensor $\mathbf{D}_i^n$, to the new time point $t^{n+1} = t^n + \Delta t$ with time increment Δt , we adopt an explicit time integration scheme using the exponential update of the inelastic right Cauchy Green tensor,

$$\mathbf{C}_i^{n+1} = (\mathbf{F}_i^n)^t \cdot \exp(2 \Delta t \mathbf{D}_i^n) \cdot \mathbf{F}_i^n \quad \text{with} \quad \mathbf{F}_i^n = +(\mathbf{C}_i^n)^{1/2}. \quad (15)$$

In Section 4, we design a neural network for the dissipative potential ϕ from which we derive the inelastic rate of deformation tensor $\mathbf{D}_i$, the inelastic right Cauchy Green tensor $\mathbf{C}_i$, and the inelastic part of the deformation gradient $\mathbf{F}_i$ at the new time point t^{n+1} .

Features. In many engineering applications, additional state variables or fields such as the density, temperature, or water content may affect the material behavior. To account for these effects, we extend the constitutive framework by introducing an input feature vector α that can modify the elastic and dissipative potentials,

$$\psi = \psi^{\text{eq}}(\mathbf{C}, \alpha) + \psi^{\text{neq}}(\mathbf{C}_e, \alpha) \quad \text{and} \quad \phi = \phi(\Sigma, \alpha). \quad (16)$$

Here, we acknowledge the biphasic nature of tofu and propose a single scalar-valued input feature, the water content $\alpha = n^F$. In Section 4, we design a feature network that we couple to the networks for the equilibrium potential ψ^{eq} , the non-equilibrium potential ψ^{neq} , and the dissipative potential ϕ to explore the dependencies of tofu on the water content.

Kinematics of finite inelastic confined cylindrical compression. Our experiment considers a cylindrical sample under confined compression in the z -direction. Its principal stretches in cylindrical coordinates, $\lambda_r = 1$, $\lambda_\theta = 1$, $\lambda_z = \lambda$, define the total deformation gradient $\mathbf{F}$ that we multiplicatively decompose into elastic and inelastic parts, $\mathbf{F}_e$ and $\mathbf{F}_i$,

$$\mathbf{F} = \text{diag} \{ 1, 1, \lambda \} = \mathbf{F}_e \cdot \mathbf{F}_i, \quad (17)$$

with the Jacobian, $J = \det(\mathbf{F}) = \lambda_r \lambda_\theta \lambda_z = \lambda$, defining the volumetric change of the sample. We assume that the inelastic deformation is isotropic and volume preserving, $J_i = \det(\mathbf{F}_i) = 1$, and that all volumetric changes are associated with the elastic deformation only, $J_e = \det(\mathbf{F}_e) = \lambda$ (Holzapfel, 2000; Reese and Govindjee, 1998; Simo, 1987), such that elastic and inelastic parts of the deformation gradient are

$$\mathbf{F}_e = \text{diag} \{ \lambda_i^{1/2}, \lambda_i^{1/2}, \lambda/\lambda_i \} \quad \text{and} \quad \mathbf{F}_i = \text{diag} \{ \lambda_i^{-1/2}, \lambda_i^{-1/2}, \lambda_i \}. \quad (18)$$

These explicit definitions introduce the right Cauchy Green tensor $\mathbf{C}$ and its elastic counterpart $\mathbf{C}_e$,

$$\mathbf{C} = \mathbf{F}^t \cdot \mathbf{F} = \text{diag} \{ 1, 1, \lambda^2 \} \quad \text{and} \quad \mathbf{C}_e = \mathbf{F}_e^t \cdot \mathbf{F}_e = \text{diag} \{ \lambda_i, \lambda_i, \lambda^2/\lambda_i^2 \} \quad (19)$$

with invariants

$$\begin{aligned} I_1^{\mathbf{C}} &= 2 + \lambda^2 & I_2^{\mathbf{C}} &= 1 + 2\lambda^2 & I_3^{\mathbf{C}} &= \lambda^2 \\ \tilde{I}_1^{\mathbf{C}} &= 2/\lambda^{2/3} + \lambda^{4/3} & \tilde{I}_2^{\mathbf{C}} &= 1/\lambda^{4/3} + 2\lambda^{2/3} & \tilde{I}_3^{\mathbf{C}} &= \lambda^2 \\ I_1^{\mathbf{C}_e} &= 2\lambda_i/\lambda^{2/3} + \lambda^{4/3}/\lambda_i^2 & I_2^{\mathbf{C}_e} &= 2\lambda^{2/3}/\lambda_i + \lambda_i^2/\lambda^{4/3} & I_3^{\mathbf{C}_e} &= \lambda^2. \end{aligned} \quad (20)$$

We note that, consistent with axisymmetric confined compression, all tensors are diagonal in the principal directions $\{r, \theta, z\}$, i.e., there are no shear components, and the radial and circumferential components are identical since the deformation is axisymmetric. For this special case of inelastic confined cylindrical compression, we can explicitly calculate the equilibrium stress in the axial direction,

$$S^{\text{eq}} = \frac{4}{3} \left[1 - \frac{1}{\lambda^2} \right] \left[\frac{1}{\lambda^{2/3}} \right] \frac{\partial \psi^{\text{eq}}}{\partial \tilde{I}_1^{\mathbf{C}}} + \frac{4}{3} \left[1 - \frac{1}{\lambda^2} \right] \left[\frac{1}{\lambda^{4/3}} \right] \frac{\partial \psi^{\text{eq}}}{\partial \tilde{I}_2^{\mathbf{C}}} + 2 \frac{\partial \psi^{\text{eq}}}{\partial I_3^{\mathbf{C}}}, \quad (21)$$

and the non-equilibrium stress in axial direction,

$$S^{\text{neq}} = \frac{4}{3} \left[\frac{1}{\lambda_i^2} - \frac{\lambda_i}{\lambda^2} \right] \left[\frac{1}{\lambda^{2/3}} \right] \frac{\partial \psi^{\text{neq}}}{\partial \tilde{I}_1^{\mathbf{C}_e}} + \frac{4}{3} \left[\frac{1}{\lambda_i} - \frac{\lambda_i^2}{\lambda^2} \right] \left[\frac{1}{\lambda^{4/3}} \right] \frac{\partial \psi^{\text{neq}}}{\partial \tilde{I}_2^{\mathbf{C}_e}} + 2 \frac{\partial \psi^{\text{neq}}}{\partial I_3^{\mathbf{C}_e}}, \quad (22)$$

solely in terms of total and inelastic stretches, λ and λ_i , and the derivatives of the elastic potentials, ψ^{eq} and ψ^{neq} , that we will now discover using constitutive neural networks.

4. Neural network model

To discover the model and parameters that best describe our experimental data, we introduce a combination of four constitutive neural networks for the equilibrium potential ψ^{eq} (Linka and Kuhl, 2023), the non-equilibrium potential ψ^{neq} (Holthausen et al., 2025a), the dissipative potential ϕ (Holthausen et al., 2025b), and the weight vector F^{max} . This architecture combines two feed forward neural networks for the equilibrium and non-equilibrium potentials with a recurrent neural network for the dissipative potential within the framework of inelastic constitutive neural networks, and, if desired, combines the resulting overall architecture with a parallel feed forward network for the weight vector. It takes the known deformation state $\mathbf{F}^n$ and the internal variables $\mathbf{F}_i^n$ from the previous time point t^n as input, integrates them in time using the exponential update (15), stores the internal variables $\mathbf{F}_i^{n+1}$, and outputs the second Piola Kirchhoff stress $\mathbf{S}^{n+1}$ at the new time point t^{n+1} .

Elastic networks. To model the equilibrium potential ψ^{eq} (Linka and Kuhl, 2023) and the non-equilibrium potential ψ^{neq} (Holthausen et al., 2025b), we introduce two structurally similar but input-independent feed forward networks. The equilibrium network uses the invariants of the total right Cauchy Green tensor $\tilde{I}_1^{\mathbf{C}}, \tilde{I}_2^{\mathbf{C}}, I_3^{\mathbf{C}}$ from Eq. (3) (Linka et al., 2023a; Martonová et al., 2024; Peirlinck et al., 2025) and the non-equilibrium network uses the invariants of the elastic right Cauchy Green tensor $\tilde{I}_1^{\mathbf{C}_e}, \tilde{I}_2^{\mathbf{C}_e}, I_3^{\mathbf{C}_e}$ from Eq. (4) (Holthausen et al., 2024b,a, 2025a) as inputs. For both networks, we ensure expressiveness and interpretability by adopting a sparse network architecture with two hidden layers, where the activation functions f^I of the first layer are the first and second powers and the activation functions f^{II} of the second layer are the identity, the exponential function, and the natural logarithm,

$$f^I \in \left\{ \begin{pmatrix} (\circ)^1 \\ (\circ)^2 \end{pmatrix} \right\} \quad \text{and} \quad f^{II} \in \left\{ \begin{pmatrix} (\circ) \\ \exp(\circ) - 1 \\ (\circ) - 1 - \ln(\circ) \end{pmatrix} \right\}.$$

Fig. 3 visualizes the architecture of both elastic networks that a priori satisfy the thermodynamic requirements of polyconvexity, growth, and normalization (Boes et al., 2024; Holthausen et al., 2025b; Linka and Kuhl, 2023). The equilibrium and non-equilibrium networks each contain 14 trainable parameters, which are visualized in Fig. 3 through solid lines. We constrain all weights to be non-negative and add Lasso regularization to the weights of the final layer. Importantly, although the two feed forward networks for the equilibrium and non-equilibrium potentials are input-independent, they are coupled through their derivatives.

Dissipative network. To model the dissipative potential ϕ (Holthausen et al., 2025b), we introduce a recurrent neural network that uses the invariants of the Mandel stress $I_1^{\Sigma}, J_2^{\Sigma}, J_3^{\Sigma}$ from Eq. (11) as input. Specifically, to maintain consistent physical units, we take their first, second, and third roots as input. We adopt a fully connected network architecture with three hidden layers, where

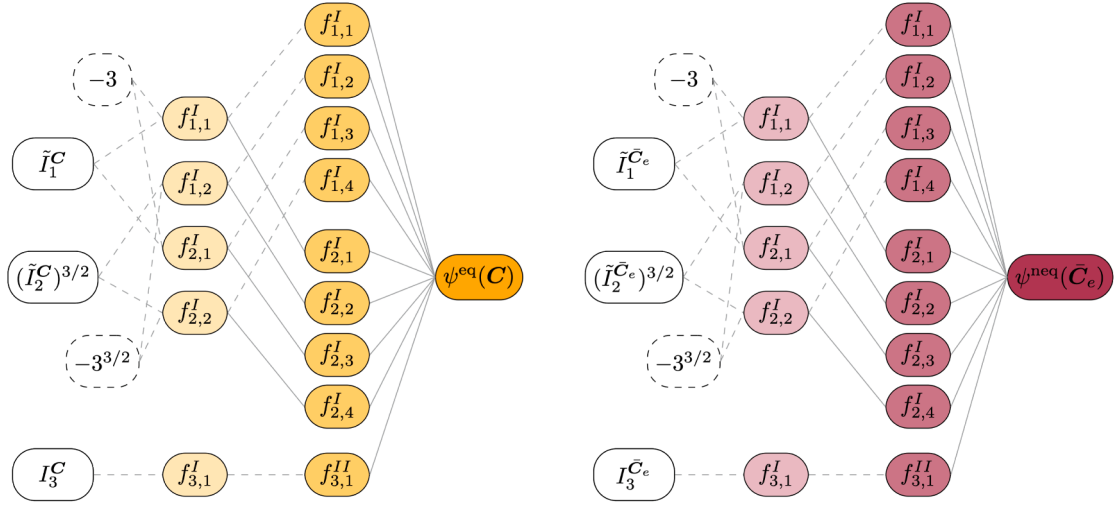


Fig. 3. Elastic networks. The elastic networks have two hidden layers with 5 and 9 nodes per layer. The feed forward networks take the invariants of the total right Cauchy Green tensor $\bar{I}_1^C, \bar{I}_2^C, I_3^C$, left, or of the elastic right Cauchy Green tensor $\bar{I}_1^{C_e}, \bar{I}_2^{C_e}, I_3^{C_e}$, right, as input; the first layer applies first or second powers, the second layer applies the identity, exponential function, or natural logarithm; the networks learn the equilibrium or non-equilibrium potentials ψ^{eq} , left, or ψ^{neq} , right, as output. Activation functions $f_{\beta,\gamma}^\delta$ are labeled by activation type β , node index γ , and layer index δ ; solid lines denote trainable weights, dashed lines indicate fixed weights equal to one, and dashed ellipses denote constant, non-trainable biases.

the activation functions f^I of the first layer are trainable powers p of the absolute input values $|(\circ)|$, and the activation functions f^{II} and f^{III} of the second and third layers are the maximum or exponential functions,

$$f^I \in \left\{ |(\circ)|^{p_1+1}, |(\circ)|^{p_2+1}, |(\circ)|^{p_3+1} \right\} \quad \text{and} \quad f^{II} \in \left\{ \max((\circ), 0), \exp((\circ) - 1) \right\} \quad \text{and} \quad f^{III} \in \left\{ \max((\circ), 0), \exp((\circ) - 1) \right\}.$$

Fig. 4 visualizes the architecture of the dissipative network that a priori satisfies thermodynamic consistency by remaining non-negative, convex, and zero-valued at equilibrium to ensure that the reduced dissipation inequality $\mathcal{D}_{\text{red}} \geq 0$ (13) is satisfied. In the first layer, we smoothen the absolute-value functions $|(\circ)|$ near the origin for differentiability, constrain the exponents p_1, p_2, p_3 to be non-negative, and allow the network weights to take positive or negative values (Holthusen et al., 2025b). In the second and third layers, we apply additional non-positive biases to the maximum functions and constrain the weights to be non-negative. To guarantee non-negativity, we introduce a final layer that mirrors the third layer, but uses the maximum function without additional weights. After evaluating multiple network configurations with varying numbers of neurons per layer, we select the architecture that achieves the best performance in a preliminary study: a three-layered network with $\{18, 8, 8\}$ nodes per layer. The resulting dissipative network has a total number of 281 trainable parameters. Appendix B provides additional sensitivity analyses for varying network architectures and varying activation functions.

Feature network. To model the impact of the feature vector on the mechanical response, we introduce a feed forward network that uses the water content as a single scalar-valued input feature α . We adopt a fully connected network architecture with three hidden layers, where the activation functions F^I, F^{II}, F^{III} are the maximum, identity, hyperbolic tangent, and softplus functions and the output of the final layer F^{max} applies the maximum function,

$$F^I, F^{II}, F^{III} \in \left\{ \begin{array}{l} \max((\circ), 0) \\ (\circ) \\ \tanh((\circ)) \\ \tanh((\circ))^2 \end{array} \right\} \quad \text{and} \quad F^{\text{max}} \in \left\{ \max((\circ), 0) \right\}.$$

Fig. 5 schematically visualizes the architecture of the feature network. To a priori ensures non-negative, non-zero-valued outputs F^{max} , we add biases to all maximum functions. After evaluating different network architectures, we select feature networks with $\{14, 15, 11\}$ nodes per activation function per layer for the equilibrium potential, $\{13, 17, 15, 9\}$ for the non-equilibrium potential, and $\{20, 6, 6\}$ for the dissipative potential. The resulting feature network adds a total number of 19,494 trainable parameters. The number of nodes in the final layer n^{max} matches those of the mechanical network to which it is coupled, with $n^{\text{max}} = 9$ nodes for the elastic networks in Fig. 3 and $n^{\text{max}} = 8$ nodes for the dissipative network in Fig. 4.

Fig. 6 illustrates the coupling between one of the three mechanical networks and the feature network. The mechanical networks from Figs. 3 or 4 receive a strain- or stress-like input $\mathbf{A}$ and the feature network from Fig. 5 receives a feature-like input α . The networks

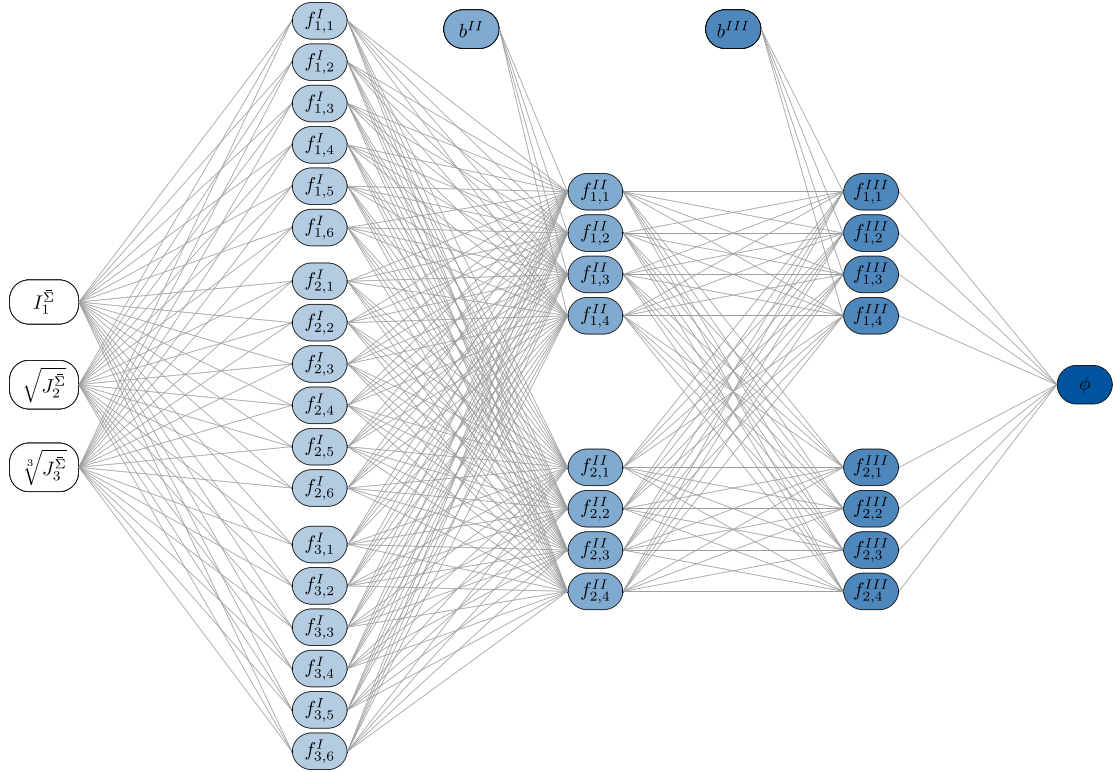


Fig. 4. Dissipative network. The dissipative network has three hidden layers with 18, 8, and 8 nodes per layer. It takes the invariants of the Mandel stress I_1^Σ , J_2^Σ , J_3^Σ as input; the first layer applies trainable powers p to the absolute input values, the second and third layers apply the maximum or exponential functions; the network learns the dissipative potential ϕ as output. Activation functions $f_{\beta,\gamma}^\delta$ are labeled by activation type β , node index γ , and layer index δ ; biases b^δ are applied to the maximum functions in the second and third hidden layers.

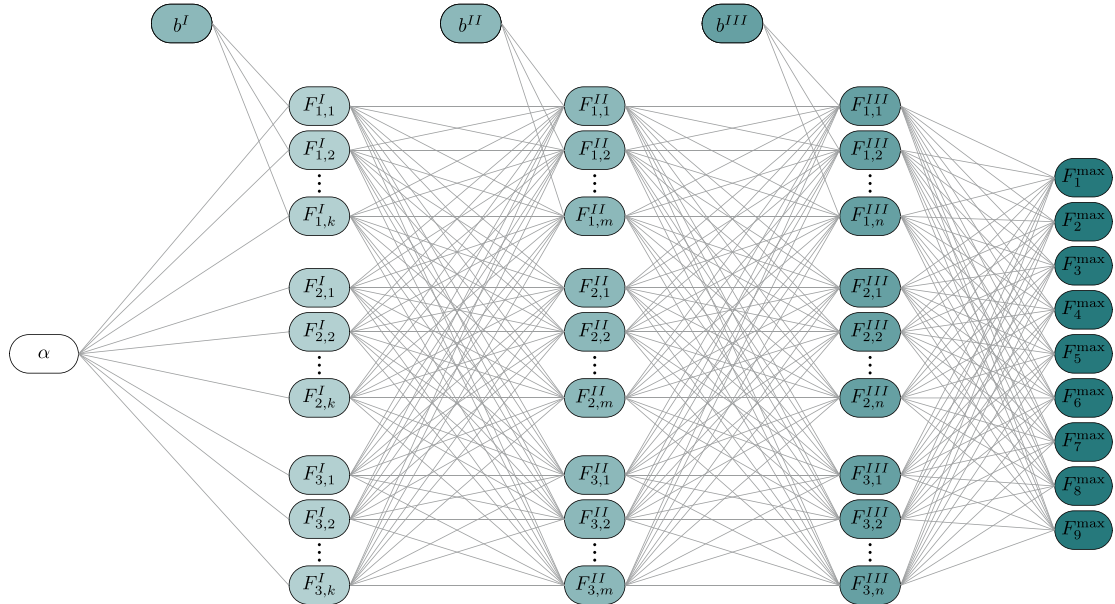


Fig. 5. Feature network. The feature network has three hidden layers with $k \times 4$ nodes per layer. It takes the feature α , in our case the scalar-valued water content, as input; the three hidden layer apply the maximum, identity, hyperbolic tangent, and softplus functions; the output of the final layer applies the maximum function; the network learns the feature weight vector $F^{\max}$ as output. Activation functions $f_{\beta,\gamma}^\delta$ are labeled by activation type β , node index γ , and layer index δ ; biases b^δ are applied to all maximum functions.

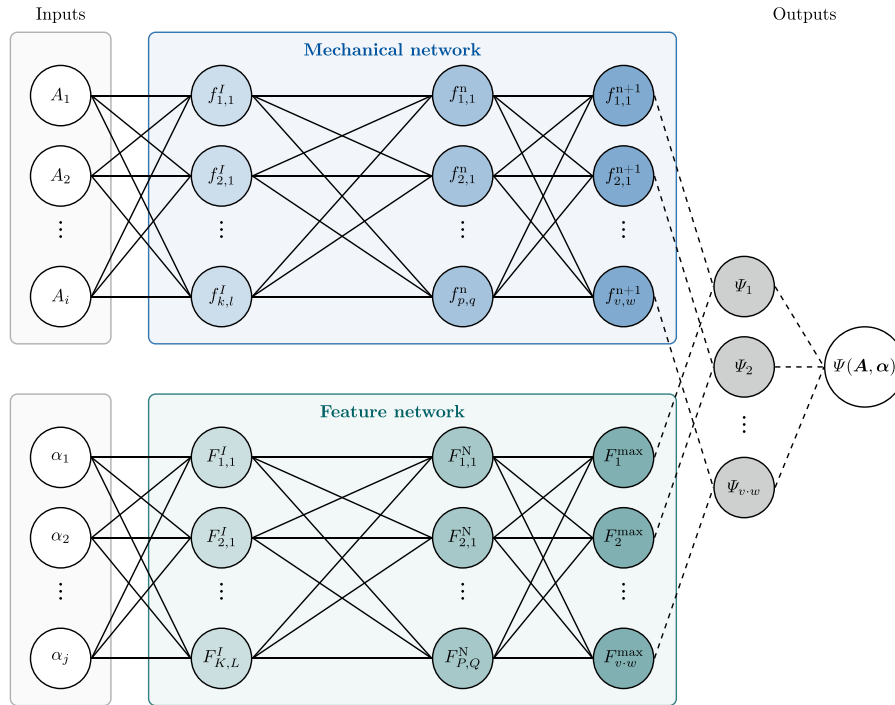


Fig. 6. Coupling of mechanical and feature networks. The mechanical network from Figs. 3 or 4 receives a strain- or stress-like input $\mathbf{A}$ and the feature network from Fig. 5 receives a feature-like input α . The networks are multiplicatively coupled in the nodes of their final layers. The sum of the final layer defines the network output, the feature-weighted mechanical potential $\Psi(\mathbf{A}, \alpha)$. Activation functions f_γ^δ and F_γ^δ of the mechanical and feature networks are labeled by node index γ and layer index δ .

are multiplicatively coupled by summing up the product of the nodes in their final layers to obtain feature-weighted potential Ψ ,

$$\Psi(\mathbf{A}, \alpha) = \sum_{i=1}^{n^{\max}} f_i^{\max}(\mathbf{A}) \cdot F_i^{\max}(\alpha) \quad \text{with} \quad \Psi = \{\psi^{\text{eq}}, \psi^{\text{neq}}, \phi\} \quad \text{and} \quad \mathbf{A} = \{\mathbf{C}, \mathbf{C}^e, \Sigma\},$$

where the potential Ψ is either the equilibrium or non-equilibrium potential ψ^{eq} or ψ^{neq} , or the dissipative potential ϕ , and the input $\mathbf{A}$ is either the total or the elastic right Cauchy Green tensor $\mathbf{C}$ or $\mathbf{C}^e$ or the Mandel stress Σ . Multiplicative feature weighting inherently preserves polyconvexity and other thermodynamic requirements.

Network training. To discover the best model and parameters to characterize tofu, we train and test our networks using the experimental data from Section 2. We introduce the loss function as the mean squared error between the model stress S and the experimental stress $\hat{S}$ in the direction of loading, averaged across all n_{exp} experiments and all n_{data} data points. To induce sparsity and improve generalization, we supplement the loss with L_1 -regularization or Lasso and L_2 -regularization or ridge regression across all n_w weights, where α_1 and α_2 are the regularization coefficients. The total loss function reads

$$L(S, \mathbf{w}) = \frac{1}{n_{\text{exp}}} \sum_{k=1}^{n_{\text{exp}}} \frac{1}{n_{\text{data}}} \sum_{m=1}^{n_{\text{data}}} \left\| S_m(\mathbf{C}_m, \mathbf{F}_{i,m}, \mathbf{w}) - \hat{S}_m \right\|_k^2 + \sum_{n=1}^{n_w} \alpha_1 |w_n| + \sum_{n=1}^{n_w} \alpha_2 (w_n)^2. \quad (23)$$

We perform two types of model discovery: First, we analyze *each tofu type independently* by using only the elastic and dissipative networks to discover *three individual tofu models* for each different tofu type. Then, we analyze *all three tofu types combined* by integrating the elastic and dissipative networks with the feature network to discover a *single unified tofu model* with the water content as input feature. Appendix B provides additional sensitivity analyses for varying regularization coefficients.

Discovering three individual tofu models. First, we train the elastic and dissipative networks to discover individual tofu models using the stress–strain data from Section 2. We neglect water-content dependency and discover three independent material models by simultaneously learning the equilibrium and non-equilibrium potentials ψ^{eq} and ψ^{neq} and the dissipative potential ϕ . We use the triple-compression tests at three different strain rates as training data and the double-compression tests at intermediate strain rates as test data. To capture history-dependent effects, we adopt a staggered training strategy that progressively increases the batch size and gradually introduces additional training data along the load path. We adopt the Adam optimizer with a learning rate of 10^{-3} and a clip norm of 10^{-3} . The training proceeds in two phases: In the first phase, we only include the first $n = 50$ data points of the first cycle of the triple-compression test. In the second phase, we include all cycles and apply a regularization with $\alpha_1 = 0.0001$ and to all layers of all networks and $\alpha_2 = 0.0001$ to the first layer of the dissipative network. The weights and biases from the first phase initialize the second phase to ensure continuity and convergence throughout the process.

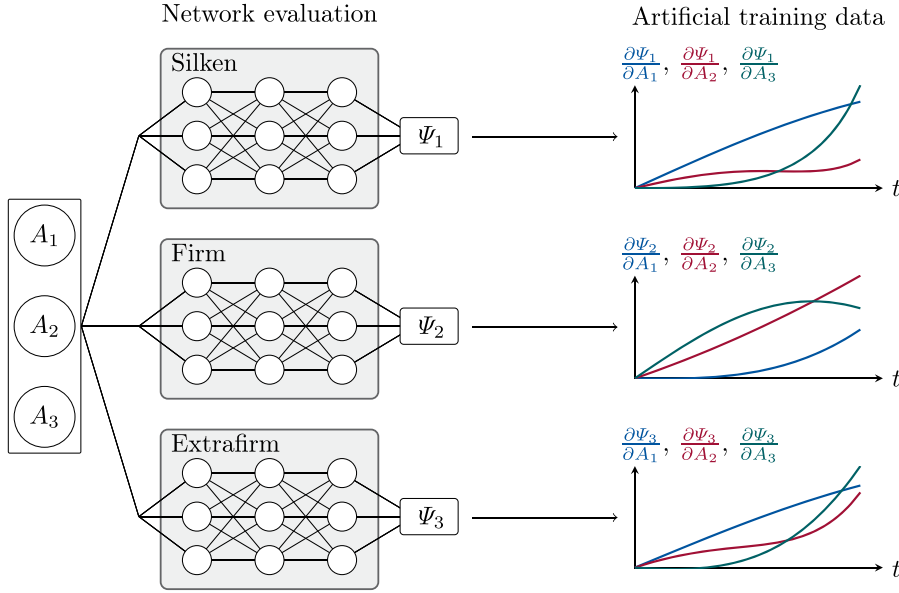


Fig. 7. Generating synthetic training data. After discovering the best models and parameters for silken, firm, and extrafirm tofu by training the equilibrium and non-equilibrium networks in Fig. 3 and the dissipative network in Fig. 4 on the experimental data, we generate synthetic data to train the coupled mechanical and feature networks in Fig. 6. We evaluate the three tofu models for sets of invariants I_i^A , and their derivatives $\hat{\Psi}/\partial I_i^A$ provide the training data for the unified feature-mechanical model.

Discovering a single unified tofu model. Then, we train the feature networks to discover a unified tofu model for the stress-strain data from Section 2. We incorporate the water-content as a scalar-valued input feature and discover a single material model by learning feature-weighted terms of the equilibrium, non-equilibrium, and dissipative potentials, ψ^{eq} and ψ^{neq} and ϕ . Yet, unlike the elastic and dissipative networks, we do not train the feature network directly on the experimental data. Instead, we train on the *outputs of the discovered models*. We generate synthetic training data by evaluating our three discovered models for various sets of invariants $I_i^A = \{I_i^C, I_i^{C_e}, I_i^\Sigma\}$ of our network input $\mathbf{A} = \{\mathbf{C}, \mathbf{C}_e, \Sigma\}$, the total and elastic right Cauchy Green tensors $\mathbf{C}$ and $\mathbf{C}_e$ and the Mandel stress Σ . According to Eqs. (8), (9), and (14), the derivatives of the potentials $\Psi = \{\psi^{\text{eq}}, \psi^{\text{neq}}, \phi\}$ with respect to these invariants, $\partial\Psi/\partial I_i^A$, determine both the stress response and the inelastic flow direction $\{\mathbf{S}^{\text{eq}}, \mathbf{S}^{\text{neq}}, \mathbf{D}_1\}$. We train the combined feature-mechanical network on these derivatives, which contain the essential physical information.

Fig. 7 illustrates how we generate our synthetic training data $\partial\hat{\Psi}_\alpha/\partial I_i^A$: We evaluate each of the three discovered network potentials $\hat{\Psi}_\alpha$ for a series of invariants I_i^A , where $\alpha = \{n_{\text{silken}}, n_{\text{firm}}, n_{\text{extra}}\}$ is the water content of silken, firm, and extrafirm tofu and $i = \{1, 2, 3\}$ are the three isotropic invariants. The derivatives of the potentials with respect to the invariants, $\partial\hat{\Psi}_\alpha/\partial I_i^A$, represent the training data that we seek to approximate by the unified model $\Psi(\mathbf{A}, \alpha, \mathbf{w})$ of the feature-mechanical network that takes both the invariants of $\mathbf{A}$ and the feature α as input. During training, the networks learn the weights $\mathbf{w}$ by minimizing the loss function that measures the error between the derivatives of the unified model $\partial\Psi_{\alpha,m}/\partial I_i^A$ and the derivatives of the synthetic training data $\partial\hat{\Psi}_{\alpha,m}/\partial I_i^A$ at $m = 1, \dots, n_{\text{data}}$ data points,

$$L(\partial\Psi/\partial I_i^A, \mathbf{w}) = \sum_{\alpha=1}^{n_{\text{tofu}}} \sum_{m=1}^{n_{\text{data}}} \sum_{i=1}^3 \left\| \frac{\partial\Psi_m(\mathbf{A}_m, \alpha_m, \mathbf{w})}{\partial I_i^A} - \frac{\partial\hat{\Psi}_m}{\partial I_i^A} \right\|. \quad (24)$$

We adopt the Adam optimizer and train the equilibrium network for 5000 epochs using 500 training samples with a learning rate of 10^{-3} and a clip norm of 10^{-3} , and the non-equilibrium and dissipative networks for 15,000 and 10,000 epochs with a learning rate of 10^{-4} . This approach eliminates the need to retrain the entire recurrent network on experimental stress-strain pairs and significantly accelerates the learning process. Importantly, it produces a *single tofu model* that captures tofu types of different water content within a unified constitutive framework.

5. Results

Mechanical testing. Figs. 8 and 9 summarize the applied stretch and stress histories $\lambda(t)$ and $S(t)$ during the triple-compression and double-compression tests. Each row corresponds to one tofu type—silken, firm, and extrafirm—and each column represents a different loading condition. The first three columns show the triple-compression tests at different strain rates and the final column shows the double-compression tests.

The curves display high reproducibility across all tofu types and loading conditions and confirm that we can robustly record stretch and stress histories at different stretch magnitudes and rates. However, the testing device displays limitations at higher strain

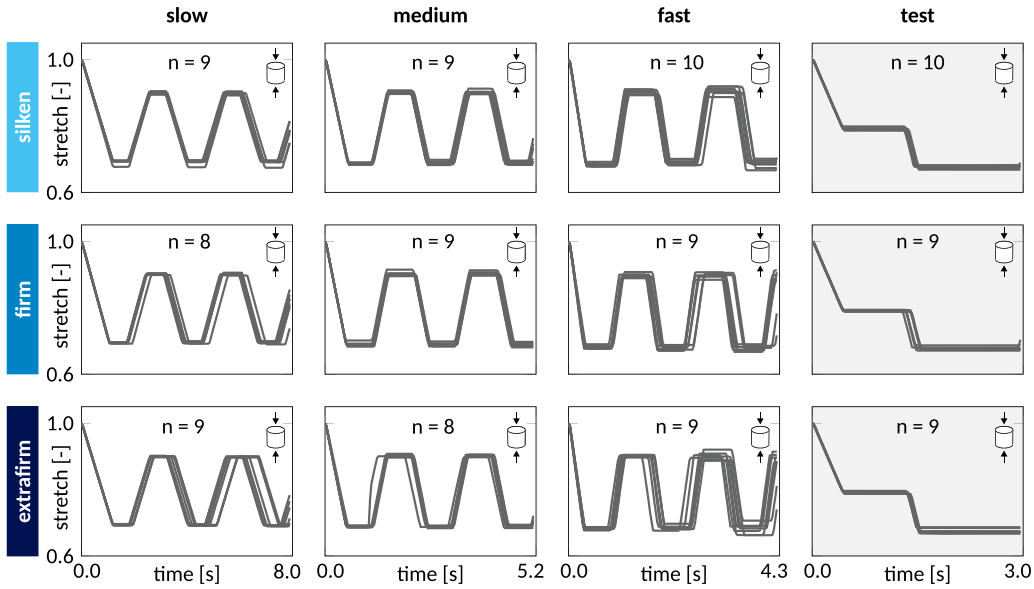


Fig. 8. Mechanical testing - Experimentally applied stretches. Stretch histories $\lambda(t)$ during triple-compression and double-compression tests for silken (top), firm (middle), and extrafirm (bottom) tofu. The first three columns show the triple-compression tests at slow, medium, and fast strain rates of 20%/s, 40%/s, and 60%/s (from left to right), where each specimen is loaded to a stretch of 0.7 and unloaded to 0.9. The last column shows the double-compression tests at a strain rates of 30%/s to a stretch of 0.8 followed by a strain rate of 50%/s to 0.7.

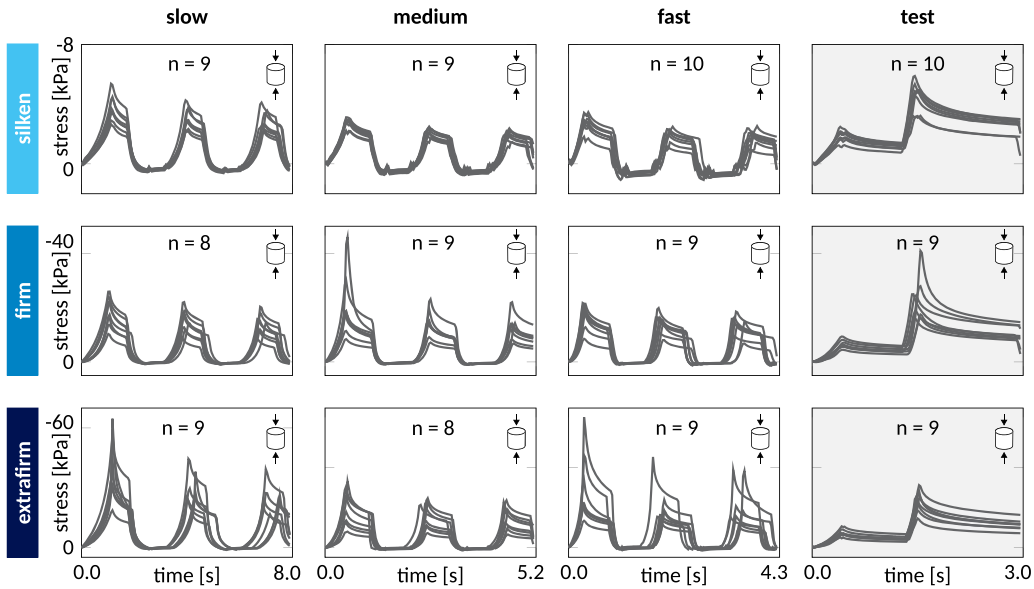


Fig. 9. Mechanical testing - Experimentally recorded stresses. Stress histories $S(t)$ during triple-compression and double-compression tests for silken (top), firm (middle), and extrafirm (bottom) tofu. The first three columns show the triple-compression tests at slow, medium, and fast strain rates of 20%/s, 40%/s, and 60%/s (from left to right), where each specimen is loaded to a stretch of 0.7 and unloaded to 0.9. The last column shows the double-compression tests at a strain rates of 30%/s to a stretch of 0.8 followed by a strain rate of 50%/s to 0.7.

rates. As a result, we observe slight variations in the applied stretch profiles resulting in slight inconsistencies in the stress profiles, which complicates direct averaging of the curves. We thus report each curve as an independent measurement. As a general trend, we observe a non-linear stress response that is sensitive to both strain rate and tofu type: extrafirm tofu is the stiffest, closely followed by firm tofu, and silken tofu is the softest. The stiffness decreases with increasing water content, consistent with the transition from solid-like to fluid-like behavior.

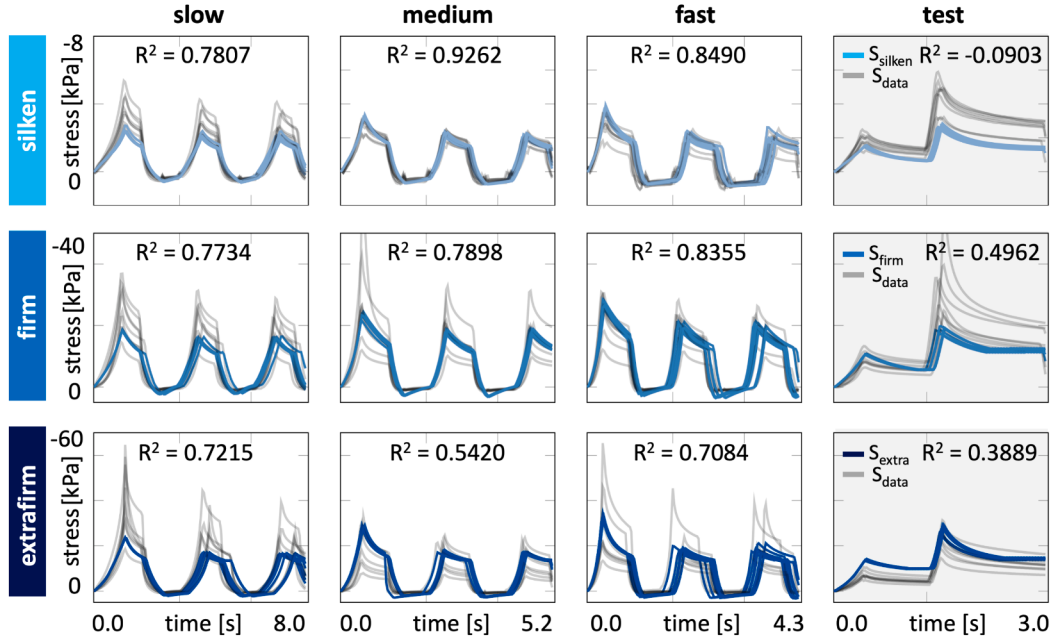


Fig. 10. Model discovery - Experimentally recorded and computationally predicted stresses of the three individual tofu models. Stress histories $S(t)$ for silken (top), firm (middle), and extrafirm (bottom) tofu. The first three columns show the triple-compression tests used as training data. The last column shows double-compression tests used as test data. Experimental data are plotted in gray, model predictions for silken, firm, and extrafirm tofu in light to dark blue; R^2 values indicate the goodness of fit.

Discovering three individual tofu models.

Fig. 10 compares the computationally predicted stresses in blue and the experimentally recorded stresses in gray. The rows correspond to silken, firm, and extrafirm tofu, and the columns represent distinct loading cases with varying stretch magnitudes and rates. The three discovered models and parameter sets for each tofu type accurately reproduce the experimental trends: For the training data in the first three columns, the modeled stresses generally fall within the measured stress ranges and accurately capture the mean behavior during loading, holding, and unloading, with coefficients of determination ranging from $R^2_{\min} = 0.5420$ to $R^2_{\max} = 0.9262$. Appendix A provides additional details about the errors. For the test data in the last column, the silken tofu model accurately reproduces the first peak and the relaxation behavior, but slightly underestimates the second peak; the firm tofu model displays a similar trend; and the extrafirm model aligns closely with all experimental curves. The test results show a decrease in the coefficient of determination, but the overall values of $R^2_{\text{silken}} = 0.7105$, $R^2_{\text{firm}} = 0.7471$ and $R^2_{\text{extra}} = 0.6564$ suggest good fits for all three tofu types. Across all tofu types, the discovered models capture a rate dependent behavior: the higher the strain rate, the stiffer the response and the larger the peak stress. In a few cases, the slowest strain rate generates higher peak stresses than the intermediate rate, but overall, the discovered models correctly reproduce strain-rate stiffening. For silken tofu, the discovered model features a quadratic second-invariant term and a logarithmic third-invariant term for the equilibrium potential and a linear second-invariant term for the non-equilibrium potential,

$$\begin{aligned}\psi_{\text{silken}}^{\text{eq}}(\mathbf{C}) &= \frac{1}{2} \mu_7 [(\tilde{I}_2^{\text{C}})^{3/2} - 3^{3/2}]^2 + \alpha_9 [(I_3^{\text{C}})^{\beta_9} - 1 - \beta_9 \ln(I_3^{\text{C}})] \\ [1.\text{pt}] \psi_{\text{silken}}^{\text{neq}}(\mathbf{C}_e) &= \frac{1}{2} \mu_5 [(\tilde{I}_2^{\text{C}_e})^{3/2} - 3^{3/2}],\end{aligned}$$

with network weights $\mu_7 = 0.158$ kPa, $\alpha_9 = 0.052$ kPa, $\beta_9 = 5.507$ for the equilibrium part and $\mu_5 = 1.179$ kPa for the non-equilibrium part that have the physical unit of stress and a clear physical interpretation as stiffness-like parameters. For firm tofu, the discovered model features a quadratic first-invariant term and a linear second-invariant term for the equilibrium potential and a linear second-invariant term and a logarithmic third-invariant term for the non-equilibrium potential,

$$\begin{aligned}\psi_{\text{firm}}^{\text{eq}}(\mathbf{C}) &= \frac{1}{2} \mu_3 [\tilde{I}_1^{\text{C}} - 3]^2 + \frac{1}{2} \mu_5 [(\tilde{I}_2^{\text{C}})^{3/2} - 3^{3/2}] \\ [1.\text{pt}] \psi_{\text{firm}}^{\text{neq}}(\mathbf{C}_e) &= \frac{1}{2} \mu_5 [(\tilde{I}_2^{\text{C}_e})^{3/2} - 3^{3/2}] + \alpha_9 [(I_3^{\text{C}_e})^{\beta_9} - 1 - \beta_9 \ln(I_3^{\text{C}_e})],\end{aligned}$$

with network weights $\mu_3 = 5.250$ kPa, $\mu_5 = 0.136$ kPa for the equilibrium part and $\mu_5 = 4.952$ kPa, $\alpha_9 = 0.867$ kPa, $\beta_9 = 1.439$ for the non-equilibrium part, again, all with a clear physical interpretation. For extrafirm tofu, the discovered model features a quadratic second-invariant term and a logarithmic third-invariant term for the equilibrium potential and a linear and a quadratic second-

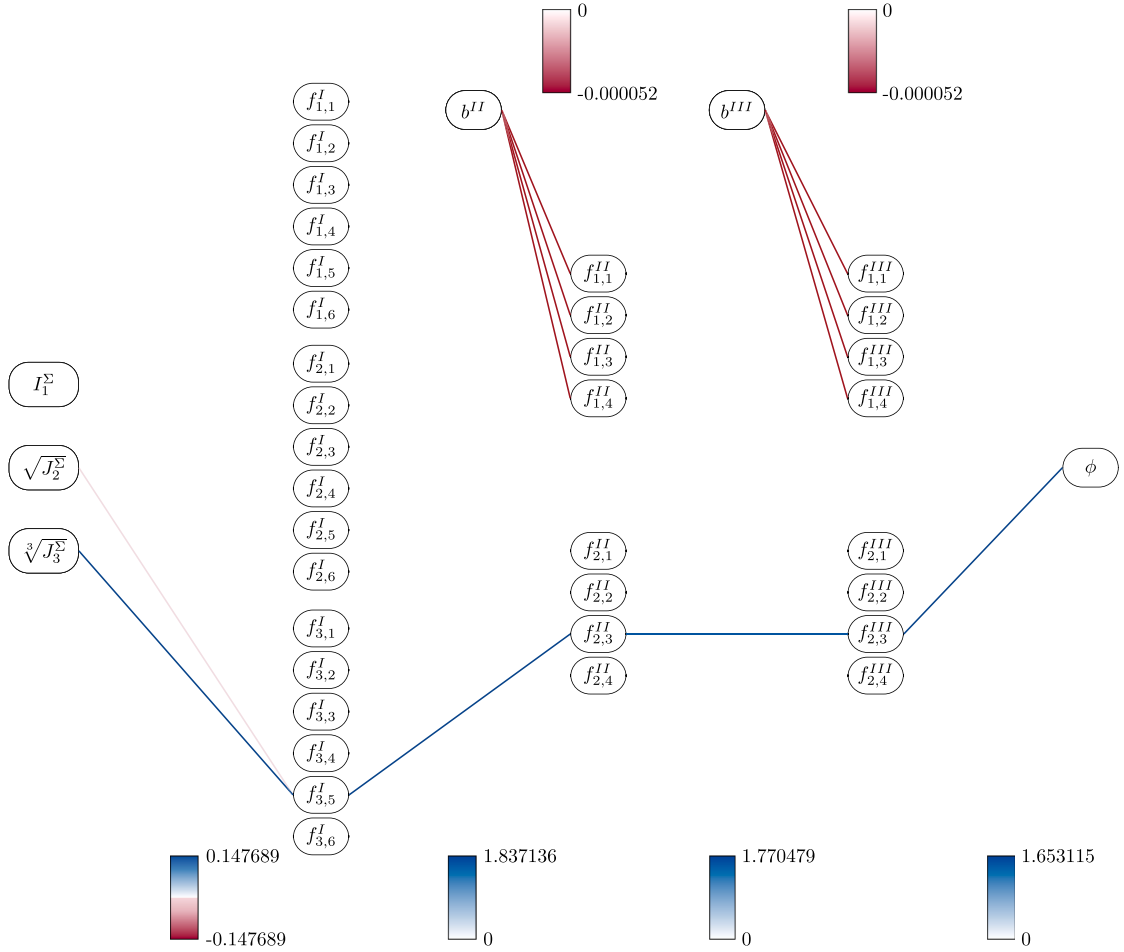


Fig. 11. Model discovery - Trained dissipative network for silken tofu. The colors represent the magnitude of the final weights and biases, with blue indicating positive and red negative values, and white indicating that this path is inactive. The network discovers a single-term potential for silken tofu that only depends on the third deviatoric invariant J_3^Σ of the Mandel stress Σ .

invariant term for the non-equilibrium potential,

$$\begin{aligned} \psi_{\text{extra}}^{\text{eq}}(\mathbf{C}) &= \frac{1}{2} \mu_7 [(\tilde{I}_2^{\mathbf{C}})^{3/2} - 3^{3/2}]^2 + \alpha_9 [(I_3^{\mathbf{C}})^{\beta_9} - 1 - \beta_9 \ln(I_3^{\mathbf{C}})] \\ [1.\text{pt}] \psi_{\text{extra}}^{\text{neq}}(\mathbf{C}_e) &= \frac{1}{2} \mu_5 [(\tilde{I}_2^{\mathbf{C}_e})^{3/2} - 3^{3/2}] + \frac{1}{2} \mu_7 [(\tilde{I}_2^{\mathbf{C}_e})^{3/2} - 3^{3/2}]^2 \end{aligned}$$

with $\mu_7 = 0.151$ kPa, $\alpha_9 = 0.103$ kPa, $\beta_9 = 4.920$ for the equilibrium part and $\mu_5 = 8.160$ kPa, $\mu_7 = 2.735$ kPa for the non-equilibrium part. Notably, the most prevalent invariant across all six potentials is the second invariant $\tilde{I}_2^{\mathbf{C}_e}$ featured in seven terms, followed by the third invariant $I_3^{\mathbf{C}_e}$ in three terms, and the first invariant $I_1^{\mathbf{C}_e}$ in only one term. None of the models features any of the four exponential terms, neither linear nor quadratic, neither in the first nor in the second invariant, and none of the models features a quadratic first-invariant term.

The dissipative part of our model discovery uses the dissipative network in Fig. 4 to discover the dissipative potential ϕ . Figs. 11–13 display the three discovered potentials for silken, firm, and extrafirm tofu. The colors represent the magnitude of the final weights and enable direct interpretation of active pathways. For silken tofu in Fig. 11, the dissipative potential is driven by both the second and third deviatoric invariants J_2^Σ and J_3^Σ ,

$$\phi_{\text{silken}} = \max(1.653 [\exp(1.770 \cdot \exp(1.837 \cdot | -0.058 (J_2^\Sigma)^{1/2} + 0.148 (J_3^\Sigma)^{1/3} |^{1.14}) - 1) - 1], 0).$$

For firm tofu in Fig. 12, the dissipative potential is driven exclusively by the second deviatoric invariant J_2^Σ ,

$$\phi_{\text{firm}} = \max(0.780 [\exp(0.814 \cdot \exp(0.857 \cdot |0.457 (J_2^\Sigma)^{1/2} |^{1.036}) - 1) - 1], 0).$$

For extrafirm tofu in Fig. 13, the dissipative potential is driven exclusively by the third deviatoric invariant J_3^Σ ,

$$\phi_{\text{extra}} = \max(1.698 [\exp(1.710 \cdot \exp(1.717 \cdot |0.156 (J_3^\Sigma)^{1/3} |^{2.42}) - 1) - 1], 0).$$

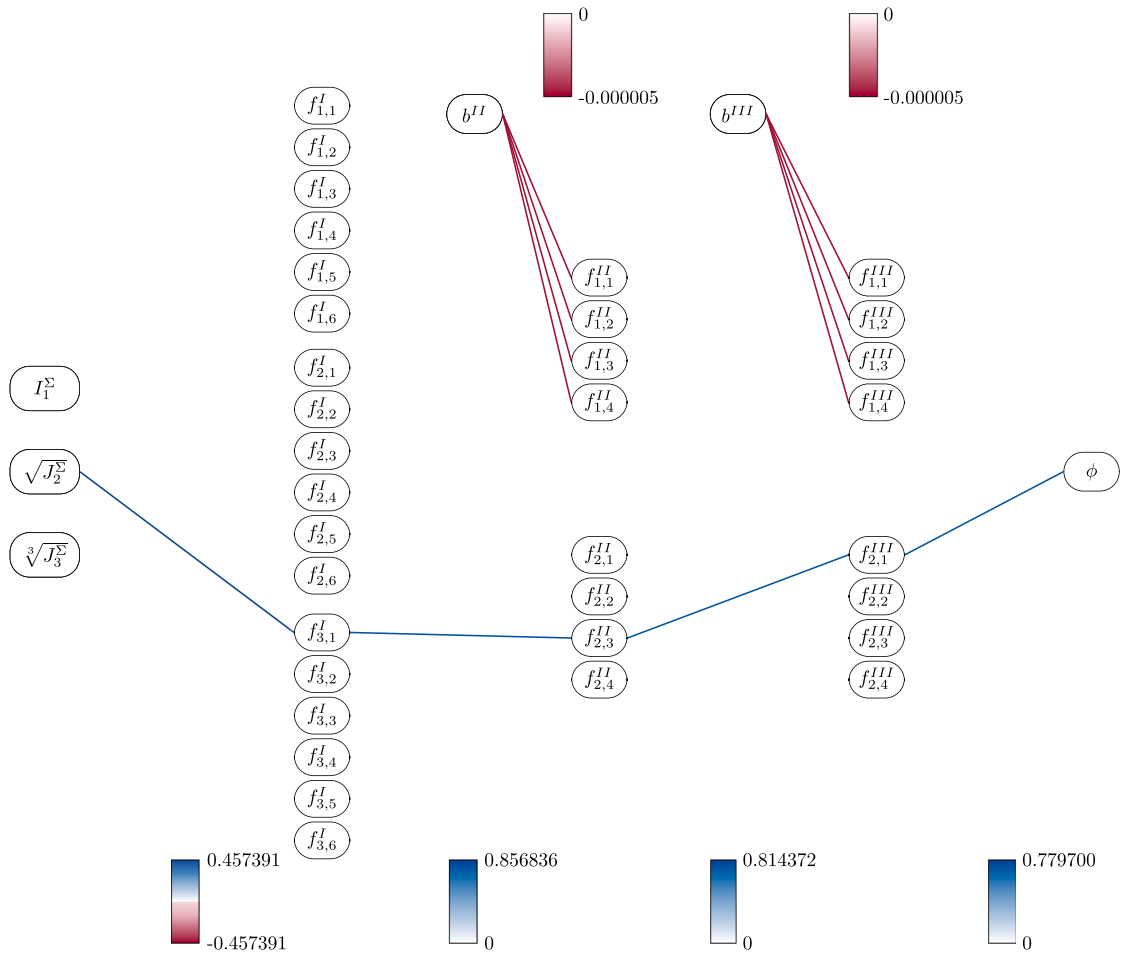


Fig. 12. Model discovery - Trained dissipative network for firm tofu. The colors represent the magnitude of the final weights and biases, with blue indicating positive and red negative values, and white indicating that this path is inactive. The network discovers a single-term potential for firm tofu that depends on the second and third deviatoric invariants J_2^Σ and J_3^Σ of the Mandel stress Σ .

Strikingly, for all three tofu types, the networks are sparse and discover the same functional form: a single term that contains weighted powers of the second and third deviatoric invariants J_2^Σ and J_3^Σ embedded in two nested exponential functions. None of the discovered potentials depends on the volumetric invariant I_1^Σ , and none of them features any biases that would result in additive contributions. Fig. 14 illustrates the stresses and the inelastic deformation rates of the discovered models for silken, firm, and extrafirm tofu for a strain rate of 60%/s. The first row shows the applied stretches λ , followed by the total stress S , its equilibrium and non-equilibrium parts S^{eq} and S^{neq} , and the inelastic deformation rate $\dot{C}_i$. The color shades denote contributions from the three invariants $\{\tilde{I}_1^C, \tilde{I}_2^C, \tilde{I}_3^C\}$ and $\{\tilde{I}_1^{C^e}, \tilde{I}_2^{C^e}, \tilde{I}_3^{C^e}\}$ and $\{I_1^\Sigma, J_2^\Sigma, J_3^\Sigma\}$. Strikingly, the predicted peak stress S in the second row increases notably with decreasing water content, from -3.90 kPa for silken to -19.62 kPa for firm and -33.85 kPa for extrafirm tofu, from left to right. The peak equilibrium stress S^{eq} in the third row increases also, from -1.26 kPa for silken to -2.44 kPa for firm and -2.76 kPa for extrafirm tofu, but less drastically than the total stress S . Notably, the peak non-equilibrium stress S^{neq} in the fourth row is almost entirely made up of second-invariant terms in $\tilde{I}_2^C$ across all three tofu types. Similarly, the evolution of the inelastic right Cauchy tensor $\dot{C}_i$ in the bottom row is made up only of second- and third-invariant terms in J_2^Σ and J_3^Σ , suggesting that the inelastic deformation remains entirely deviatoric. Importantly, as the water content decreases, from left to right, the equilibrium contribution S^{eq} to the total stress S decreases: It accounts for 1/3 of the total stress in silken, but for only 1/8 in firm and 1/12 in extrafirm tofu.

Discovering a single unified tofu model.

Fig. 15 summarizes the water-content dependency of the single unified tofu model. The rows correspond to the equilibrium, non-equilibrium, and dissipative potentials, the columns show the discovered dependencies of the individual terms, and the dashed lines indicate the water contents of our extrafirm, firm, and silken tofu. Interestingly, the majority of terms, six of eight, depend on the water content of tofu, while two terms do not: The equilibrium potential ψ^{eq} features water-content-independent quadratic first-invariant and logarithmic third-invariant terms, combined with water-content-dependent linear and quadratic second-invariant terms, where the first increases and the second decreases with water content. The non-equilibrium potential ψ^{neq} features water-content-dependent

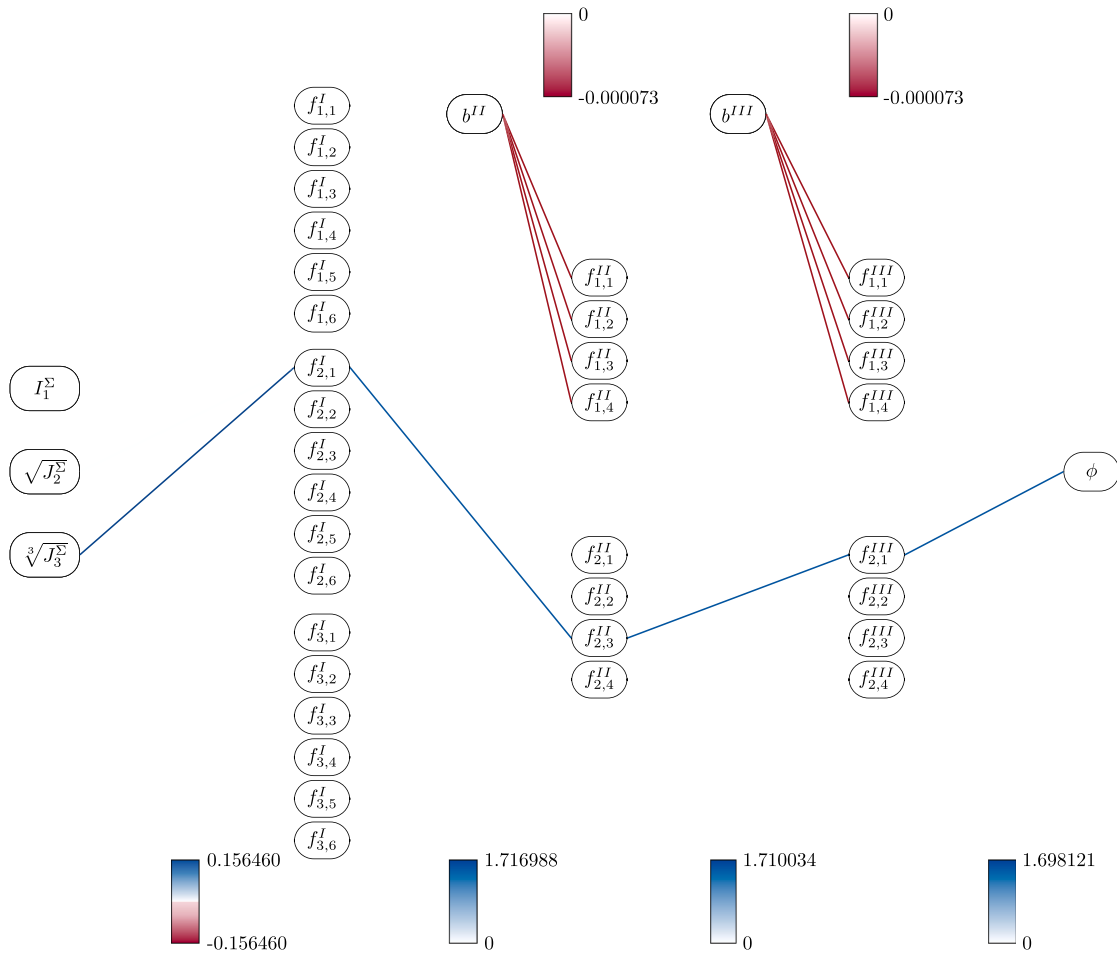


Fig. 13. Model discovery - Trained dissipative network for extrafirm tofu. The colors represent the magnitude of the final weights and biases, with blue indicating positive and red negative values, and white indicating that this path is inactive. The network discovers a four-term potential for extrafirm tofu that depends on complex combinations of the second and third deviatoric invariants J_2^Σ and J_3^Σ of the Mandel stress Σ .

quadratic first-invariant and linear second-invariant terms, where both increase with water content. The dissipative potential ϕ features water-content-dependent exponential second- and third-invariant terms, where the first increases and the second decreases with water content.

Fig. 16 compares the computationally predicted stresses in green and the experimentally recorded stresses in gray. The rows correspond to silken, firm, and extrafirm tofu, and the columns represent distinct loading cases with varying stretch magnitudes and rates. In contrast to Fig. 10, the stress predictions now result from a single model that includes the water content as an explicit input feature. This unified model results in a smaller coefficient of determination with $R^2 = 0.6900$ compared to the single models with $R^2 = 0.7449$, but still captures the underlying trends effectively. Appendix A provides additional details about the errors. Although the model is trained only on synthetic data, it reproduces the major experimental trends: For firm and extrafirm tofu, in the second and third rows, the computationally predicted stresses align closely with the experimentally recorded stresses across all strain rates, especially during loading and unloading. During holding, the predicted stresses follow the experimental averages, but display slightly less nonlinearity than the experimental measurements. For silken tofu, in the first row, the model overestimates the compressive peak stresses by a factor two, which is reflected in its negative R^2 -values. Despite this discrepancy, the silken tofu model captures the overall trends of nonlinear relaxation and decreasing stress magnitude with successive load cycles.

6. Discussion

Tofu is a nonlinear viscoelastic material. Tofu is an exceptional material made up of only soy beans and water. This simple yet tunable model system offers unique conditions to study composition–mechanics relationships. Our mechanical testing of three different tofu types across different water contents, strain magnitudes, and rates in Figs. 8 and 9 reveals several interesting rheological characteristics: First, during the entirely loading path—loading, holding, and unloading—the response of tofu is markedly *nonlinear*. Second, during holding, stresses relax at constant stretch, and the response of tofu is notably *time-dependent*. During

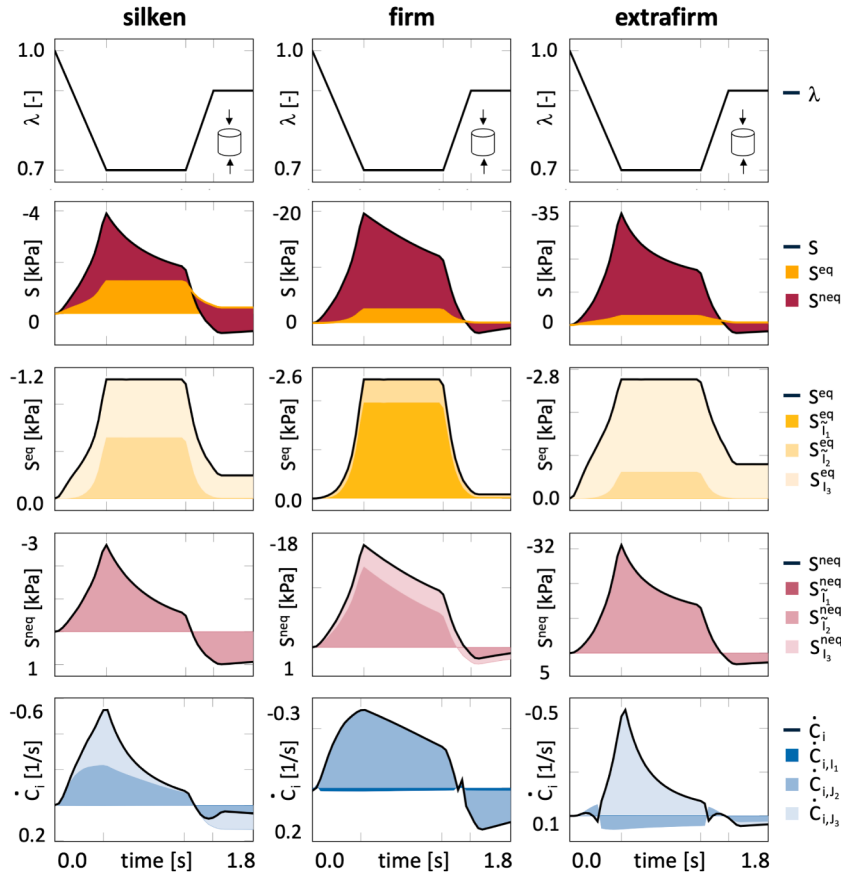


Fig. 14. Model discovery - Predicted stresses and inelastic deformation rates of the three individual tofu models. Stress histories and evolution of the inelastic deformation rate for silken (left), firm (middle), and extrafirm (right) tofu for a strain rate of 60%/s. Applied stretches λ (top); total stress S , equilibrium stress S^{eq} , and non-equilibrium stress S^{neq} (middle), and inelastic deformation rate $\dot{C}_i$ (bottom). Distinct patterns denote contributions from the three invariants $\{I_1^C, I_2^C, I_3^C\}$ and $\{\bar{I}_1^C, \bar{I}_2^C, \bar{I}_3^C\}$ and $\{I_1^Z, I_2^Z, I_3^Z\}$.

consecutive load cycles with similar peak stretches, the recorded peak stresses decrease, which supports the existence of stress relaxation. We conclude that the material behavior of tofu is made up of a history-independent *elastic* part and a history-dependent *inelastic* part. Our observations agree well with rheological tests of different foods throughout the past four decades (Lee et al., 1983; Rao and Steffe, 1992). Specifically, our confined compression setup imposes non-permeable boundaries that prevent water from escaping the sample. This suggests that the observed stress relaxation arises primarily from the viscoelastic behavior of the solid (Budday et al., 2017a), rather than from the fluid-filled matrix (Franceschini et al., 2006). We therefore first adopt the *theory of finite viscoelasticity* (Reese and Govindjee, 1998; Simo, 1987) to explain the observed nonlinear time-dependent response.

Tofu is loading-rate and water-content dependent. We further observe two general trends: stresses increase with increasing loading rate and with decreasing the water content. Fig. 9 of the experimental data and Fig. 10 of the discovered model highlight the loading-rate dependence from left to right and the water-content dependence from top to bottom. The *loading-rate dependence* agrees with our previous findings of fungi-based steak, where the stiffness increased by a factor five when varying the loading rate across four orders of magnitude (Vervenne et al., 2025b). This trend is visible in Figs. 9 and 10, although less pronounced, since our loading rates only vary by a factor three. The *water-content dependence* is more obvious, for example, in the first column, where mean peak stresses are lowest in silken tofu with -3.74 ± 0.86 kPa, followed by firm tofu with -18.91 ± 5.50 kPa, and extrafirm tofu with -42.84 ± 16.22 kPa. These stress ranges agree well with our previous tension, compression, and shear tests of firm and extrafirm tofu with mean stiffnesses of 26.4 ± 4.0 kPa and 27.5 ± 5.6 kPa (St Pierre et al., 2024) and with our rheologically measured storage moduli of 5.7 ± 0.5 kPa and 11.0 ± 1.0 kPa (Dunne et al., 2025). Our observations also agree with previous viscoelastic tests on lab-made tofu with different water contents (Chen et al., 2002), where the elastic, viscous, and relaxation parameters increased notably with decreasing water content (Chen et al., 2005). We can potentially explain these observations with the *theory of finite poroelasticity* (de Boer, 2020; Mow et al., 1984), and propose to supplement the theory of finite viscoelasticity with volume fractions of the solid and fluid phases.

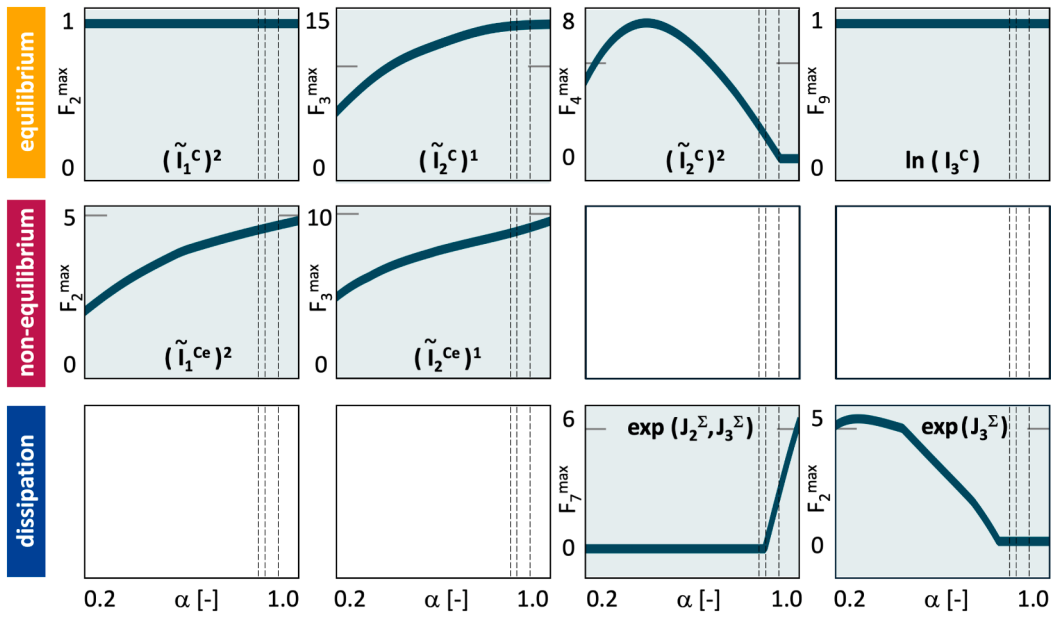


Fig. 15. Model discovery - Water-content dependency of the single unified tofu model. Functional dependency of the equilibrium potential ψ^{eq} (top), non-equilibrium potential ψ^{neq} (middle), and dissipative potential ϕ (bottom) on the water content α . Columns show the discovered dependencies on the individual terms. Dashed lines highlight the water contents of extrafirm, firm, and silken tofu.

Inelastic constitutive neural networks differentiate elastic and inelastic effects. Our experiments suggest that the constitutive behavior of tofu consists of elastic and inelastic parts. Yet, their individual contributions to the overall stresses remain poorly understood. To discover which fraction of the experimentally recorded stress is reversibly elastic and which is not, we utilize an *inelastic constitutive neural network* (Abdolazizi et al., 2024; Holthussen et al., 2024b) made up of feed forward networks for the equilibrium and non-equilibrium potentials combined with a recurrent neural network for the dissipative potential (Holthussen et al., 2024a). The consistent agreement between our computationally predicted and experimentally measured stretch-strain curves, across all three tofu types, for both training and test data in Fig. 10, confirms that our network accurately captures the essential mechanical features of tofu. From the visualization of the equilibrium and non-equilibrium contributions to the overall stress in Fig. 14, we conclude that the network autonomously differentiates elastic and inelastic effects. This differentiation provides an important physical insight: The equilibrium stress increases, from -1.26 kPa for silken to -2.44 kPa for firm and -2.76 kPa for extra firm tofu, as the water content decreases. This trend mirrors the classical constitutive theory for fluids, where a perfect fluid at rest cannot support shear stresses. In other words, its rheological model lacks a Hookean spring and only consists of a Maxwell element. Several rheological models with a single Maxwell model, a generalized Maxwell model, and a four-element Maxwell model have been proposed to explain the stress relaxation of tofu with different water contents (Chen et al., 2002). Our automated discovery process captures these Maxwell-type characteristics (Holthussen et al., 2024a), and naturally reproduces the gradual transition from a viscoelastic solid to a nearly fluid-like response.

The elastic behavior of tofu is dominated by the second isochoric invariant. To characterize the isotropic elastic behavior of tofu, we have parameterized its equilibrium and non-equilibrium potentials in terms of *two isochoric invariants* and *one volumetric invariant* of the total and elastic right Cauchy Green tensors (Holzapfel, 2000), $\{\bar{I}_1^C, \bar{I}_2^C, \bar{I}_3^C\}$ and $\{\bar{I}_1^{Ce}, \bar{I}_2^{Ce}, \bar{I}_3^{Ce}\}$. Yet, their individual contributions to the two elastic potentials are unknown. To discover which invariants drive the equilibrium and non-equilibrium responses, we consult our elastic networks in Fig. 3 to identify the most relevant contributions out of $2^9 = 512$ possible combinations of terms (Holthussen et al., 2024b,a). Despite quantitative differences, the discovered models for all three tofu types, silken, firm, and extrafirm, share several interesting trends: Across all tofu types, both elastic potentials, equilibrium and non-equilibrium, are dominated by the *second isochoric invariants*, $\bar{I}_2^C$ or $\bar{I}_2^{Ce}$, which measure the isochoric area change. The equilibrium potential ψ^{eq} features a quadratic second-invariant term for silken and extrafirm tofu, and a linear second-invariant term for firm tofu. The non-equilibrium potential ψ^{neq} features a linear second-invariant term for all three tofu types, and an additional quadratic second-invariant term for extrafirm tofu. The complete *absence of exponential terms*—in both potentials and across all three tofu types—is in line with our general understanding that exponential terms represent the strain-stiffening of elastin and collagen fibers in soft biological tissues (Holzapfel, 2000), which are not present in tofu. The *dominance of the second invariant* (Kuhl and Goriely, 2024) is in line with previously discovered models for human brain (Linka et al., 2023a) and plant-based meat (Linka and Kuhl, 2024), and suggests that these hydrated soft materials share common rheological trends. Our results parallel classical constitutive modeling, where a single constitutive model can describe a family of materials simply by adjusting its parameter values. For example, in previous studies on lab-made tofu, a four-element Burgers model, a rheological model with two parallel dashpot-springs,

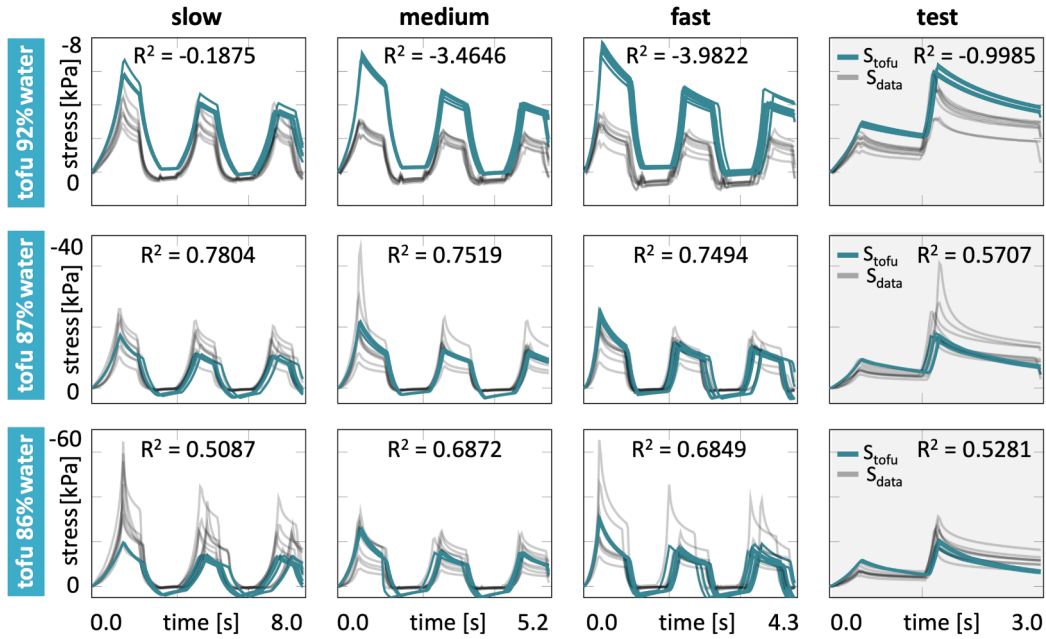


Fig. 16. **Model discovery - Experimentally recorded and computationally predicted stresses of the single unified tofu model.** Stress histories $S(t)$ for silken (top), firm (middle), and extrafirm (bottom) tofu. The first three columns show the triple-compression tests used as training data. The last column shows double-compression tests used as test data. Experimental data are plotted in gray, model predictions in green; R^2 values indicate the goodness of fit.

was able to explain the rheological behavior of tofu with varying soymilk concentrations of $\{5\%, 6\%, 7\%, 8\%, 9\%\}$, simply by varying the elastic and viscous parameter values and the relaxation time (Chen et al., 2005). Our discovered models naturally reflect this principle: Tofu with varying solid volume fractions of $\{8\%, 13\%, 14\%\}$ results in varying parameterizations of the model, but not in fundamentally different functional forms.

The dissipative behavior of tofu is driven by the second and third deviatoric invariants. To characterize the dissipative behavior of tofu, we have parameterized its dissipative potentials in terms of *one volumetric invariant* and *two deviatoric invariants* of the Mandel stress (Holthusen et al., 2024b), $\{I_1^S, J_2^S, J_3^S\}$. Yet, their individual contributions to the dissipative potential are unknown. To discover which invariants drive the dissipative responses, we consult our dissipative network in Fig. 4. Strikingly, across all three tofu types, the discovered dissipative potentials ϕ feature a single term that contains weighted powers of the second and third deviatoric invariants, J_2^S and J_3^S , embedded in two nested exponential functions. None of the discovered potentials depends on the volumetric invariant I_1^S , and none of them features any biases that would result in additive contributions. These observations confirm that our dissipative network autonomously eliminates irrelevant terms, and discovers sparse and interpretable models (McCulloch et al., 2024): The first stress invariant I_1^S measures the hydrostatic stress, and its absence in the dissipative potential indicates a *pressure insensitivity*, unlike in the classical Drucker-Prager model (Drucker and Prager, 1952). The second stress invariant J_2^S measures the magnitude of shear, and its presence in the dissipative potential indicates a *shear sensitivity*, like in the classical von Mises model (von Mises, 1913). The third stress invariant J_3^S measures the mode of shear via the ratio between the principal stresses, and its presence in the dissipative potential indicates a *stress ratio sensitivity*, like in the classical Tresca (Tresca, 1864) and Mohr-Coulomb models (Coulomb, 1776; Mohr, 1900). While simple gels are often modeled as von-Mises type viscoelastic materials that depend exclusively on the second invariant, our dissipative network discovers a Lode-angle sensitive behavior (Lode, 1925) that depends on both the second and third invariants and distinguishes between tension-shear, pure shear, and compression-shear of the same shear magnitude (Bruhns, 2024). These considerations underscore that our discovered models not only accurately reproduce the experimental behavior of tofu, but also offer interpretable, data-driven insight into its nonlinear viscoelastic behavior.

The mechanics of tofu depend nonlinearly on its water content. Tofu is one of the simplest man-made foods; it consists of only two ingredients: soy beans and water. We can easily tune its consistency by adjusting its processing parameters: the higher the pressure and the longer its duration, the lower the water content (Shurtleff and Aoyagi, 1975). Making and testing tofu with varying water content provides valuable insight into how we can modulate the rheology of food simply by adding or removing water (Chen et al., 2005). From our three individually discovered tofu models, we conclude that the nonlinear and time-dependent behavior of tofu is highly sensitive to its water content. Yet, the explicit impact of water on the individual tofu models is unknown. A naive ad hoc explanation would be to consult the biphasic theory (Mow et al., 1984), adopt the concept of volume fractions (de Boer and Ehlers, 1990), associate the equilibrium behavior with the solid and the non-equilibrium behavior with the fluid, and weigh both with their individual volume fractions (Ehlers, 1993), $S = v_s \cdot S^{\text{eq}} + v_f \cdot S^{\text{neq}}$. In Table 2, the volume fractions of the fluid and

solid are $v_f = \{0.92, 0.87, 0.86\}$ and $v_s = 1 - v_f = \{0.08, 0.13, 0.14\}$, for silken, firm, and extrafirm tofu. We can easily see that the *biphasic theory alone fails* to explain the observed stiffness variations: For example, the peak stresses in the first column of Fig. 9, $S_{\text{peak}} = \{-3.74 \pm 0.86 \text{ MPa}, -18.91 \pm 5.50 \text{ MPa}, -42.84 \pm 16.22 \text{ MPa}\}$ vary by more than an order of magnitude across the three tofu types, while the volume fractions vary by only six percent. This agrees with our previous observations of firm and extrafirm tofu for which the stiffness, storage, and loss moduli varied by a factor two, while the water contents differed by only one percent (Dunne et al., 2025). We conclude that volume-fraction-based scaling might be appropriate for linear mixtures with low-to-moderate water contents (de Boer, 2020), but a more complex nonlinear water-content dependence is critical to accurately model highly hydrated soft materials like tofu.

Feature network modeling reveals correlations between water content and mechanics. Feature networks learn nonlinear mappings from raw input features to weights that parameterize or modulate terms of another network model. To explore how water content modulates the mechanics of tofu, we introduce a feature network that takes the scalar-valued water content α as input, and outputs the weights for each node of the three mechanical networks. Our results in Fig. 15 confirm that the individual terms of the equilibrium, non-equilibrium, and dissipative potentials display different types of water-content dependency: Two terms are entirely independent of the water content, while four increase and two decrease with increasing water content. We conclude that incorporating the *water content as a feature* can successfully introduce material-specific variability into a unified constitutive model. The close agreement between the computationally predicted and experimental recorded stresses in Fig. 16, especially for firm and extrafirm tofu, confirms that the feature network can indeed effectively modulate the mechanical response. Comparisons with other hydrated soft biological tissues suggest that neither a purely viscoelastic (Budday et al., 2017b), nor a purely poroelastic model (Franceschini et al., 2006) can accurately capture this characteristic behavior, but that a poroviscoelastic model is needed instead (Kainz et al., 2023). Indeed, hydrated soft solids including the cornea, meniscus, intervertebral disk, articular cartilage, and hydrogels have long been known to display a complex nonlinear poroviscoelastic behavior (Holmes and Mow, 1990). Our feature network provides the inherent flexibility to model these complex rheologies. It naturally combines microstructural information and macrostructural experiments into a data driven, scalable approach that represents families of hydrated soft materials within a single feature-weighted model. More generally, our results suggest that feature network modeling is a flexible approach to integrate multimodal data and discover how various multiphysics inputs—water content, density, temperature, or fiber angles—modulate the mechanics and physics of soft matter systems.

Limitations. We note a few limitations of our study: First, we have assumed that the *inelastic deformation* is isotropic and *volume preserving*, and that volumetric changes are associated with the elastic deformation only (Holzapfel, 2000; Reese and Govindjee, 1998; Simo, 1987). In a follow up study, we are currently performing both confined and unconfined compression tests on tofu to explore to which extent this assumption is justified. Second, our currently findings are restricted the *compressive loading* only. Although compression is the dominant loading mode during chewing and most relevant to our sensory perception of texture (Vervenne et al., 2025a), we have previously tested firm and extrafirm tofu in compression and shear (Dunne et al., 2025), and in tension, compression, and shear (St Pierre et al., 2024); yet, testing silken tofu in tension remains challenging. Third, we have only tested over a *small range of loading rates*. Testing food at loading rates across several orders of magnitude and observing fluid flow can help delineate poroelastic and viscoelastic phenomena (Vervenne et al., 2025a). Fourth, we have assumed that water in tofu exists only in a single state (de Boer, 2020), freely moving. In reality, tofu is a hydrated soy protein gel that contains water in three distinct states, *bound*, *immobilized*, and *free* (Yuan et al., 2025), where bound and immobilized water modulate the elastic response, while free water modulates the inelastic response (Lucey, 2002). A next logical step would be to characterize the volume fractions of bound, immobilized, and free water (Kainz et al., 2023), then integrate all three volume fractions as distinct features of the network, and discover how each contribution modulates the behavior of tofu, and of hydrated soft solids in general. Finally—and this is probably more a philosophical question than a true limitation—in line with the literature (Brunton et al., 2016; Flaschel et al., 2021; Linka and Kuhl, 2023), we have used the terminology *discover* in the context of discovering sparse, interpretable functional representations of potentials within a pre-specified constitutive ansatz. Following the common approach, we reverse-engineer this pre-defined ansatz from popular constitutive terms and functions and generalize it to a library of activation functions that satisfy thermodynamics and physical constraints by design (Kalina et al., 2023; Klein et al., 2023; Masi et al., 2021).

Conclusion

Tofu is a hydrated protein gel composed only of soybeans and water. It provides a minimal model system to explore how material composition governs the mechanics and physics of food. A series of compression tests of three tofu types reveals its strong nonlinearity, its pronounced stress relaxation, and its more than ten-fold increase in peak stress with only six percent decrease in water content. Using automated model discovery, we discover inelastic constitutive models and parameters that accurately explain both our training and test data of silken, firm, and extrafirm tofu across three loading magnitudes and five loading rates. The discovered models autonomously separate elastic and inelastic behavior and reveal that the elastic response is dominated by the second isochoric invariant, while the inelastic response is governed by a combination of the second and third deviatoric invariants. Water content modulates the magnitude of these mechanisms, but not their functional form. This motivates the integration of all results into a single water-content weighted model for different tofu types. Our results position tofu as a simple quantitative benchmark for nonlinear poroviscoelastic solids and demonstrate how physics-informed machine learning can expose the underlying constitutive structure of hydrated soft materials. Knowing the precise rheology of tofu is essential to reliably reproduce, engineer, and design the texture of alternative protein products in the benefit of human and planetary health. More broadly, our results suggest that combined

feature–mechanics networks can uncover complex multiphysics behavior in hydrated soft materials, including biopolymer hydrogels, engineered protein networks, and hydrated porous composites Figs. B.20, B.21.

CRedit authorship contribution statement

Birte Boes: Writing – review & editing, Writing – original draft, Validation, Software, Methodology, Investigation, Formal analysis, Data curation, Conceptualization; **Jaen-Willem Simon:** Writing – review & editing, Funding acquisition; **Hagen Holthusen:** Writing – review & editing, Software, Methodology, Formal analysis, Conceptualization; **Ellen Kuhl:** Writing – review & editing, Resources, Conceptualization.

Data availability

All data, raw data, analyzed data, code to analyze the data, and results will be made available on our GitHub repository upon publication. Our source code and examples are available at <https://doi.org/10.5281/zenodo.16993236>.

Declaration of competing interest

The authors declare that they have no known competing financial interests or personal relationships that could have appeared to influence the work reported in this paper.

Acknowledgment

We thank Sofie Danowski for assisting with the hyperparameter optimization as part to this study. This work was supported by the DFG Project SI 1959/12-1 466117814, by the DFG TRR 280 417002380, by the publication prize of the University of Wuppertal, by the Stanford Bio-X Snack Grant 2025, by the NSF CMMI Award 2320933, and by the ERC Advanced Grant 101141626.

Appendix A. Errors of discovered individual and unified models

To quantify the performance of our discovered models, we calculate the error between the stress of the model S and the data $\hat{S}$ across all data points n_{data} . In addition to the coefficients of determination, $R^2 = 1 - \sum_{i=1}^m [S_i - \hat{S}_i]^2 / \sum_{i=1}^m [S_i - \bar{S}]^2$ in Figs. 10 and 16, we report the mean absolute error, $\text{MAE} = \sum_{m=1}^{n_{\text{data}}} |S_m - \hat{S}_m| / n_{\text{data}}$, the root mean squared error, $\text{RMSE} = [\sum_{m=1}^{n_{\text{data}}} [S_m - \hat{S}_m]^2 / n_{\text{data}}]^{1/2}$, normalized root mean squared error, $\text{NRMSE} = \text{RMSE} / \max\{\hat{S}\}$, where $\max\{\hat{S}\}$ denotes the maximum experimental stress, and the relative L^2 error, $L^2 = [\sum_{i=1}^m [S_i - \hat{S}_i]^2]^{1/2} / [\sum_{i=1}^m S_i^2]^{1/2}$. Table A summarizes the errors for the discovered individual tofu models in Fig. 10 and for the discovered unified tofu model in Fig. 16.

Table A.2

Errors of discovered models. Mean absolute error MAE, root mean squared error RMSE, normalized root mean squared error NRMSE, and relative L^2 error for the discovered individual tofu models in Fig. 10, left, and for the discovered unified tofu model in Fig. 16, right.

individual	error	slow	medium	fast	test	unified	error	slow	medium	fast	test
silken	RMSE	0.633	0.298	0.460	1.206	silken	RMSE	1.473	2.318	2.643	1.633
	MAE	0.402	0.223	0.322	0.970		MAE	1.182	1.893	2.150	1.435
	NRMSE	0.118	0.095	0.133	0.205		NRMSE	0.274	0.742	0.762	0.277
	L^2	0.381	0.228	0.344	0.479		L^2	0.886	1.780	1.973	0.648
firm	RMSE	2.911	2.891	2.476	4.203	firm	RMSE	2.867	3.140	3.056	3.880
	MAE	1.871	1.684	1.625	2.707		MAE	2.131	2.231	2.321	2.728
	NRMSE	0.112	0.063	0.113	0.103		NRMSE	0.111	0.068	0.141	0.095
	L^2	0.361	0.352	0.836	0.496		L^2	0.356	0.382	0.373	0.355
extrafirm	RMSE	5.746	4.877	5.200	4.644	extrafirm	RMSE	7.632	4.030	5.405	4.080
	MAE	3.400	3.386	3.659	3.922		MAE	4.982	3.290	3.642	3.297
	NRMSE	0.089	0.147	0.080	0.148		NRMSE	0.118	0.122	0.083	0.131
	L^2	0.402	0.521	0.419	0.411		L^2	0.534	0.431	0.436	0.361

To interpret the sources of error, Fig. A.17 highlights the training performance on individual curves extracted from the datasets. We incrementally increase the amount of training data, from slow loading, to slow and medium, to slow, medium and fast, from top to bottom. Fig. A.18 shows the model prediction vs. ground truth data for these three training scenarios. For training on a single curve, top row, the fit is nearly perfect, and it is evident that the remaining error originates primarily from the data themselves. The performance on the training data with $R^2 = 0.9543$ is notably better than the performance on the test data with $R^2 = 0.7287$ and $R^2 = 0.8416$. For training on two curves, middle row, the fit of the second curve improves significantly, from $R^2 = 0.7287$ to $R^2 = 0.9495$, while the first shows a minor reduction in fit quality. For training on three curves, bottom row, the fit of the third curve improves, from $R^2 = 0.8416$ to $R^2 = 0.9457$, while the first and second curve fits remain largely unchanged. The plain testing performance, last column, remains weaker than the slow, medium, and fast columns, since the model is evaluated on previously unseen loading magnitudes and rates.

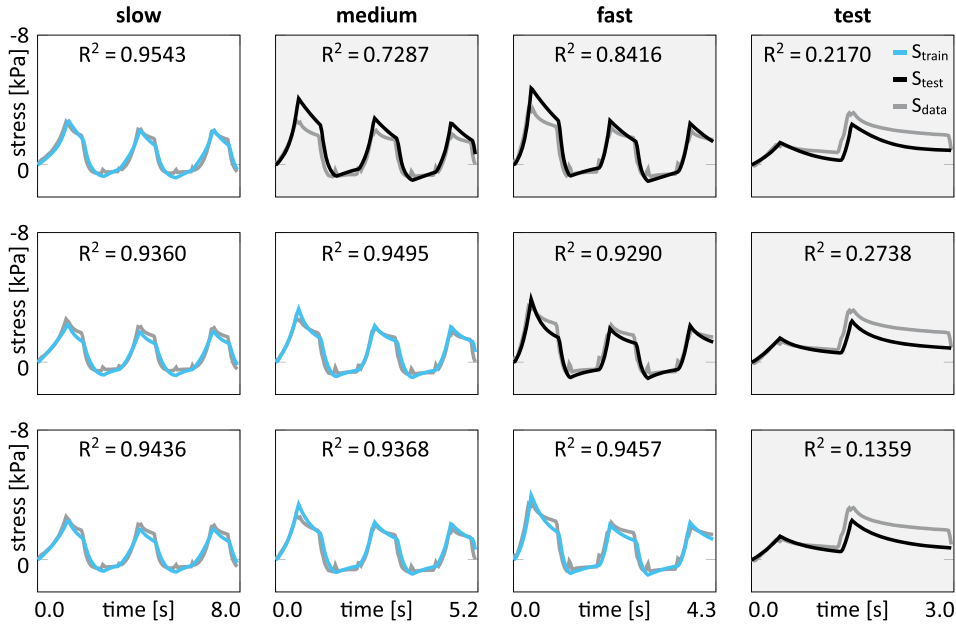


Fig. A.17. Error measures for varying training sets. Single stress histories $S(t)$ for silken tofu trained on the slow strain rate (top), slow and medium rates (middle), and slow, medium, and fast rates (bottom). Predictions during training shown in blue, testing in black.

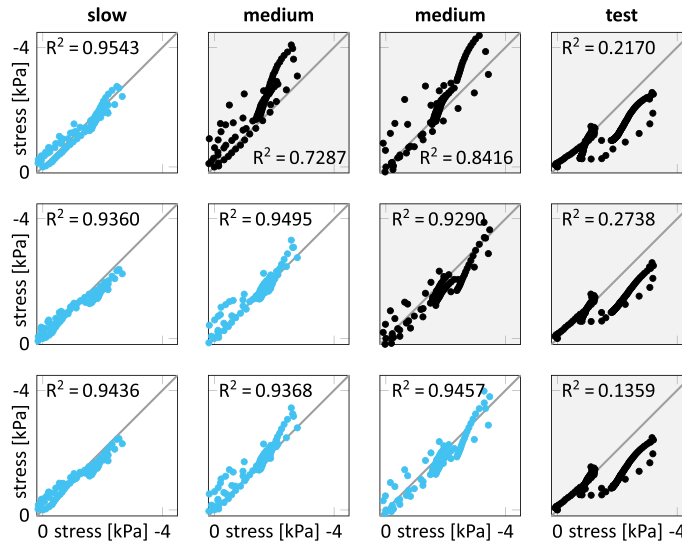


Fig. A.18. Error measures for varying training sets. Ground truth vs. model predictions for silken tofu trained on the slow strain rate (top), slow and medium rates (middle), and slow, medium, and fast rates (bottom), corresponding to Fig. A.17. Predictions during training shown in blue, testing in black.

Interestingly, the predictions remain accurate for slower loading rates, while larger deviations occur at higher loading rates. This low predictive performance indicates that the current training data do not sufficiently capture the full range of the material behavior under different rates. Finally, the reported R^2 values in Figs. 10 and Figs. 16 and in Table A additionally reflect the fact that here, a single model represents the collective response of $n = 8$ to $n = 10$ individual curves, with a large variance between individual curves. Taken together, from Figs. A.17 and A.18, we conclude that the R^2 value for training on single curves is on the order of 0.95%, and that the remaining error of 5% reflects measurement inconsistencies in the training data. The R^2 values for testing on single curves vary between 0.73% and 0.93%, and the remaining error reflects the collective effects of measurement inconsistencies and prediction errors to unseen loading rates. This error increases further when making predictions into unseen loading magnitudes and rates. The

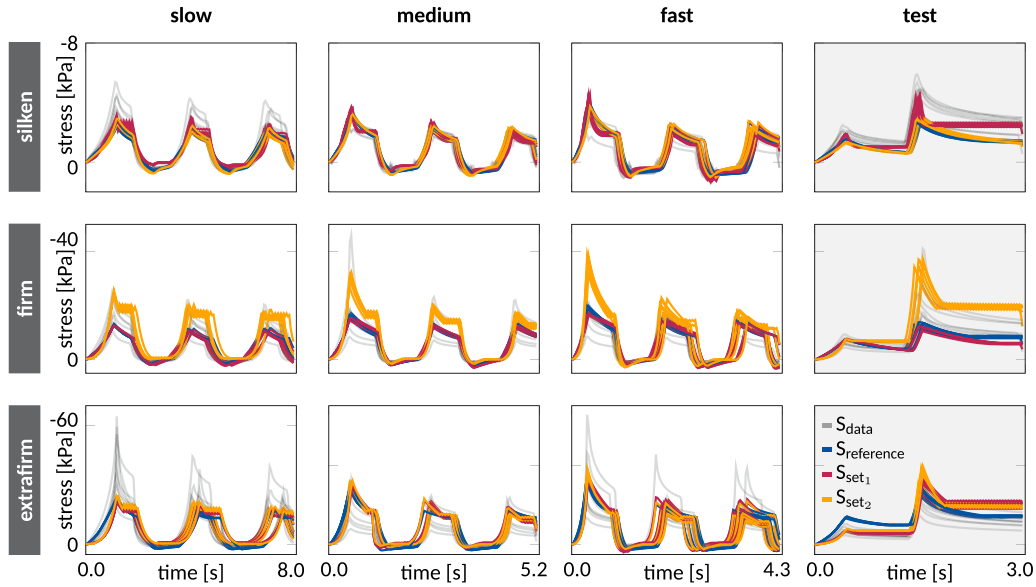


Fig. B.19. Preliminary sensitivity analysis for varying network architectures. Stress histories $S(t)$ for silken (top), firm (middle), and extrafirm (bottom) tofu. Experimental data are plotted in gray; reference data from Fig. 10 with $\{18, 8, 8\}$ neurons per layer are plotted in blue; predictions with $\{30, 16, 12\}$ and $\{4, 4, 2\}$ neurons per layer are shown in red as set 1 and in orange as set 2.

seemingly low R^2 values in Figs. 10 and Figs. 16 and in Table A reflect a *collective result of measurement inconsistencies* in single curves, a *high variance* across different curves, and *prediction errors* into previously unseen regimes.

Appendix B. Preliminary sensitivity analyses

Prior to generating the main results of this study, we performed a systematic series of sensitivity analyses. Here we showcase three exemplary studies that, amongst many others, have guided our selection of network architectures, activation functions, and hyperparameters.

Network architectures. Fig. B.19 summarizes our preliminary sensitivity analysis for varying network architectures. Specifically, we study a finer network with $\{30, 16, 12\}$ neurons per layer and a coarser network with $\{4, 4, 2\}$ neurons per layer, compared to our reference network with $\{18, 8, 8\}$ neurons per layer that achieved a goodness of fit of $R^2 = 0.7449$. The finer $\{30, 16, 12\}$ network results in a reduced goodness of fit of $R^2 = 0.7085$, associated with the stress predictions in red. The coarser $\{4, 4, 2\}$ network results in a similarly reduced goodness of fit of $R^2 = 0.7091$, associated with the stress predictions in orange. Interestingly, both network architectures lead to numerical instabilities, most visible in the results for firm and extrafirm tofu. These instabilities could be a result of time-step sensitivity related to the explicit time integration and might require a reduced time step size or more steps in the staggered training approach.

Activation functions.

Fig. B.20 summarizes our preliminary sensitivity analysis for varying activation functions. Specifically, we vary the activation functions in the first layer of the dissipative network by fixing the powers to $p = 1$ and $p = 2$, instead of leaving p as a trainable parameter which achieved a goodness of fit of $R^2 = 0.7449$. Fixing the powers to $p = 1$ and $p = 2$ reveals similar trends and a notable reduction of the goodness of fit to $R^2 = 0.7212$ for $p = 1$, associated with the red curves, and even further to $R^2 = 0.6586$ for $p = 2$, associated with the orange curves. For the latter, non-linear effects almost vanish entirely for firm and extrafirm tofu. Overall, our selection of activation functions is guided by thermodynamical considerations, and is in line with previous elastic (Linka and Kuhl, 2023; Linka et al., 2023a,b) and dissipative (Holthusen et al., 2024b, 2025a,b) networks similar to Figs. 3 and 4 that have demonstrated high robustness and reproducibility.

Regularization parameters.

Fig. B.21 summarizes our preliminary sensitivity analysis for varying regularization coefficients α_1 and α_2 that fine-tune the level of L_1 and L_2 regularization. Specifically, we set both hyperparameters a magnitude higher and lower than our selected reference values of $\alpha_1 = \alpha_2 = 0.0001$ that achieved a goodness of fit of $R^2 = 0.7449$. Higher regularization coefficients of $\alpha_1 = \alpha_2 = 0.001$ result in a reduced goodness of fit of $R^2 = 0.6556$ with stress predictions that display almost no nonlinearity for firm and extrafirm tofu, shown in red. Lower regularization coefficients of $\alpha_1 = \alpha_2 = 0.00001$ result in a similar goodness of fit of $R^2 = 0.7510$ as the reference model, with stress predictions, shown in orange, that predict similar features as the reference model, shown in blue.

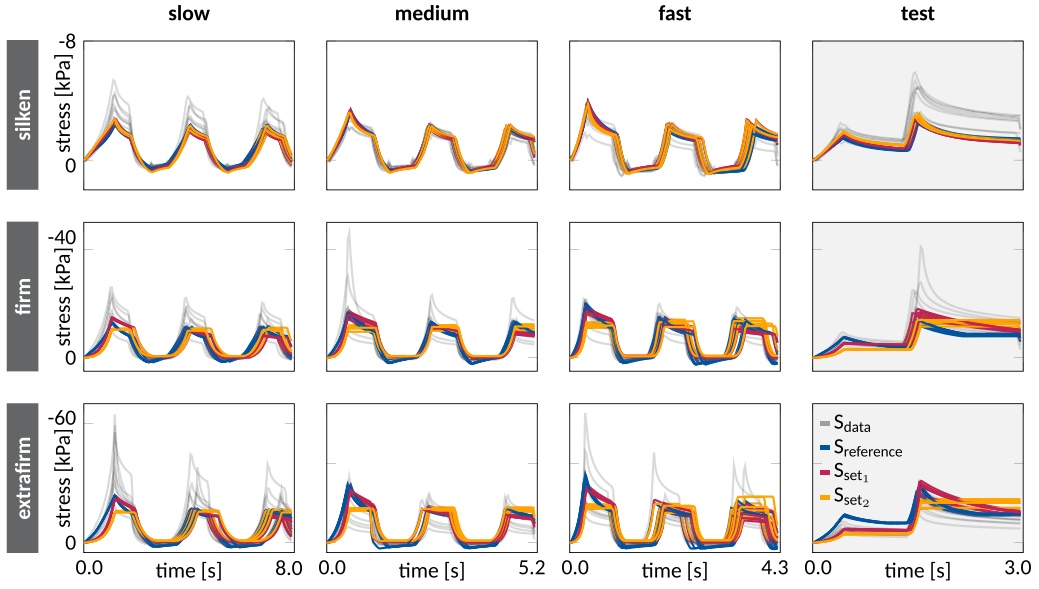


Fig. B.20. Preliminary sensitivity analysis for varying activation functions. Stress histories $S(t)$ for silken (top), firm (middle), and extrafirm (bottom) tofu. Experimental data are plotted in gray; reference data from Figure 10 with trainable parameter p are plotted in blue; predictions with fixed parameters $p = 1$ and $p = 2$ are shown in red as set 1 and in orange as set 2.

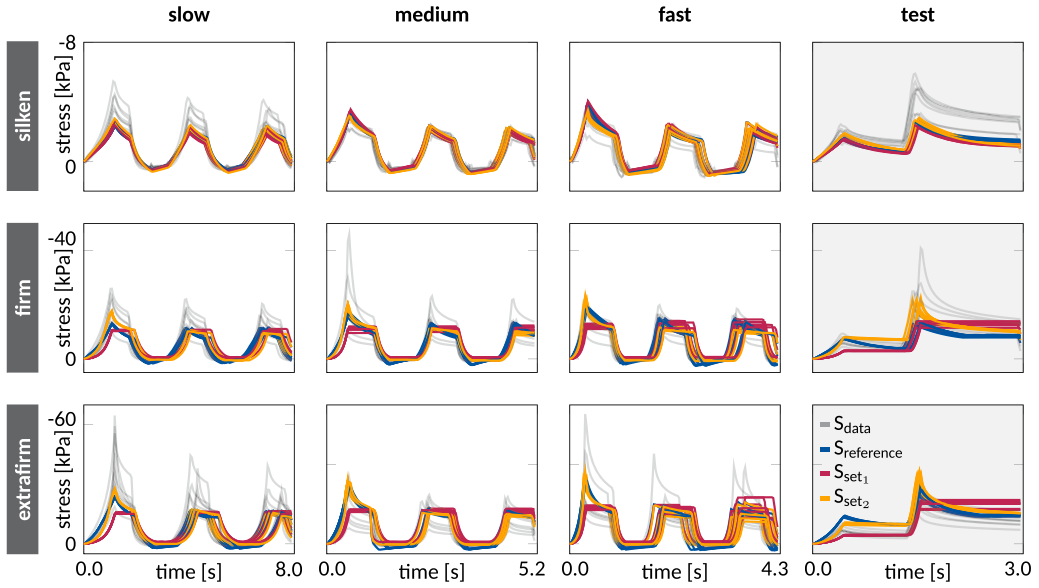


Fig. B.21. Preliminary sensitivity analysis for varying regularization parameters. Stress histories $S(t)$ for silken (top), firm (middle), and extrafirm (bottom) tofu. Experimental data are plotted in gray; reference data from Figure 10 with regularization parameters $\alpha_1 = \alpha_2 = 0.0001$ are plotted in blue; predictions with $\alpha_1 = \alpha_2 = 0.001$ and $\alpha_1 = \alpha_2 = 0.00001$ are shown in red as set 1 and in orange as set 2.

Taken together, these results illustrate the sensitivity of network training with respect to the network architecture, activation functions, and regularization parameters. While we only show a few selected cases here, we have performed a wide variety of preliminary studies and recommend systematic sensitivity analyses before training the final model.

References

- Abdolazizi, K., Linka, K., Cyron, C., 2024. Viscoelastic constitutive artificial neural networks (VCANNS) – a framework for data-driven anisotropic nonlinear finite viscoelasticity. *J. Comput. Phys.* 499, 112704.
- Amos, B., Xu, L., Kolter, J.Z., 2017. Input convex neural networks. In: Precup, D., Teh, Y.W. (Eds.), *Proceedings of the 34th International Conference on Machine Learning*. Vol. 70 of *PMLR*, pp. 146–155.
- As'ad, F., Avery, P., Farhat, C., 2022. A mechanics-informed artificial neural network approach in data-driven constitutive modeling. *Int. J. Numer. Methods Eng.* 123 (12), 2738–2759.
- de Boer, R., 2020. *Theory of Porous Media: highlights in Historical Development and Current State*. Springer, Berlin.
- de Boer, R., Ehlers, W., 1990. The development of the concept of effective stresses. *Acta Mech.* 83, 77–92.
- Boes, B., Simon, J., Holthusen, H., 2024. Accounting for plasticity: an extension of inelastic constitutive artificial neural networks. *Eur. J. Mech. A/Solids* 117, 105998.
- Bourne, M., 1968. Texture profile of ripening pears. *J. Food Sci.* 33 (2), 223–226.
- Bruhns, O., 2024. Some remarks on lode angle and lode parameter. *Proc. Appl. Math. Mech.* 24, e202400004.
- Brunton, S., Proctor, J., Kutz, J., 2016. Discovering governing equations from data by sparse identification of nonlinear dynamical systems. *Proc. Nat. Acad. Sci.* 113, 3932–3937.
- Budday, S., Oveart, T., Holzapfel, G., Steinmann, P., Kuhl, E., 2020. Fifty shades of brain: a review on the mechanical testing and modeling of brain tissue. *Arch. Comp. Meth. Eng.* 27, 1187–1230.
- Budday, S., Sommer, G., Haybaeck, J., Steinmann, P., Holzapfel, G., Kuhl, E., 2017a. Rheological characterization of human brain tissue. *Acta Biomat.* 60, 315–329.
- Budday, S., Sommer, G., Holzapfel, G., Steinmann, P., Kuhl, E., 2017b. Viscoelastic parameter identification of human brain tissue. *J. Mech. Beh. Biomed. Mat.* 74, 463–476.
- Chen, C., Hsieh, J., Kuo, M., 2023. Insight into the processing, gelation and functional components of tofu: a review. *Processes* 11 (1), 202.
- Chen, Y., Shimizu, N., Kimura, T., 2002. Texture evaluation of tofu (soybean curd) by viscoelastic tests. *J. Japanese Soc. Agric. Mach.* 64 (6), 137–144.
- Chen, Y., Shimizu, N., Kimura, T., 2005. The viscoelastic properties of soybean curd (tofu) as affected by soymilk concentration and type of coagulant. *Int. J. Food Sci. Tech.* 40, 385–390.
- Coulomb, C., 1776. Essai sur une application des règles des maximis et minimis à quelques problèmes de statique relatifs à l'architecture. *Mémoires de l'Académie R. Sci.* 7, 343–387.
- Datta, B., Buehler, M., Chow, Y., Gligoric, K., Jurafsky, D., Kaplan, D., Ledesma-Amaro, R., Del Missier, G., Neidhardt, L., Pichara, K., Sanchez-Lengeling, B., Schlagen, M., St Pierre, S., Tagkopoulos, I., Thomas, A., Watson, N., Kuhl, E., 2025. AI for sustainable future foods. *arXiv:2509.21556*
- Drucker, D., Prager, W., 1952. Soil mechanics and plastic analysis or limit design. *Quart. Appl. Math.* 10, 157–165.
- Dunne, R., Darwin, E., Perez Medina, V., Levenston, M., St Pierre, S., Kuhl, E., 2025. Texture profile analysis and rheology of plant-based and animal meat. *Food Res. Int.* 205, 115876.
- Ehlers, W., 1993. Constitutive equations for granular materials in geomechanical context. In: *Continuum Mechanics in Environmental Sciences and Geophysics*. Springer, pp. 313–402.
- Flaschel, M., Kumar, S., De Lorenzis, L., 2021. Unsupervised discovery of interpretable hyperelastic constitutive laws. *Comput. Methods Appl. Mech. Eng.* 381, 113852.
- Franceschini, G., Bigoni, D., Regitnig, P., Holzapfel, G., 2006. Brain tissue deforms similarly to filled elastomers and follows consolidation theory. *J. Mech. Phys. Solids* 54, 2592–2620.
- Friedman, H., Whitney, J., Szczesniak, A., 1963. The texturometer – a new instrument for objective texture measurement. *J. Food Sci.* 28 (4), 390–396.
- Holmes, M., Mow, V., 1990. The nonlinear characteristics of soft gels and hydrated connective tissues in ultrafiltration. *J. Biomech.* 23, 1145–1156.
- Holthusen, H., Brepols, T., Linka, K., Kuhl, E., 2025a. Automated model discovery for tensional homeostasis: constitutive machine learning in growth and remodeling. *Comput. Biol. Med.* 186, 109691.
- Holthusen, H., Kuhl, E., 2026. A complement to neural networks for anisotropic inelasticity at finite strains. *Comput. Methods Appl. Mech. Eng.* 450, 118612.
- Holthusen, H., Lamm, L., Brepols, T., Reese, S., Kuhl, E., 2024a. Polyconvex inelastic constitutive artificial neural networks. *PAMM* 24 (3), e202400032.
- Holthusen, H., Lamm, L., Brepols, T., Reese, S., Kuhl, E., 2024b. Theory and implementation of inelastic constitutive artificial neural networks. *Comput. Methods Appl. Mech. Eng.* 428, 117063.
- Holthusen, H., Linka, K., Kuhl, E., Brepols, T., 2025b. A generalized dual potential for inelastic constitutive artificial neural networks: a jax implementation at finite strains. *J. Mech. Phys. Solids* 206, 106337.
- Holzapfel, G., 2000. *Nonlinear Solid Mechanics*. John Wiley & Sons, New York.
- Hou, J., Chen, X., Wu, T., Kuhl, E., Wang, X., 2024. Automated data-driven discovery of material models based on symbolic regression: a case study on the human brain cortex. *Acta Biomater.* 188, 276–296.
- Kainz, M., Greiner, A., Hinrichsen, J., Kolb, D., Comellas, Steinmann, P., Budday, S., Terzano, M., Holzapfel, G., 2023. Poro-viscoelastic material parameter identification of brain-tissue mimicking hydrogels. *Front. Bioeng. Biotech.* 11, 1143304.
- Kalina, K., Gebhart, P., Brummund, J., Linden, L., Sun, W., Kästner, M., 2024. Neural network-based multiscale modeling of finite strain magneto-elasticity with relaxed convexity criteria. *Comput. Methods Appl. Mech. Eng.* 421, 116739.
- Kalina, K., Linden, L., Brummund, J., Kästner, M., 2023. Fe-ann: an efficient data-driven multiscale approach based on physics-constrained neural networks and automated data mining. *Comput. Mech.* 71 (5), 827–851.
- Klein, D., Roth, F., Valizadeh, I., Weeger, O., 2023. Parametrized polyconvex hyperelasticity with physics-augmented neural networks. *Data Centric Eng.* 4, e25.
- Kohyama, K., Sano, Y., Doi, E., 1995. Rheological characteristics and gelation mechanism of tofu (soybean curd). *J. Agric. Food Chem.* 43 (7), 1808–1812.
- Kuhl, E., 2025. AI for food: accelerating and democratizing discovery and innovation. *npj Sci. Food* 9, 82.
- Kuhl, E., Goriely, A., 2024. I too love i2: a new class of hyperelastic isotropic incompressible models based solely on the second invariant. *J. Mech. Phys. Solids* 188, 105670.
- Lee, Y., Rosenau, J., Peleg, M., 1983. Rheological characterization of tofu. *J. Texture Stud.* 14 (2), 143–154.
- Li, T., Rui, X., Li, W., Chen, X., Jiang, M., Dong, M., 2014. Water distribution in tofu and application of t2 relaxation measurements in determination of tofu's water-holding capacity. *J. Agric. Food Chem.* 62, 8594–8601.
- Linka, K., Hillgärtner, M., Abdolazizi, K., Aydin, R., Itskov, M., Cyron, C., 2021. Constitutive artificial neural networks: a fast and general approach to predictive data-driven constitutive modeling by deep learning. *J. Comput. Phys.* 429, 110010.
- Linka, K., Kuhl, E., 2023. A new family of constitutive artificial neural networks towards automated model discovery. *Comput. Methods Appl. Mech. Eng.* 403, 115731.
- Linka, K., Kuhl, E., 2024. Best-in-class modeling: a novel strategy to discover constitutive models for soft matter systems. *Extreme Mech. Lett.* 70, 102181.
- Linka, K., St Pierre, S., Kuhl, E., 2023a. Automated model discovery for human brain using constitutive artificial neural networks. *Acta Biomater.* 160, 134–151.
- Linka, K., Tepole, A., Holzapfel, G., Kuhl, E., 2023b. Automated model discovery for skin: discovering the best model, data, and experiment. *Comput. Methods Appl. Mech. Eng.* 410, 116007.
- Lode, W., 1925. Versuche, über den Einfluß der mittleren hauptspannung auf die fließ grenze. *Z. Ang. Math. Mech.* 5, 142–144.
- Lucey, J., 2002. Formation and physical properties of milk protein gels. *J. Dairy Sci.* 85, 281–294.
- Martonová, D., Peirlinck, M., Linka, K., Holzapfel, G., Leyendecker, S., Kuhl, E., 2024. Automated model discovery for human cardiac tissue: discovering the best model and parameters. *Comput. Methods Appl. Mech. Eng.* 428, 117078.
- Masi, F., Stefanou, I., Vannucci, P., Maffi-Berthier, V., 2021. Thermodynamics-based artificial neural networks for constitutive modeling. *J. Mech. Phys. Solids* 147, 104277.
- MATLAB, 2022. MATLAB version 9.13.0 (R2022b). The MathWorks, Inc., Natick, Massachusetts. <https://www.mathworks.com>.
- McCulloch, J., St Pierre, S., Linka, K., Kuhl, E., 2024. On sparse regression, lp-regularization, and automated model discovery. *Int. J. Num. Meth. Eng.* 125, e7481.
- von Mises, R., 1913. *Mechanik der festen Körper im plastisch deformablen Zustand*. Weidmann.

- Mohr, O., 1900. Über die darstellung des spannungszustandes und des bruchverhaltens fester körper. *Z. Ver. Deutsch. Ing.* 44, 1524–1527.
- Mow, C., Holmes, M., Lai, W., 1984. Fluid transport and mechanical properties of articular cartilage: a review. *J. Biomechanics* 17, 377–394.
- Paredes, J., Cortizo-Lacalle, D., Imaz, A., Aldazabal, J., Vila, M., 2022. Application of texture analysis methods for the characterization of cultured meat. *Sci. Rep.* 12 (1), 77785.
- Peirlinck, M., Linka, K., Kuhl, E., 2025. Atrial constitutive neural networks. In: *Proc. Int. Conf. Functional Imaging and Modeling of the Heart*. Springer, pp. 249–259.
- Poore, J., Nemecek, T., 2018. Reducing food's environmental impacts through producers and consumers. *Science* 360, 987–992.
- Qin, P., Wang, T., Luo, Y., 2022. A review on plant-based proteins from soybean: health benefits and soy product development. *J. Agricult. Food Res.* 7, 100265.
- Rao, M., Steffe, J., 1992. *Viscoelastic Properties of Foods*. Elsevier Applied Science, London.
- Reese, S., Govindjee, S., 1998. A theory of finite viscoelasticity and numerical aspects. *Int. J. Solids Struct.* 35, 3455–3482.
- Rizzo, G., Baroni, L., 2018. Review: soy, soy foods and their role in vegetarian diets. *Nutrients* 10 (43).
- Schreuders, F. K.G., Schlangen, M., Kyriakopoulou, K., Boom, R., van der Goot, A., 2021. Texture methods for evaluating meat and meat analogue structures: a review. *Food Contr.* 127, 108103.
- Shurtleff, W., Aoyagi, A., 1975. *The Book of Tofu: Protein Source of the Future*. Soyinfo Center, Lafayette.
- Simo, J., 1987. On a fully three-dimensional finite-strain viscoelastic damage model: formulation and computational aspects. *Comput. Methods Appl. Mech. Eng.* 60, 153–173.
- St Pierre, S., Darwin, E., Adil, D., Aviles, M., Date, A., Dunne, R., Lall, Y., Parra Vallecillo, M., Perez Medina, V., Linka, K., Kuhl, E., 2024. The mechanical and sensory signature of plant-based and animal meat. *npj Sci. Food* 8 (1), 94.
- St Pierre, S., Rajasekharan, D., Darwin, E., Linka, K., Levenston, M., Kuhl, E., 2023. Discovering the mechanics of artificial and real meat. *Comput. Methods Appl. Mech. Eng.* 415, 116236.
- Tac, V., Gardner, C., Kuhl, E., 2023. Generative artificial intelligence creates delicious, sustainable, and nutritious burgers. *arXiv:2602.03092*
- Tresca, H., 1864. Mémoire sur l'écoulement des corps solides soumis à de fortes pressions. *C. R. l'Académie Sci.* 59, 754–758.
- van der Velden, T., Boes, B., Brepols, T., Kuhl, E., Holthuisen, H., 2025. A note on constitutive artificial neural networks for finite strain plasticity. 5588650.
- Vervenne, T., Peirlinck, M., Famaey, N., Kuhl, E., 2025a. Constitutive neural networks for main pulmonary arteries: discovering the undiscovered. *Biomech. Model. Mechanobiol.* 24, 615–634.
- Vervenne, T., St Pierre, S., Famaey, N., Kuhl, E., 2025b. Probing mycelium mechanics and taste: the moist and fibrous signature of fungi steak. *Acta Biomater.* 202, 341–351.
- Wang, L., Linka, K., Kuhl, E., 2023. Automated model discovery for muscle using constitutive recurrent neural networks. *J. Mech. Behav. Biomed. Mat.* 145, 106021.
- Wang, X., Luo, K., Liu, S., Zeng, M., Adhikari, B., He, Z., Chen, J., 2018. Textural and rheological properties of soy protein isolate tofu-type emulsion gels: influence of soybean variety and coagulant type. *Food Biophys.* 13 (3), 324–332.
- Watanabe, T., 1997. *The Science of Tofu*. Food Journal Co., Kyoto.
- Yuan, Y., Zhang, Q., Lin, S., Li, J., 2025. Water: the soul of hydrogels. *Prog. Mat. Sci.* 148, 101378.
- Zheng, L., Regenstein, J., Teng, F., Li, Y., 2020. Tofu products: a review of their raw materials, processing conditions, and packaging. *Compr. Rev. Food Sci. Food Safety* 19, 3683–3714.