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

Birte Boes[a,*], Jaan-Willem Simon[a], Hagen Holthusen[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*

**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 viscoelasticy, 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 in 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.

*Keywords:* inelasticity, viscoelasticity, poroelasticity, automated model discovery, feature network, food, tofu

*"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 [68]. It is one of the simplest and most influential plant-based foods in human history [80]. Beyond its cultural role, tofu promotes human health [63], contributes to planetary sustainability [62], and remains the cornerstone of a modern plant-based diet [66]. Yet, the material itself is remarkably simple [41]: Soymilk and a mineral coagulant form a hydrated protein gel with only two essential ingredients, soybeans and water [84]. Whether we make our own tofu at home or buy it in supermarkets, Figure 1 illustrates that its texture can range from silken to extrafirm, determined almost entirely by its water content [13] and by the specifics of its coagulation process [81]. 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 [26]. In texture profile analysis , mechanical double-compression tests simulate the behavior of chewing [5], we can easily extract its parameters such as stiffness, hardness, cohesiveness, springiness, resilience, and chewiness from the resulting force–time curves [22]. Although texture profile analysis continues to be widely used [60], the lack of standardized testing protocols limits comparability across the different studies and labs [67]. 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.

*Corresponding author
*Email addresses:* birte.boes@uni-wuppertal.de (Birte Boes), jaan.simon@uni-wuppertal.de (Jaan-Willem Simon), hagen.holthusen@fau.de (Hagen Holthusen), ekuhl@stanford.edu (Ellen Kuhl)

Mechanical testing allows us to interpret texture through the lens of mechanics and to extract fundamental material properties directly from experimental data [71]. 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 [12]. Grounding texture evaluation in the principles of material modeling [77] enables a more rigorous and interpretable mechanistic analysis [72]. 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 [17]. 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 [45]. 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 [52]. This nonlinear, often rate-dependent response calls for new approaches that are capable of identifying the governing equations directly from experimental data [43]. Recent advances in data-driven mechanics have introduced frameworks that automatically discover constitutive relationships [47] without pre-selecting the specific functional form of the model [28]. Instead of repeatedly fitting and refining candidate models, automated model discovery infers the governing equations in a self-consistent and physics-informed manner [24]. 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 [24] as well as mechanics-informed [3], physics-constrained [38], physics-augmented [39, 40], thermodynamics-based [54], input convex neural networks [2], and input monotonic neural networks [32]. Among these, the framework of constitutive artificial neural networks has proven particularly effective because it hardwires continuum mechanical principles directly into the network architecture [46]. This design ensures that the resulting models remain both physically sound and interpretable [47]. 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 [49], arteries [77], the heart [53], and the brain [34, 48], and, most recently, plant-based and animal meats [70]. 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 [1, 82]. 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 [28]. 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 [29], growth modeling [30], and plasticity [4, 76]. Importantly, inelastic constitutive neural networks autonomously discover whether inelastic effects are necessary, and reduce unnecessary model complexity when the data do not justify it [31].

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.

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 1 summarizes their serving sizes, calories, and macronutrient compositions per serving according

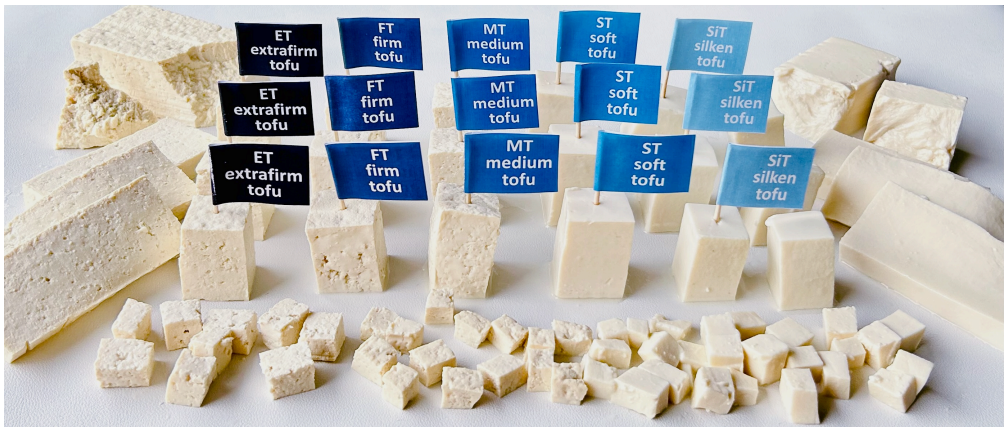


Figure 1: **Five different types of tofu.** Extra-firm, firm, medium, soft, and silken tofu illustrate how small changes in composition and processing create large differences in mechanical behavior. Tofu combines soy beans with water to form a biphasic protein gel whose stiffness, porosity, and fracture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

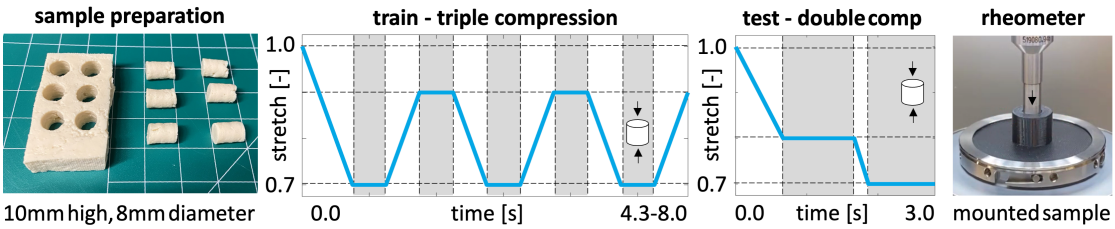


Figure 2: **Mechanical testing.** Preparation of cylindrical samples; triple compression test for training; double compression test for testing; and rheometer with mounted sample. During triple compression, we load and unload between $\lambda = 0.7$ and $\lambda = 0.9$ at slow, medium, and fast stretch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

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

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 |

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 $w_\mathrm{f}$, we divide the water mass, $m_\mathrm{f} = \{\,76\mathrm{g}, 71\mathrm{g}, 70\mathrm{g}\,\}$, by the total mass per serving of 85g, resulting in $w_\mathrm{f} = \{\,0.89, 0.84, 0.82\,\}$ for the three tofu types. To calculate the volume fraction of the fluid $\nu_\mathrm{f} = V_\mathrm{f}/[\,V_\mathrm{f} + V_\mathrm{s}\,]$, we divide the volume of the fluid, $V_\mathrm{f} = w_\mathrm{f}/\rho_\mathrm{f}$, by the total volume of fluid and solid, $V_\mathrm{f} + V_\mathrm{s}$, with $V_\mathrm{s} = [\,1 - w_\mathrm{f}\,]/\rho_\mathrm{s}$, assuming fluid and solid densities of $\rho_\mathrm{f} = 1.0\,\mathrm{g/ml}$ and $\rho_\mathrm{s} = 1.3\,\mathrm{g/ml}$, resulting in $\nu_\mathrm{f} = \{\,0.92, 0.87, 0.86\,\}$ 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. Figure 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, Delware). Figure 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 Figure 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 $\dot{\lambda} = \{\, 0.2/\mathrm{s}, 0.4/\mathrm{s}, 0.6/\mathrm{s} \,\}$. During double compression in Figure 2 middle right, we compress each specimen first to a stretch of $\lambda = 0.8$ at a rate of $\dot{\lambda} = 0.3/\mathrm{s}$, hold for 1.0 s, further compress to $\lambda = 0.7$ at a rate of $\dot{\lambda} = 0.5/\mathrm{s}$, and hold for another 1.0 s.

**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 [55] with a filter width of $\sigma_{\mathrm{gauss}} = \{\, 8, 3, 3 \,\}$ 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 $A = \pi \cdot (4\,\mathrm{mm})^2$ and the stretch $\lambda$.

## 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 $\boldsymbol{F}$ into elastic and inelastic parts, $\boldsymbol{F}_{\mathrm{e}}$ and $\boldsymbol{F}_{\mathrm{i}}$,

$$\boldsymbol{F} = \boldsymbol{F}_{\mathrm{e}} \cdot \boldsymbol{F}_{\mathrm{i}} \,. \tag{1}$$

As deformation measures, we introduce the right Cauchy Green deformation tensor $\boldsymbol{C}$ and its elastic and inelastic counterparts, $\boldsymbol{C}_{\mathrm{e}}$ and $\boldsymbol{C}_{\mathrm{i}}$,

$$\boldsymbol{C} = \boldsymbol{F}^{\mathrm{t}} \cdot \boldsymbol{F} \quad \text{and} \quad \boldsymbol{C}_{\mathrm{e}} = \boldsymbol{F}_{\mathrm{e}}^{\mathrm{t}} \cdot \boldsymbol{F}_{\mathrm{e}} = \boldsymbol{F}_{\mathrm{i}}^{-\mathrm{t}} \cdot \boldsymbol{C} \cdot \boldsymbol{F}_{\mathrm{i}}^{-1} \quad \text{and} \quad \boldsymbol{C}_{\mathrm{i}} = \boldsymbol{F}_{\mathrm{i}}^{\mathrm{t}} \cdot \boldsymbol{F}_{\mathrm{i}} \,. \tag{2}$$

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

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

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

$$\tilde{I}_1^{\boldsymbol{C}_{\mathrm{e}}} = \mathrm{tr}(\boldsymbol{C}_{\mathrm{e}})/(I_3^{\boldsymbol{C}_{\mathrm{e}}})^{1/3} \quad \text{and} \quad \tilde{I}_2^{\boldsymbol{C}_{\mathrm{e}}} = \mathrm{tr}(\mathrm{cof}(\boldsymbol{C}_{\mathrm{e}}))/(I_3^{\boldsymbol{C}_{\mathrm{e}}})^{2/3} \quad \text{and} \quad I_3^{\boldsymbol{C}_{\mathrm{e}}} = \det(\boldsymbol{C}_{\mathrm{e}}) \,. \tag{4}$$

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

$$\psi = \psi^{\mathrm{eq}}(\boldsymbol{C}) + \psi^{\mathrm{neq}}(\boldsymbol{C}_{\mathrm{e}}) = \psi^{\mathrm{eq}}(\tilde{I}_1^{\boldsymbol{C}}, \tilde{I}_2^{\boldsymbol{C}}, I_3^{\boldsymbol{C}}) + \psi^{\mathrm{neq}}(\tilde{I}_1^{\boldsymbol{C}_{\mathrm{e}}}, \tilde{I}_2^{\boldsymbol{C}_{\mathrm{e}}}, I_3^{\boldsymbol{C}_{\mathrm{e}}}) \,. \tag{5}$$

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

$$\mathcal{D} := \left[\, \boldsymbol{S} - 2\,\frac{\partial \psi^{\mathrm{eq}}}{\partial \boldsymbol{C}} - 2\,\boldsymbol{F}_{\mathrm{i}}^{-1} \cdot \frac{\partial \psi^{\mathrm{neq}}}{\partial \boldsymbol{C}_{\mathrm{e}}} \cdot \boldsymbol{F}_{\mathrm{i}}^{-\mathrm{t}} \,\right] : \frac{1}{2}\dot{\boldsymbol{C}} + 2\,\boldsymbol{C}_{\mathrm{e}} \cdot \frac{\partial \psi^{\mathrm{neq}}}{\partial \boldsymbol{C}_{\mathrm{e}}} : \boldsymbol{D}_{\mathrm{i}} \geq 0 \,, \tag{6}$$

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

$$\boldsymbol{S} = 2\,\frac{\partial\psi^{\mathrm{eq}}}{\partial\boldsymbol{C}} + 2\,\boldsymbol{F}_{\mathrm{i}}^{-1} \cdot \frac{\partial\psi^{\mathrm{neq}}}{\partial\boldsymbol{C}_{\mathrm{e}}} \cdot \boldsymbol{F}_{\mathrm{i}}^{-\mathrm{t}} = \boldsymbol{S}^{\mathrm{eq}} + \boldsymbol{S}^{\mathrm{neq}}\,, \tag{7}$$

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

$$\boldsymbol{S}^{\mathrm{eq}} = 2\,\frac{\partial\psi^{\mathrm{eq}}}{\partial\tilde{I}_1^{\boldsymbol{C}}}\,[\,\boldsymbol{I} - \tfrac{1}{3}I_1^{\boldsymbol{C}}\,\boldsymbol{C}^{-1}]/(I_3^{\boldsymbol{C}})^{1/3} + 2\,\frac{\partial\psi^{\mathrm{eq}}}{\partial\tilde{I}_2^{\boldsymbol{C}}}\,[\,I_1^{\boldsymbol{C}}\boldsymbol{I} - \boldsymbol{C} - \tfrac{2}{3}I_2^{\boldsymbol{C}}\,\boldsymbol{C}^{-1}]/(I_3^{\boldsymbol{C}})^{2/3} + 2\,\frac{\partial\psi^{\mathrm{eq}}}{\partial I_3^{\boldsymbol{C}}}\,I_3^{\boldsymbol{C}}\,\boldsymbol{C}^{-1}\,, \tag{8}$$

and its non-equilibrium part $\boldsymbol{S}^{\mathrm{neq}} := 2\,\boldsymbol{F}_{\mathrm{i}}^{-1} \cdot \partial\psi^{\mathrm{neq}}/\partial\boldsymbol{C}_{\mathrm{e}} \cdot \boldsymbol{F}_{\mathrm{i}}^{-\mathrm{t}}$,

$$\boldsymbol{S}^{\mathrm{neq}} = 2\frac{\partial\psi^{\mathrm{neq}}}{\partial\tilde{I}_1^{\boldsymbol{C}_{\mathrm{e}}}}[\boldsymbol{C}_{\mathrm{i}}^{-1} - \tfrac{1}{3}I_1^{\boldsymbol{C}_{\mathrm{e}}}\,\boldsymbol{C}^{-1}]/(I_3^{\boldsymbol{C}_{\mathrm{e}}})^{1/3} + 2\frac{\partial\psi^{\mathrm{neq}}}{\partial\tilde{I}_2^{\boldsymbol{C}_{\mathrm{e}}}}[I_1^{\boldsymbol{C}_{\mathrm{e}}}\boldsymbol{C}_{\mathrm{i}}^{-1} - \boldsymbol{C}_{\mathrm{i}}^{-1} \cdot \boldsymbol{C} \cdot \boldsymbol{C}_{\mathrm{i}}^{-1} - \tfrac{2}{3}I_2^{\boldsymbol{C}_{\mathrm{e}}}\boldsymbol{C}^{-1}]/(I_3^{\boldsymbol{C}_{\mathrm{e}}})^{2/3} + 2\frac{\partial\psi^{\mathrm{neq}}}{\partial I_3^{\boldsymbol{C}_{\mathrm{e}}}}I_3^{\boldsymbol{C}_{\mathrm{e}}}\boldsymbol{C}^{-1}. \tag{9}$$

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

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

$$\boldsymbol{\Sigma} = 2\,\frac{\partial\psi^{\mathrm{neq}}}{\partial\tilde{I}_1^{\boldsymbol{C}_{\mathrm{e}}}}\,[\,\boldsymbol{C}_{\mathrm{e}} - \tfrac{1}{3}I_1^{\boldsymbol{C}_{\mathrm{e}}}\boldsymbol{I}\,]/(I_3^{\boldsymbol{C}_{\mathrm{e}}})^{1/3} + 2\,\frac{\partial\psi^{\mathrm{neq}}}{\partial\tilde{I}_2^{\boldsymbol{C}_{\mathrm{e}}}}\,[\,I_1^{\boldsymbol{C}_{\mathrm{e}}}\boldsymbol{C}_{\mathrm{e}} - \boldsymbol{C}_{\mathrm{e}}^2 - \tfrac{2}{3}I_2^{\boldsymbol{C}_{\mathrm{e}}}\boldsymbol{I}\,]/(I_3^{\boldsymbol{C}_{\mathrm{e}}})^{2/3} + 2\,\frac{\partial\psi^{\mathrm{neq}}}{\partial I_3^{\boldsymbol{C}_{\mathrm{e}}}}\,I_3^{\boldsymbol{C}_{\mathrm{e}}}\boldsymbol{I}\,. \tag{10}$$

To characterize the dissipative volume-changing and volume-preserving behavior of tofu, we introduce the volumetric invariant $I_1^{\boldsymbol{\Sigma}}$ and the deviatoric invariants $J_2^{\boldsymbol{\Sigma}}$ and $J_3^{\boldsymbol{\Sigma}}$ of the Mandel stress $\boldsymbol{\Sigma}$,

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

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

$$\phi = \phi(\boldsymbol{\Sigma}) = \phi(I_1^{\boldsymbol{\Sigma}}, J_2^{\boldsymbol{\Sigma}}, J_3^{\boldsymbol{\Sigma}})\,. \tag{12}$$

To a priori satisfy the reduced Clausius–Planck inequality,

$$\mathcal{D}_{\mathrm{red}} = \boldsymbol{\Sigma} : \boldsymbol{D}_{\mathrm{i}} \geq 0 \tag{13}$$

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

$$\boldsymbol{D}_{\mathrm{i}} = \frac{\partial\phi}{\partial I_1^{\boldsymbol{\Sigma}}}\,\boldsymbol{I} + \frac{\partial\phi}{\partial J_2^{\boldsymbol{\Sigma}}}\,\mathrm{dev}(\boldsymbol{\Sigma}) + \frac{\partial\phi}{\partial J_3^{\boldsymbol{\Sigma}}}\,\mathrm{dev}(\mathrm{dev}(\boldsymbol{\Sigma})^2)\,. \tag{14}$$

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

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

$$\boldsymbol{C}_{\mathrm{i}}^{\mathrm{n+1}} = (\boldsymbol{F}_{\mathrm{i}}^{\mathrm{n}})^{\mathrm{t}} \cdot \exp(2\,\Delta t\,\boldsymbol{D}_{\mathrm{i}}^{\mathrm{n}}) \cdot \boldsymbol{F}_{\mathrm{i}}^{\mathrm{n}} \quad \text{with} \quad \boldsymbol{F}_{\mathrm{i}}^{\mathrm{n}} = +(\boldsymbol{C}_{\mathrm{i}}^{\mathrm{n}})^{1/2}\,. \tag{15}$$

In Section 4, we design a neural network for the dissipative potential $\phi$ from which we derive the inelastic rate of deformation tensor $\boldsymbol{D}_{\mathrm{i}}$, the inelastic right Cauchy Green tensor $\boldsymbol{C}_{\mathrm{i}}$, and the inelastic part of the deformation gradient

$\boldsymbol{F}_{\mathrm{i}}$ 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 $\boldsymbol{\alpha}$ that can modify the elastic and dissipative potentials,

$$\psi = \psi^{\mathrm{eq}}(\boldsymbol{C}, \boldsymbol{\alpha}) + \psi^{\mathrm{neq}}(\boldsymbol{C}_{\mathrm{e}}, \boldsymbol{\alpha}) \qquad \text{and} \qquad \phi = \phi(\boldsymbol{\Sigma}, \boldsymbol{\alpha}). \tag{16}$$

Here, we acknowledge the biphasic nature of tofu and propose a single scalar-valued input feature, the water content $\alpha = n^{\mathrm{F}}$. In Section 4, we design a feature network that we couple to the networks for the equilibrium potential $\psi^{\mathrm{eq}}$, the non-equilibrium potential $\psi^{\mathrm{neq}}$, and the dissipative potential $\phi$ 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 $\boldsymbol{F}$ that we multiplicatively decompose into elastic and inelastic parts, $\boldsymbol{F}_{\mathrm{e}}$ and $\boldsymbol{F}_{\mathrm{i}}$,

$$\boldsymbol{F} = \mathrm{diag}\,\{\,1,\,1,\,\lambda\,\} = \boldsymbol{F}_{\mathrm{e}} \cdot \boldsymbol{F}_{\mathrm{i}}\,, \tag{17}$$

with the Jacobian, $J = \det(\boldsymbol{F}) = \lambda_r\,\lambda_\theta\,\lambda_z = \lambda$, defining the volumetric change of the sample. We assume that the inelastic deformation is isotropic and volume preserving, $J_{\mathrm{i}} = \det(\boldsymbol{F}_{\mathrm{i}}) = 1$, and that all volumetric changes are associated with the elastic deformation only, $J_{\mathrm{e}} = \det(\boldsymbol{F}_{\mathrm{e}}) = \lambda$ [33, 65, 69], such that elastic and inelastic parts of the deformation gradient are

$$\boldsymbol{F}_{\mathrm{e}} = \mathrm{diag}\,\{\,\lambda_{\mathrm{i}}^{1/2}, \lambda_{\mathrm{i}}^{1/2}, \lambda/\lambda_{\mathrm{i}}\,\} \quad \text{and} \quad \boldsymbol{F}_{\mathrm{i}} = \mathrm{diag}\,\{\,\lambda_{\mathrm{i}}^{-1/2}, \lambda_{\mathrm{i}}^{-1/2}, \lambda_{\mathrm{i}}\,\}\,. \tag{18}$$

These explicit definitions introduce the right Cauchy Green tensor $\boldsymbol{C}$ and its elastic counterpart $\boldsymbol{C}_{\mathrm{e}}$,

$$\boldsymbol{C} = \boldsymbol{F}^{\mathrm{t}} \cdot \boldsymbol{F} = \mathrm{diag}\,\{\,1, 1, \lambda^2\,\} \quad \text{and} \quad \boldsymbol{C}_{\mathrm{e}} = \boldsymbol{F}_{\mathrm{e}}^{\mathrm{t}} \cdot \boldsymbol{F}_{\mathrm{e}} = \mathrm{diag}\,\{\,\lambda_{\mathrm{i}}, \lambda_{\mathrm{i}}, \lambda^2/\lambda_{\mathrm{i}}^2\,\} \tag{19}$$

with invariants

$$\begin{array}{lll} I_1^{\boldsymbol{C}} = 2 + \lambda^2 & I_2^{\boldsymbol{C}} = 1 + 2\lambda^2 & I_3^{\boldsymbol{C}} = \lambda^2 \\ \tilde{I}_1^{\boldsymbol{C}} = 2/\lambda^{2/3} + \lambda^{4/3} & \tilde{I}_2^{\boldsymbol{C}} = 1/\lambda^{4/3} + 2\lambda^{2/3} & I_3^{\boldsymbol{C}} = \lambda^2 \\ \tilde{I}_1^{\boldsymbol{C}_{\mathrm{e}}} = 2\lambda_{\mathrm{i}}/\lambda^{2/3} + \lambda^{4/3}/\lambda_{\mathrm{i}}^2 & \tilde{I}_2^{\boldsymbol{C}_{\mathrm{e}}} = 2\lambda^{2/3}/\lambda_{\mathrm{i}} + \lambda_{\mathrm{i}}^2/\lambda^{4/3} & I_3^{\boldsymbol{C}_{\mathrm{e}}} = \lambda^2. \end{array} \tag{20}$$

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,

$$S^{\mathrm{eq}} = \frac{4}{3}\left[1 - \frac{1}{\lambda^2}\right]\left[\frac{1}{\lambda^{2/3}}\right]\frac{\partial \psi^{\mathrm{eq}}}{\partial \tilde{I}_1^{\boldsymbol{C}}} + \frac{4}{3}\left[1 - \frac{1}{\lambda^2}\right]\left[\frac{1}{\lambda^{4/3}}\right]\frac{\partial \psi^{\mathrm{eq}}}{\partial \tilde{I}_2^{\boldsymbol{C}}} + 2\,\frac{\partial \psi^{\mathrm{eq}}}{\partial I_3^{\boldsymbol{C}}}\,, \tag{21}$$

and the non-equilibrium stress in axial direction,

$$S^{\mathrm{neq}} = \frac{4}{3}\left[\frac{1}{\lambda_{\mathrm{i}}^2} - \frac{\lambda_{\mathrm{i}}}{\lambda^2}\right]\left[\frac{1}{\lambda^{2/3}}\right]\frac{\partial \psi^{\mathrm{neq}}}{\partial \tilde{I}_1^{\boldsymbol{C}_{\mathrm{e}}}} + \frac{4}{3}\left[\frac{1}{\lambda_{\mathrm{i}}} - \frac{\lambda_{\mathrm{i}}^2}{\lambda^2}\right]\left[\frac{1}{\lambda^{4/3}}\right]\frac{\partial \psi^{\mathrm{neq}}}{\partial \tilde{I}_2^{\boldsymbol{C}_{\mathrm{e}}}} + 2\,\frac{\partial \psi^{\mathrm{neq}}}{\partial I_3^{\boldsymbol{C}_{\mathrm{e}}}}\,, \tag{22}$$

solely in terms of total and inelastic stretches, $\lambda$ and $\lambda_{\mathrm{i}}$, and the derivatives of the elastic potentials, $\psi^{\mathrm{eq}}$ and $\psi^{\mathrm{neq}}$, 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 $\psi^{\mathrm{eq}}$ [47], the non-equilibrium potential $\psi^{\mathrm{neq}}$ [30], the dissipative potential $\phi$ [31], 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 $\boldsymbol{F}^{\mathrm{n}}$ and the internal variables $\boldsymbol{F}_{\mathrm{i}}^{\mathrm{n}}$ from the previous time point $t^{\mathrm{n}}$ as input, integrates them in time using the exponential update (15), stores the internal variables $\boldsymbol{F}_{\mathrm{i}}^{\mathrm{n+1}}$, and outputs the second Piola Kirchhoff stress $\boldsymbol{S}^{\mathrm{n+1}}$ at the new time point $t^{\mathrm{n+1}}$.

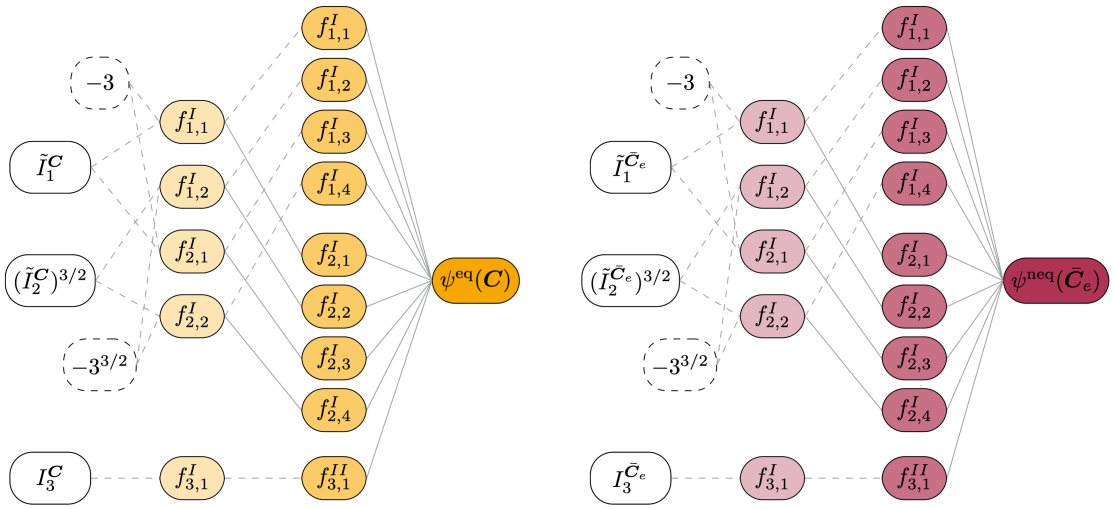


Figure 3: **Elastic networks.** The elastic networks have two hidden layers with 5 and 9 nodes per layer. The feed forward networks take the invariants of the total right Cauchy Green tensor $\tilde{I}_1^{\boldsymbol{C}}$, $\tilde{I}_2^{\boldsymbol{C}}$, $I_3^{\boldsymbol{C}}$, left, or of the elastic right Cauchy Green tensor $\tilde{I}_1^{\boldsymbol{C}_e}$, $\tilde{I}_2^{\boldsymbol{C}_e}$, $I_3^{\boldsymbol{C}_e}$, 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 $\psi^{\mathrm{eq}}$, left, or $\psi^{\mathrm{neq}}$, right, as output. Activation functions $f_{\beta,\gamma}^{\delta}$ are labeled by activation type $\beta$, node index $\gamma$, and layer index $\delta$; solid lines denote trainable weights, dashed lines indicate fixed weights equal to one, and dashed ellipses denote constant, non-trainable biases.

**Elastic networks.** To model the equilibrium potential $\psi^{\mathrm{eq}}$ [47] and the non-equilibrium potential $\psi^{\mathrm{neq}}$ [31], we introduce two structurally similar but input-independent feed forward networks. The equilibrium network uses the invariants of the total right Cauchy Green tensor $\tilde{I}_1^{\boldsymbol{C}}$, $\tilde{I}_2^{\boldsymbol{C}}$, $I_3^{\boldsymbol{C}}$ from equation (3) [48, 53, 61] and the non-equilibrium network uses the invariants of the elastic right Cauchy Green tensor $\tilde{I}_1^{\boldsymbol{C}_e}$, $\tilde{I}_2^{\boldsymbol{C}_e}$, $I_3^{\boldsymbol{C}_e}$ from equation (4) [28, 29, 30] 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,

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

Figure 3 visualizes the architecture of both elastic networks that a priori satisfy the thermodynamic requirements of polyconvexity, growth, and normalization [4, 31, 47]. The equilibrium and non-equilibrium networks each contain 14 trainable parameters, which are visualized in Figure 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 $\phi$ [31], we introduce a recurrent neural network that uses the invariants of the Mandel stress $I_1^{\boldsymbol{\Sigma}}$, $J_2^{\boldsymbol{\Sigma}}$, $J_3^{\boldsymbol{\Sigma}}$ from equation (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 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,

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

Figure 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 $\mathcal{D}_{\mathrm{red}} \geq 0$ (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 [31]. 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

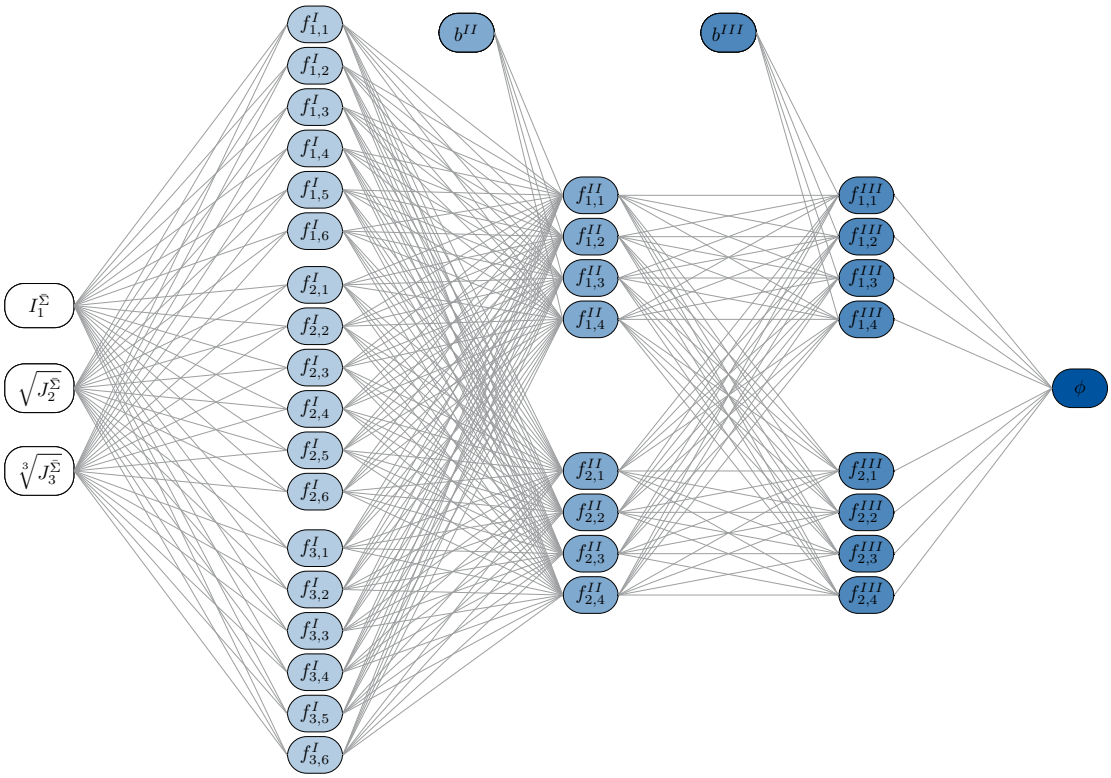


Figure 4: **Dissipative network.** The dissipative network has three hidden layers with 18, 8, and 8 nodes per layer. It take the invariants of the Mandel stress $I_1^{\boldsymbol{\Sigma}}$, $J_2^{\boldsymbol{\Sigma}}$, $J_3^{\boldsymbol{\Sigma}}$ as input; the first layer applies trainable powers $p$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

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 $\alpha$. 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,

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

Figure 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 $n^{\max} = 9$ nodes for the elastic networks in Figure 3 and $n^{\max} = 8$ nodes for the dissipative network in Figure 4.

Figure 6 illustrates the coupling between one of the three mechanical networks and the feature network. The mechanical networks from Figure 3 or 4 receive a strain- or stress-like input $\boldsymbol{A}$ and the feature network from Figure 5 receives a feature-like input $\boldsymbol{\alpha}$. The networks are multiplicatively coupled by summing up the product of the nodes in their final layers to obtain feature-weighted potential $\Psi$,

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

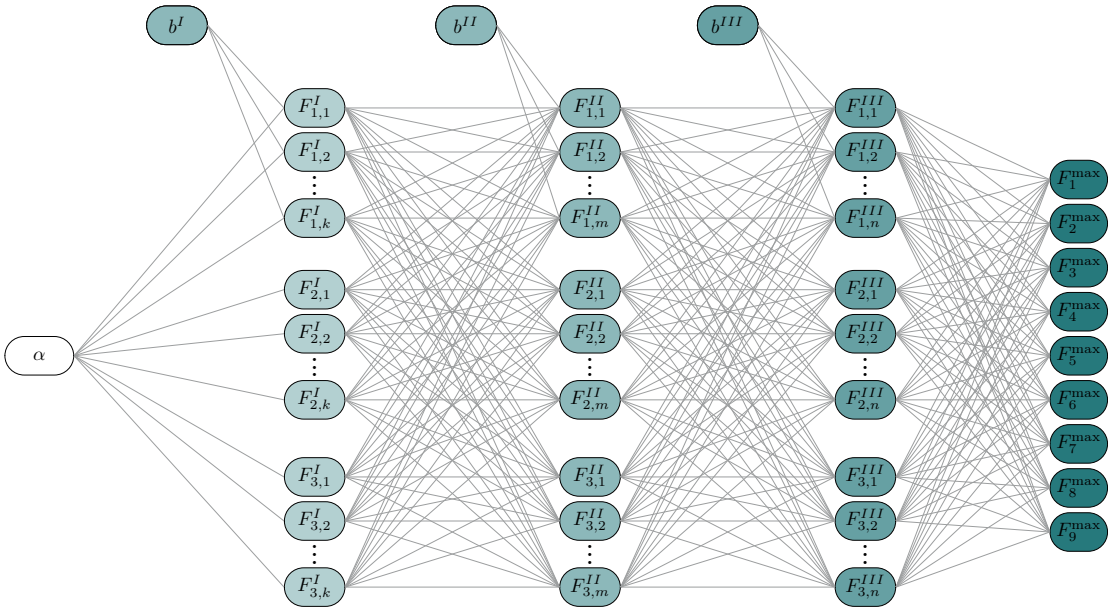


Figure 5: **Feature network.** The feature network has three hidden layers with $k \times 4$ nodes per layer. It takes the feature $\alpha$, in our case the scalar-valued water content, as input; the three hidden layer apply the maximum, identity, hyperbolic tangent, and softplus functions; the output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

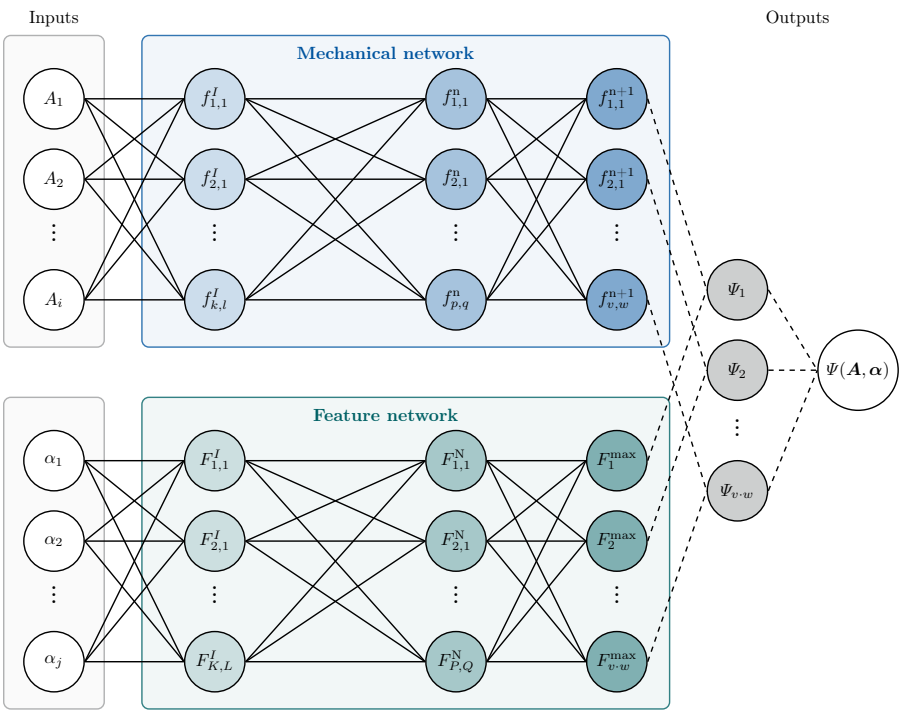


Figure 6: **Coupling of mechanical and feature networks.** The mechanical network from Figure 3 or 4 receives a strain- or stress-like input $\boldsymbol{A}$ and the feature network from Figure 5 receives a feature-like input $\boldsymbol{\alpha}$. The networks are multiplicatively coupled in the nodes of their final layers. The sum of the final layer defines the network output, the feature-weighted mechanical potential $\Psi(\boldsymbol{A},\boldsymbol{\alpha})$. Activation functions $f^{\delta}_{\gamma}$ and $F^{\delta}_{\gamma}$ of the mechanical and feature networks are labeled by node index $\gamma$ and layer index $\delta$.

where the potential $\Psi$ is either the equilibrium or non-equilibrium potential $\psi^{\mathrm{eq}}$ or $\psi^{\mathrm{neq}}$, or the dissipative potential $\phi$, and the input $\boldsymbol{A}$ is either the total or the elastic right Cauchy Green tensor $\boldsymbol{C}$ or $\boldsymbol{C}_{\mathrm{e}}$ or the Mandel stress $\boldsymbol{\Sigma}$. This multiplicative coupling ensures that the combined output is non-negative and remains zero for zero-valued mechanical

outputs. 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 $n_{\text{exp}}$ experiments and all $n_{\text{data}}$ 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 $\alpha_1$ and $\alpha_2$ are the regularization coefficients. The total loss function reads

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

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 $\psi^{\text{eq}}$ and $\psi^{\text{neq}}$ and the dissipative potential $\phi$. 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.

**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, $\psi^{\text{eq}}$ and $\psi^{\text{neq}}$ and $\phi$. 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 $I_i^A = \{ I_i^C, I_i^{C_{\text{e}}}, I_i^{\Sigma} \}$ of our network input $\boldsymbol{A} = \{ \boldsymbol{C}, \boldsymbol{C}_{\text{e}}, \boldsymbol{\Sigma} \}$, the total and elastic right Cauchy Green tensors $\boldsymbol{C}$ and $\boldsymbol{C}_{\text{e}}$ and the Mandel stress $\boldsymbol{\Sigma}$. According to equations (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 $\{ \boldsymbol{S}^{\text{eq}}, \boldsymbol{S}^{\text{neq}}, \boldsymbol{D}_{\text{i}} \}$. We train the combined feature–mechanical network on these derivatives, which contain the essential physical information.

Figure 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 $\Psi(\boldsymbol{A}, \alpha, \mathbf{w})$ of the feature-mechanical network that takes both the invariants of $\boldsymbol{A}$ and the feature $\alpha$ 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 $m = 1, ..., n_{\text{data}}$ data points,

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

We adopt the Adam optimizer and train the equilibrium network for 5,000 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.

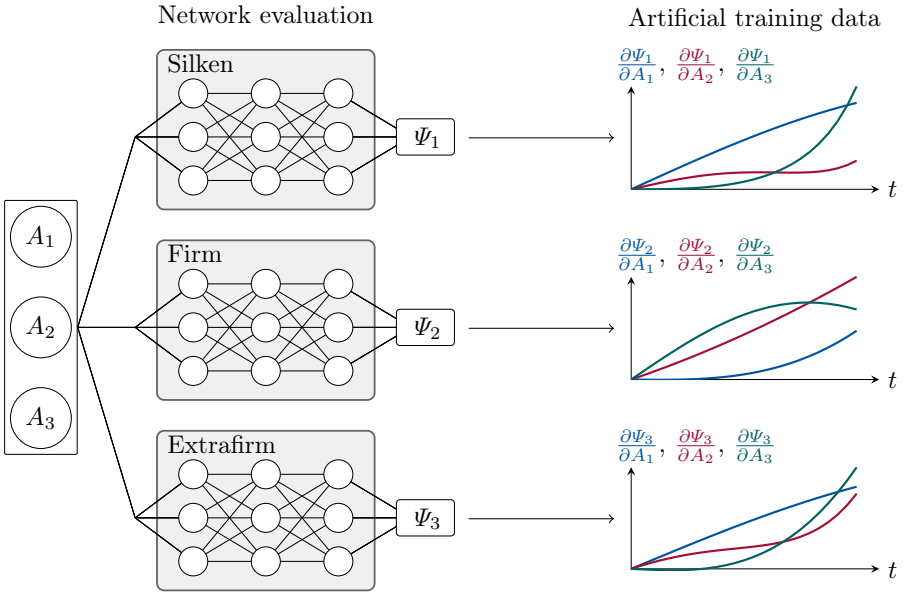


Figure 7: **Generating synthetic training data.** After discovering the best models and parameters for for silken, firm, and extrafirm tofu by training the equilibrium and non-equilibrium networks in Figures 3 and the dissipative network in Figure 4 on the experimental data, we generate synthetic data to train the coupled mechanical and feature networks in Figure 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.

## 5. Results

**Mechanical testing.** Figures 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

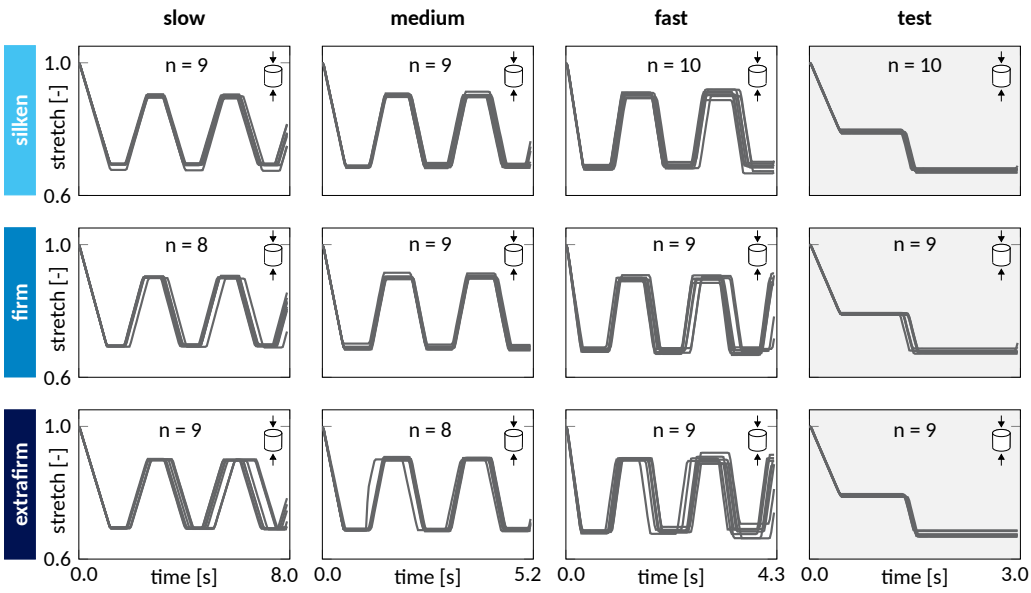


Figure 8: **Mechanical testing - Experimentally applied stretches.** Stretch histories $\lambda(t)$ during triple-compression and double-compression tests for silken (top), firm (middle), and extrafirm (bottom) tofu. The first three columns show the triple-compression tests at slow, medium, and fast 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.

all tofu types and loading conditions and confirm that we can robustly record stretch and stress histories at different

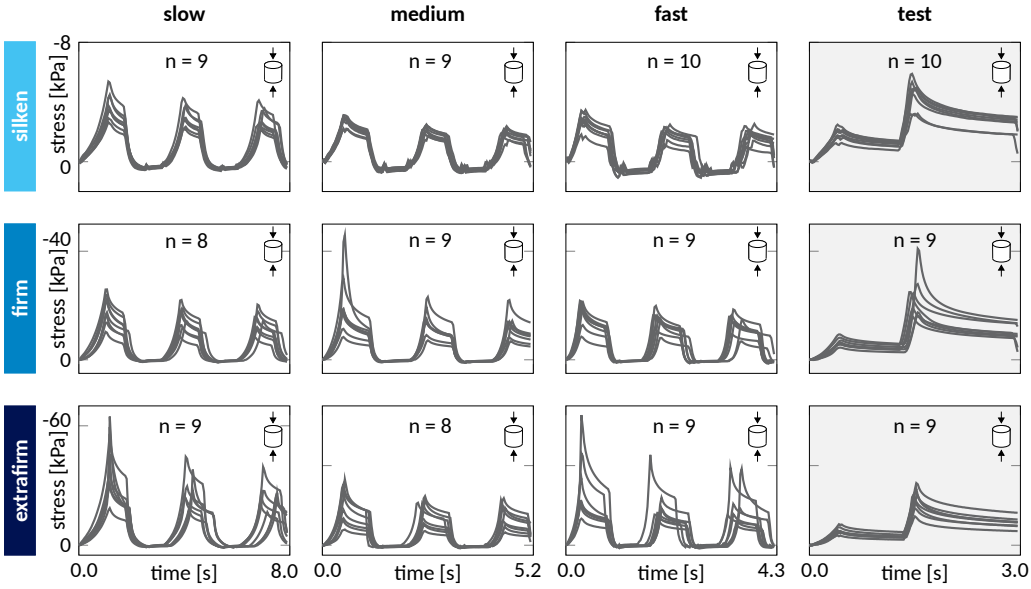


Figure 9: **Mechanical testing - Experimentally recorded stresses.** Stress histories $S(t)$ during triple-compression and double-compression tests for silken (top), firm (middle), and extrafirm (bottom) tofu. The first three columns show the triple-compression tests at slow, medium, and fast strain 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.

stretch magnitudes and rates. However, especially at higher strain rates, we observe slight variations in the applied stretch profiles resulting in slight inconsistencies in the stress profiles, which complicate 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.

**Discovering three individual tofu models.** Figure 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 $\mathsf{R}^2_{\text{min}} = 0.5420$ to $\mathsf{R}^2_{\text{max}} = 0.9262$. 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 $\mathsf{R}^2_{\text{silken}} = 0.7105$, $\mathsf{R}^2_{\text{firm}} = 0.7471$ and $\mathsf{R}^2_{\text{extra}} = 0.6564$ 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,

$$\begin{aligned}\psi^{\text{eq}}_{\text{silken}}(\boldsymbol{C}) &= \tfrac{1}{2}\,\mu_7\,[(\tilde{I}^{\boldsymbol{C}}_2)^{3/2} - 3^{3/2}]^2 + \alpha_9[(I^{\boldsymbol{C}}_3)^{\beta_9} - 1 - \beta_9 \ln(I^{\boldsymbol{C}}_3)] \\ \psi^{\text{neq}}_{\text{silken}}(\boldsymbol{C}_{\text{e}}) &= \tfrac{1}{2}\,\mu_5\,[(\tilde{I}^{\boldsymbol{C}_{\text{e}}}_2)^{3/2} - 3^{3/2}]\,,\end{aligned}$$

with network weights $\mu_7 = 0.158\,\text{kPa}$, $\alpha_9 = 0.052\,\text{kPa}$, $\beta_9 = 5.507$ for the equilibrium part and $\mu_5 = 1.179\,\text{kPa}$ 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

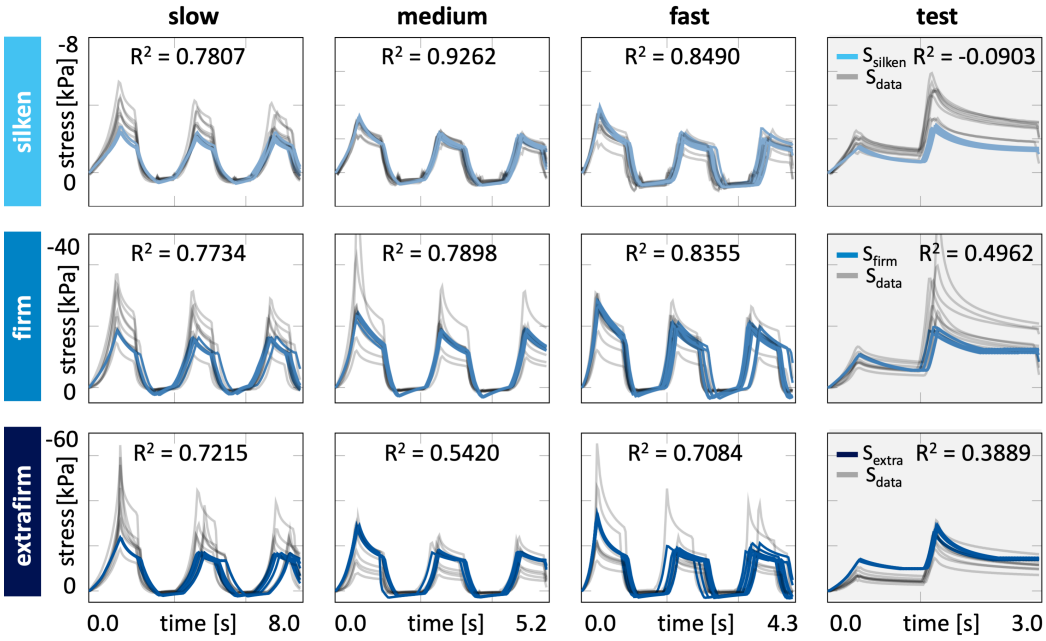


Figure 10: **Model discovery - Experimentally recorded and computationally predicted stresses of the three individual tofu models.** Stress histories $S(t)$ for silken (top), firm (middle), and extrafirm (bottom) tofu. The first three columns show the triple-compression tests used as training data. The 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.

potential,

$$\begin{aligned}\psi^{\text{eq}}_{\text{firm}}(\boldsymbol{C}) &= \tfrac{1}{2}\,\mu_3[\tilde{I}^{\boldsymbol{C}}_1 - 3]^2 \qquad\qquad + \tfrac{1}{2}\,\mu_5[(\tilde{I}^{\boldsymbol{C}}_2)^{3/2} - 3^{3/2}] \\ \psi^{\text{neq}}_{\text{firm}}(\boldsymbol{C}_{\text{e}}) &= \tfrac{1}{2}\,\mu_5[(\tilde{I}^{\boldsymbol{C}_{\text{e}}}_2)^{3/2} - 3^{3/2}] + \alpha_9[(I^{\boldsymbol{C}_{\text{e}}}_3)^{\beta_9} - 1 - \beta_9 \ln(I^{\boldsymbol{C}_{\text{e}}}_3)]\,,\end{aligned}$$

with network weights $\mu_3 = 5.250\,\text{kPa}$, $\mu_5 = 0.136\,\text{kPa}$ for the equilibrium part and $\mu_5 = 4.952\,\text{kPa}$, $\alpha_9 = 0.867\,\text{kPa}$, $\beta_9 = 1.439$ 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-invariant term for the non-equilibrium potential,

$$\begin{aligned}\psi^{\text{eq}}_{\text{extra}}(\boldsymbol{C}) &= \tfrac{1}{2}\,\mu_7[(\tilde{I}^{\boldsymbol{C}}_2)^{3/2} - 3^{3/2}]^2 + \alpha_9[(I^{\boldsymbol{C}}_3)^{\beta_9} - 1 - \beta_9 \ln(I^{\boldsymbol{C}}_3)] \\ \psi^{\text{neq}}_{\text{extra}}(\boldsymbol{C}_{\text{e}}) &= \tfrac{1}{2}\,\mu_5[(\tilde{I}^{\boldsymbol{C}_{\text{e}}}_2)^{3/2} - 3^{3/2}] + \tfrac{1}{2}\,\mu_7[(\tilde{I}^{\boldsymbol{C}_{\text{e}}}_2)^{3/2} - 3^{3/2}]^2\end{aligned}$$

with $\mu_7 = 0.151\,\text{kPa}$, $\alpha_9 = 0.103\,\text{kPa}$, $\beta_9 = 4.920$ for the equilibrium part and $\mu_5 = 8.160\,\text{kPa}$, $\mu_7 = 2.735\,\text{kPa}$ for the non-equilibrium part. Notably, the most prevalent invariant across all six potentials is the second invariant $\tilde{I}^{\boldsymbol{C}_{\text{e}}}_2$ featured in seven terms, followed by the third invariant $I^{\boldsymbol{C}_{\text{e}}}_3$ in three terms, and the first invariant $I^{\boldsymbol{C}_{\text{e}}}_3$ 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 Figure 4 to discover the dissipative potential $\phi$. Figures 11–13 display the three discovered potentials for silken, firm, and extrafirm tofu. The colors represent the magnitude of the final weights and biases and enable direct interpretation of active pathways. For silken tofu in Figure 11, the dissipative potential is driven by both the second and third deviatoric invariants $J^{\boldsymbol{\Sigma}}_2$ and $J^{\boldsymbol{\Sigma}}_3$,

$$\phi_{\text{silken}} = \max\,(\,1.653\,[\,\exp(\,1.770 \cdot \exp(\,1.837 \cdot |-0.058\,(J^{\boldsymbol{\Sigma}}_2)^{1/2} + 0.148\,(J^{\boldsymbol{\Sigma}}_3)^{1/3}\,|^{0.14}) - 1\,) - 1\,],0\,).$$

For firm tofu in Figure 12, the dissipative potential is driven exclusively by the second deviatoric invariant $J^{\boldsymbol{\Sigma}}_2$,

$$\phi_{\text{firm}} = \max\,(\,0.780\,[\,\exp(\,0.814 \cdot \exp(\,0.857 \cdot |\,0.457\,(J^{\boldsymbol{\Sigma}}_2)^{1/2}\,|^{0.036}) - 1\,) - 1\,],0\,).$$

For extrafirm tofu in Figure 13, the dissipative potential is driven exclusively by the third deviatoric invariant $J^{\boldsymbol{\Sigma}}_3$,

$$\phi_{\text{extra}} = \max\,(\,1.698\,[\,\exp(\,1.710 \cdot \exp(\,1.717 \cdot |\,0.156\,(J^{\boldsymbol{\Sigma}}_3)^{1/3}\,|^{1.42}) - 1\,) - 1\,],0\,).$$

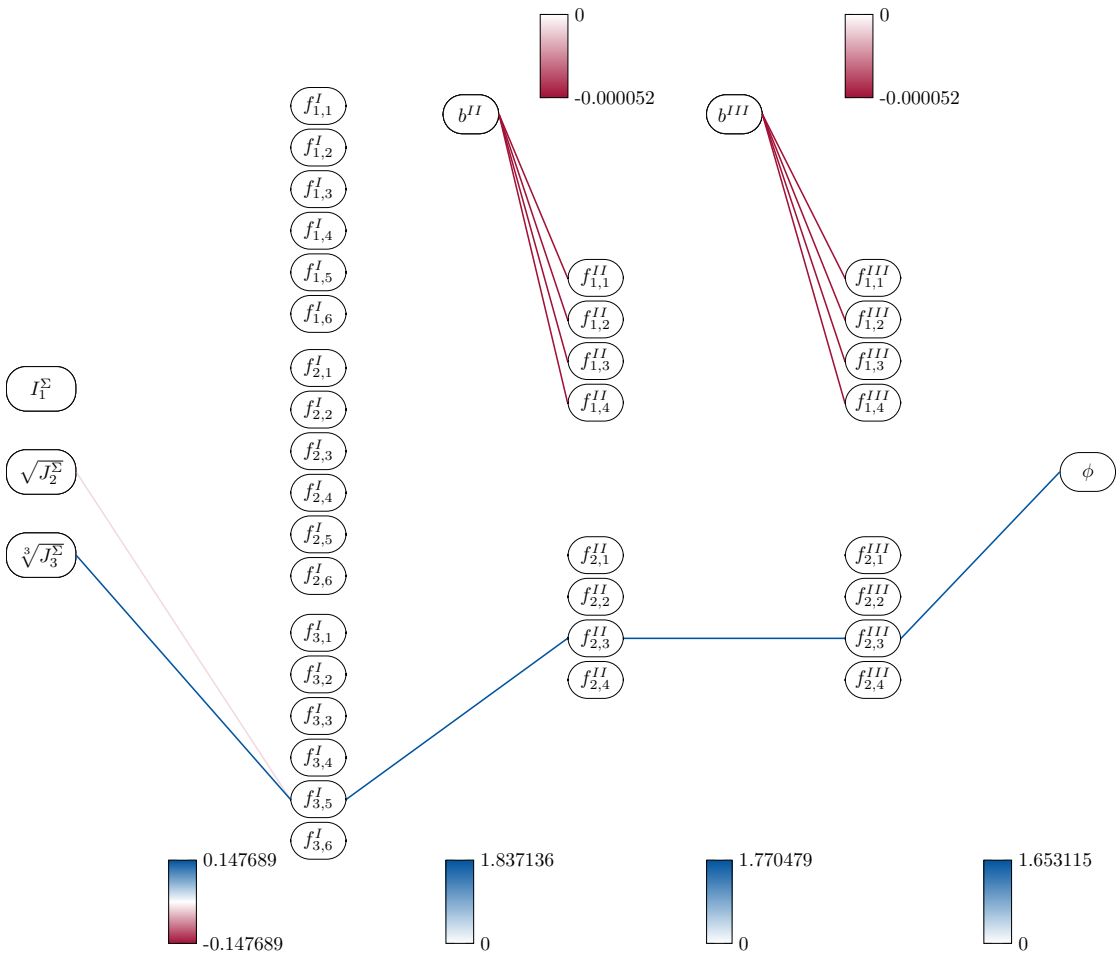


Figure 11: **Model discovery - Trained dissipative network for silken tofu.** The colors represent the magnitude of the final weights and biases, with blue indicating positive and red negative values, and white indicating that this path is inactive. The network discovers a single-term potential for silken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

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 $J_2^{\boldsymbol{\Sigma}}$ and $J_3^{\boldsymbol{\Sigma}}$ embedded in two nested exponential functions. None of the discovered potentials depends on the volumetric invariant $I_1^{\boldsymbol{\Sigma}}$, and none of them features any biases that would result in additive contributions.

Figure 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 $\lambda$, followed by the total stress $S$, its equilibrium and non-equilibrium parts $S^{\mathrm{eq}}$ and $S^{\mathrm{neq}}$, and the inelastic deformation rate $\dot{C}_{\mathrm{i}}$. The color shades denote contributions from the three invariants $\{\, \tilde{I}_1^{\boldsymbol{C}}, \tilde{I}_2^{\boldsymbol{C}}, I_3^{\boldsymbol{C}} \,\}$ and $\{\, \tilde{I}_1^{\boldsymbol{C}_{\mathrm{e}}}, \tilde{I}_2^{\boldsymbol{C}_{\mathrm{e}}}, I_3^{\boldsymbol{C}_{\mathrm{e}}} \,\}$ and $\{\, I_1^{\boldsymbol{\Sigma}}, J_2^{\boldsymbol{\Sigma}}, J_3^{\boldsymbol{\Sigma}} \,\}$. 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 $S^{\mathrm{eq}}$ 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 $S^{\mathrm{neq}}$ in the fourth row is almost entirely made up of second-invariant terms in $\tilde{I}_2^{\boldsymbol{C}}$ across all three tofu types. Similarly, the evolution of the inelastic right Cauchy tensor $\dot{C}_{\mathrm{i}}$ in the bottom row is made up only of second- and third-invariant terms in $J_2^{\boldsymbol{\Sigma}}$ and $J_3^{\boldsymbol{\Sigma}}$, suggesting that the inelastic deformation remains entirely deviatoric. Importantly, as the water content decreases, from left to right, the equilibrium contribution $S^{\mathrm{eq}}$ 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.** Figure 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,

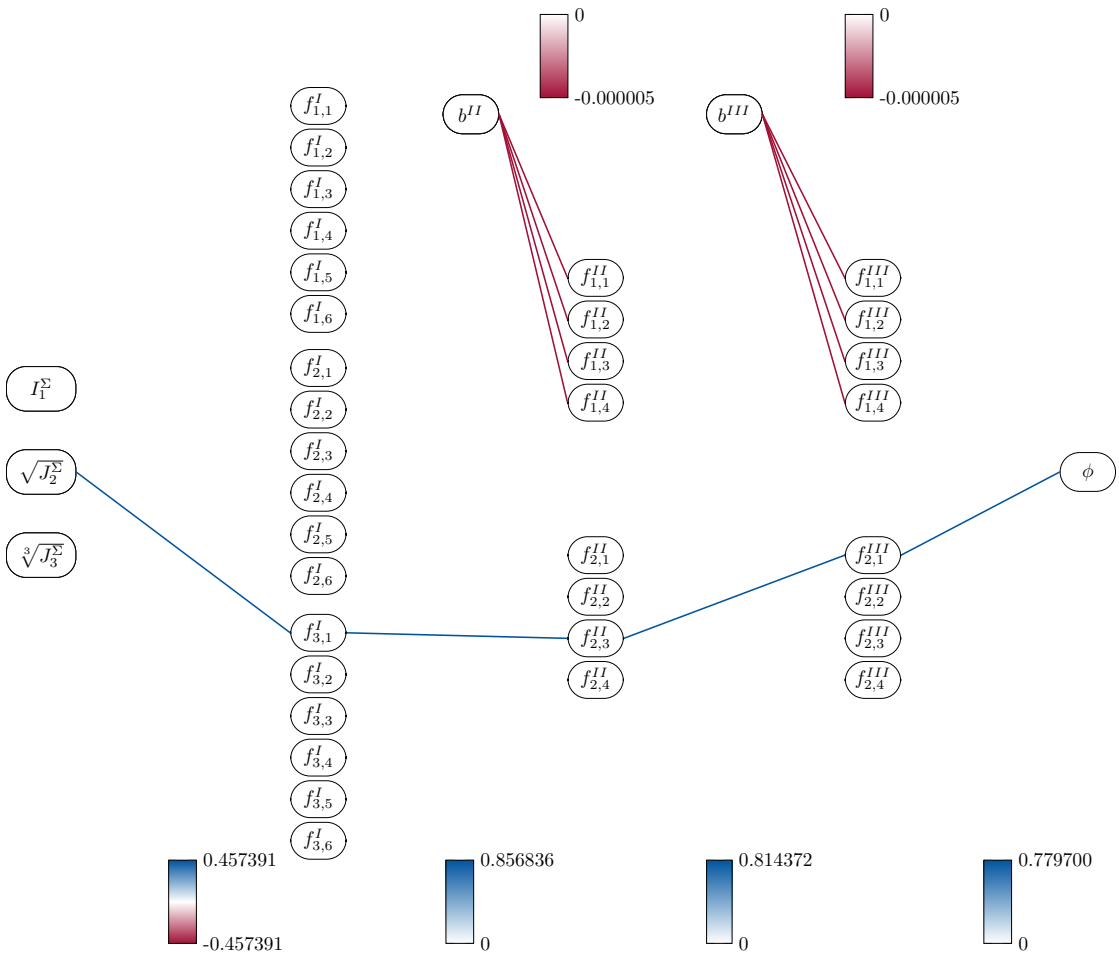


Figure 12: **Model discovery - Trained dissipative network for firm tofu.** The colors represent the magnitude of the final weights and biases, with blue indicating positive and red negative values, and white indicating that this path is inactive. The network discovers a single-term potential for firm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

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 $\psi^{\text{eq}}$ 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 $\psi^{\text{neq}}$ features water-content-dependent quadratic first-invariant and linear second-invariant terms, where both increase with water content. The dissipative potential $\phi$ features water-content-dependent exponential second- and third-invariant terms, where the first increases and the second decreases with water content.

Figure 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 Figure 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 $\mathsf{R}^2 = 0.6900$ compared to the single models with $\mathsf{R}^2 = 0.7449$, 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 $\mathsf{R}^2$-values. Despite this discrepancy, the silken tofu model captures the overall trends of nonlinear relaxation and decreasing stress magnitude with successive load cycles.

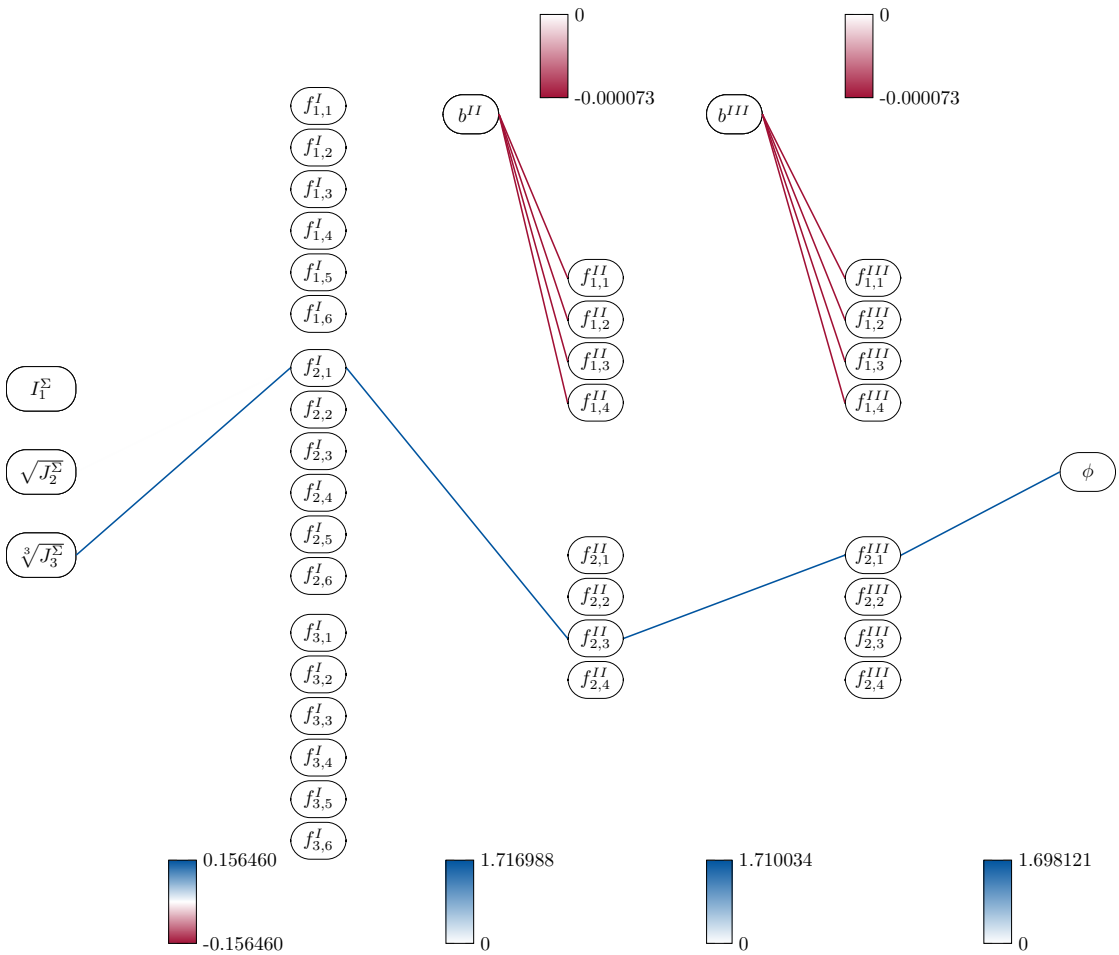


Figure 13: **Model discovery - Trained dissipative network for extrafirm tofu.** The colors represent the magnitude of the final weights and biases, with blue indicating positive and red negative values, and white indicating that this path is inactive. The network discovers a four-term potential for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

## 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 Figures 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 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 [44, 64]. 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 [10]. rather than from the fluid-filled matrix [25]. We therefore first adopt the *theory of finite viscoelasticity* [65, 69] 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. Figure 9 of the experimental data and Figure 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 [78]. This trend is visible in Figures 9 and 10, although less pronounced, since our loading rates only vary by a factor three. The *water-content dependence* is

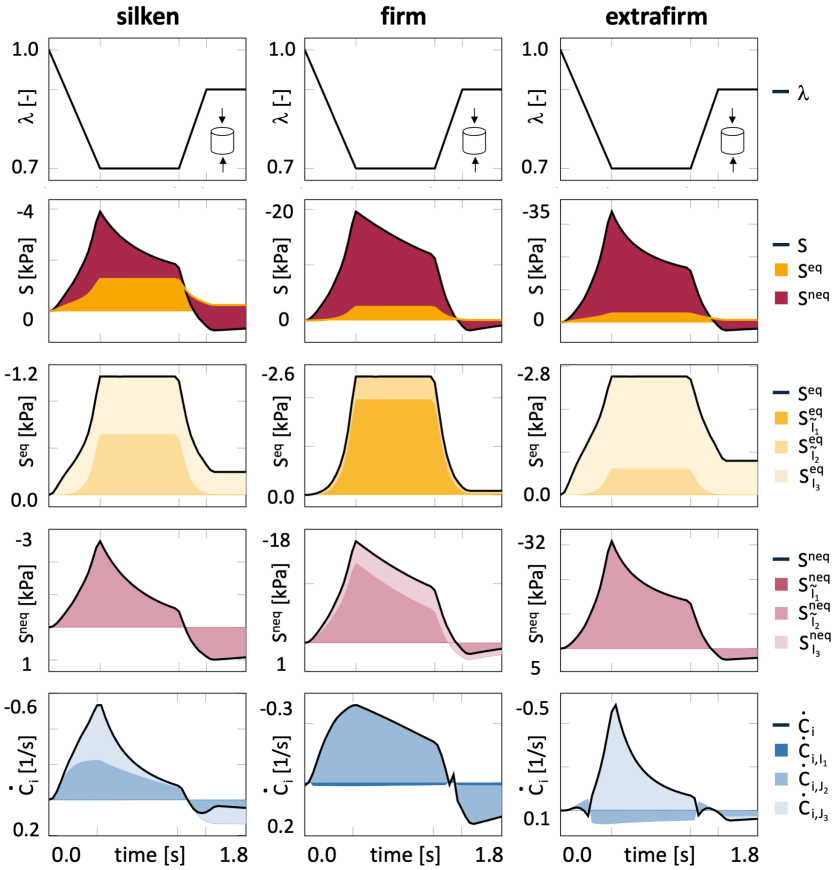


Figure 14: **Model discovery - Predicted stresses and inelastic deformation rates of the three individual tofu models.** Stress histories and evolution of the inelastic deformation rate for silken (left), firm (middle), and extrafirm (right) tofu for a strain rate of 60%/s. Applied stretches $\lambda$ (top); total stress $S$, equilibrium stress $S^{\mathrm{eq}}$, and non-equilibrium stress $S^{\mathrm{neq}}$ (middle), and inelastic deformation rate $\dot{C}_{\mathrm{i}}$ (bottom). Distinct patterns denote contributions from the three invariants $\{\,\tilde{I}_1^{\boldsymbol{C}}, \tilde{I}_2^{\boldsymbol{C}}, I_3^{\boldsymbol{C}}\,\}$ and $\{\,\tilde{I}_1^{\boldsymbol{C}_{\mathrm{e}}}, \tilde{I}_2^{\boldsymbol{C}_{\mathrm{e}}}, I_3^{\boldsymbol{C}_{\mathrm{e}}}\,\}$ and $\{\,I_1^{\boldsymbol{\Sigma}}, J_2^{\boldsymbol{\Sigma}}, J_3^{\boldsymbol{\Sigma}}\,\}$.

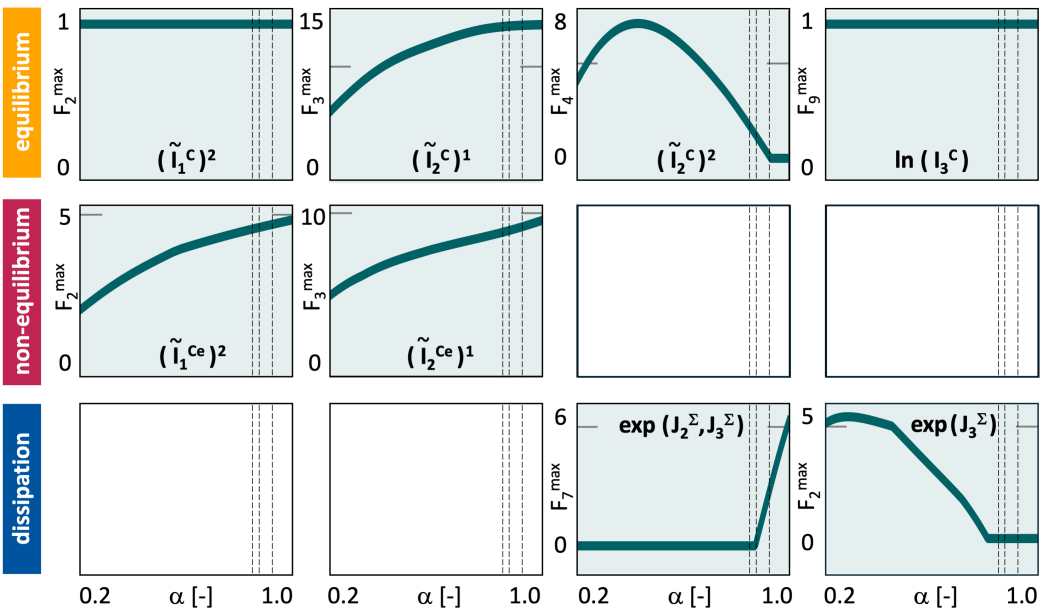


Figure 15: **Model discovery - Water-content dependency of the single unified tofu model.** Functional dependency of the equilibrium potential $\psi^{\mathrm{eq}}$ (top), non-equilibrium potential $\psi^{\mathrm{neq}}$ (middle), and dissipative potential $\phi$ (bottom) on the water content $\alpha$. Columns show the discovered dependencies on the individual terms. Dashed lines highlight the water contents of extrafirm, firm, and silken tofu.

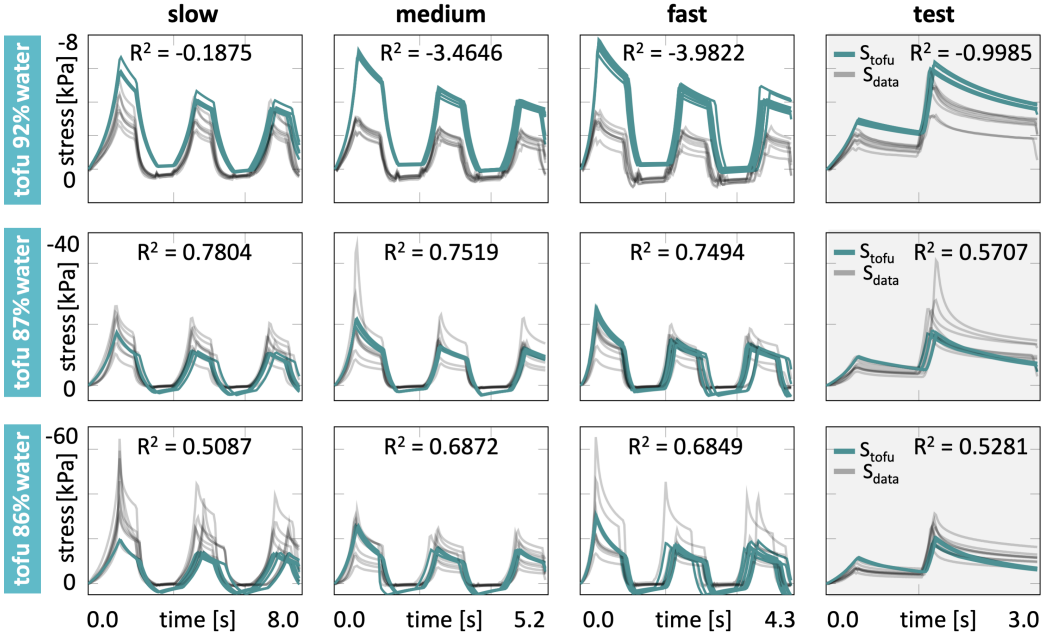


Figure 16: **Model discovery - Experimentally recorded and computationally predicted stresses of the single unified tofu model.** Stress histories $S(t)$ for silken (top), firm (middle), and extrafirm (bottom) tofu. The first three columns show the triple-compression tests used as training data. The 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.

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±16.22kPa. 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 [71] and with our rheologically measured storage moduli of 5.7±0.5 kPa and 11.0±1.0 kPa [22]. Our observations also agree with previous viscoelastic tests on lab-made tofu with different water contents [14], where the elastic, viscous, and relaxation parameters increased notably with decreasing water content [15]. We can potentially explain these observations with the *theory of finite poroelasticity* [19, 59], and propose to supplement the theory of finite viscoelasticity with volume fractions of the solid and fluid phases.

**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* [1, 28] made up of feed forward networks for the equilibrium and non-equilibrium potentials combined with a recurrent neural network for the dissipative potential [29]. 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 Figure 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 Figure 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 [14]. Our automated discovery process captures these Maxwell-type characteristics [29], 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 [33], $\{\tilde{I}_1^{\boldsymbol{C}}, \tilde{I}_2^{\boldsymbol{C}}, I_3^{\boldsymbol{C}}\}$ and $\{\tilde{I}_1^{\boldsymbol{C}_e}, \tilde{I}_2^{\boldsymbol{C}_e}, I_3^{\boldsymbol{C}_e}\}$. 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 Figure 3 to identify the most relevant contributions out of $2^9 = 512$ possible combinations of terms [28, 29]. 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*, $\tilde{I}_2^{\boldsymbol{C}}$ or $\tilde{I}_2^{\boldsymbol{C}_e}$, which measure the isochoric area change. The equilibrium potential $\psi^{\text{eq}}$ features a quadratic second-invariant term for silken and extrafirm tofu, and a linear second-invariant term for firm tofu. The non-equilibrium potential $\psi^{\text{neq}}$ 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 [33], which are not present in tofu. The *dominance of the second invariant* [42] is in line with previously discovered models for human brain [48] and plant-based meat [50], 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, 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 [15]. 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 dissipatve potentials in terms of *one volumetric invariant* and *two deviatoric invariants* of the Mandel stress [28], $\{\, I_1^{\boldsymbol{\Sigma}}, J_2^{\boldsymbol{\Sigma}}, J_3^{\boldsymbol{\Sigma}} \,\}$. Yet, their individual contributions to the dissipative potential are unknown. To discover which invariants drive the dissipative responses, we consult our dissipative network in Figure 4. Strikingly, across all three tofu types, the discovered dissipative potentials $\phi$ feature a single term that contains weighted powers of the second and third deviatoric invariants, $J_2^{\boldsymbol{\Sigma}}$ and $J_3^{\boldsymbol{\Sigma}}$, embedded in two nested exponential functions. None of the discovered potentials depends on the volumetric invariant $I_1^{\boldsymbol{\Sigma}}$, 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 [56]: The first stress invariant $I_1^{\boldsymbol{\Sigma}}$ measures the hydrostatic stress, and its absence in the dissipative potential indicates a *pressure insensitivity*, unlike in the classical Drucker-Prager model [21]. The second stress invariant $J_2^{\boldsymbol{\Sigma}}$ measures the magnitude of shear, and its presence in the dissipative potential indicates a *shear sensitivity*, like in the classical von Mises model [57]. The third stress invariant $J_3^{\boldsymbol{\Sigma}}$ 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 [73] and Mohr-Coulomb models [16, 58]. 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 [51] that depends on both the second and third invariants and distinguishes between tension-shear, pure shear, and compression-shear of the same shear magnitude [8]. 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 [68]. 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 [15]. 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 [59], adopt the concept of volume fractions [18], associate the equilibrium behavior with the solid and the non-equilibrium behavior with the fluid, and weigh both with their individual volume fractions [23], $S = \nu_{\text{s}} \cdot S^{\text{eq}} + \nu_{\text{f}} \cdot S^{\text{neq}}$. In Table 1, the volume fractions of the fluid and solid are $\nu_{\text{f}} = \{\, 0.92, 0.87, 0.86 \,\}$ and $\nu_{\text{s}} = 1 - \nu_{\text{f}} = \{\, 0.08, 0.13, 0.14 \,\}$, 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 Figure 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 [22]. We conclude that volume-fraction-based scaling might be appropriate for linear mixtures with low-to-moderate water contents [19], 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 $\alpha$ as input, and outputs the weights for each node of the three mechanical networks. Our results in Figure 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 Figure 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 [11], nor a purely poroelastic model [25] can accurately capture this characteristic behavior, but that a poroviscoelastic model is needed instead [36]. 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 [27]. 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 [33, 65, 69]. 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 [77], we have previously tested firm and extrafirm tofu in compression and shear [22], and in tension, compression, and shear [71]; 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 [77]. Fourth, we have assumed that water in tofu exists only in a single state [19], freely moving. In reality, tofu is a hydrated soy protein gel that contains water in three distinct states, *bound*, *immobilized*, and *free* [83], where bound and immobilized water modulate the elastic response, while free water modulates the inelastic response [52]. A next logical step would be to characterize the volume fractions of bound, immobilized, and free water [36, 37], 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 [9, 24, 47], 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 [38, 40, 54].

## 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.

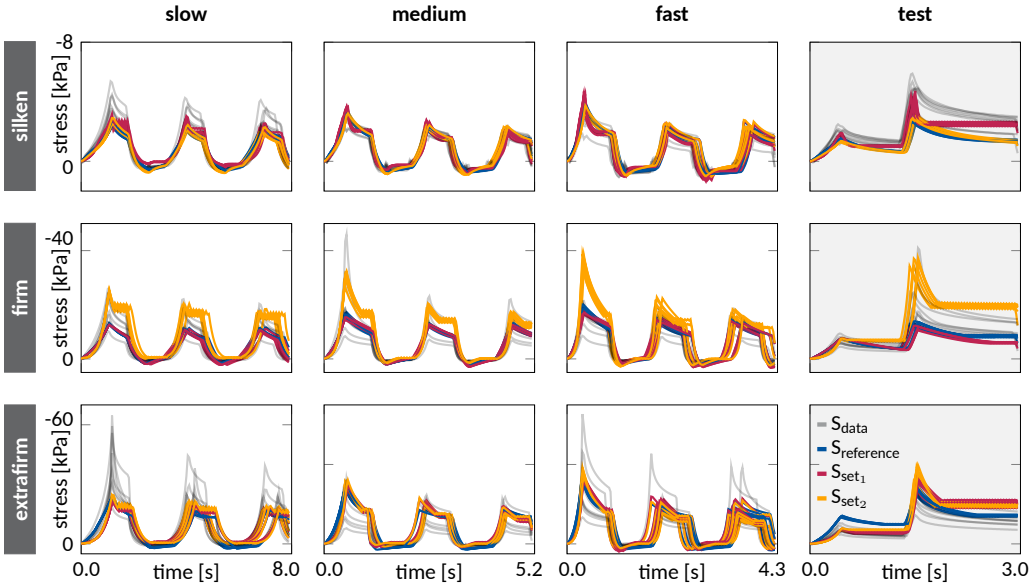


Figure B.17: **Preliminary sensitivity analysis for varying network architectures.** Stress histories $S(t)$ for silken (top), firm (middle), and extrafirm (bottom) tofu. Experimental data are plotted in gray; reference data from Figure 10 with {18, 8, 8} neurons per layer are plotted in blue; predictions with {30, 16, 12} and {4, 4, 2} neurons per layer are shown in red as set 1 and in orange as set 2.

## 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 $n_{\text{data}}$. In addition to the coefficients of determination $\mathsf{R}^2$ reported in Figures 10 and 16, we report the mean absolute error, $\mathsf{MAE} = \sum_{m=1}^{n_{\text{data}}} |S_m - \hat{S}_m| / n_{\text{data}}$, the root mean squared error, $\mathsf{RMSE} = [\sum_{m=1}^{n_{\text{data}}} [S_m - \hat{S}_m]^2 / n_{\text{data}}]^{1/2}$, and normalized root mean squared error, $\mathsf{NRMSE} = \mathsf{RMSE}/\max\{\hat{S}\}$, where $\max\{\hat{S}\}$ denotes the maximum experimental stress. Table A.2 summarizes the errors for the discovered individual tofu models in Figure 10 and for the discovered unified tofu model in Figure 16.

Table A.2: **Errors of discovered models.** Mean absolute error MAE, root mean squared error RMSE, and normalized root mean squared error NRMSE for the discovered individual tofu models in Figure 10, left, and for the discovered unified tofu model in Figure 16, right.

| individual | error | slow | medium | fast | test |
|---|---|---|---|---|---|
| silken | RMSE | 0.633 | 0.298 | 0.460 | 1.206 |
| | MAE | 0.402 | 0.223 | 0.322 | 0.970 |
| | NRMSE | 0.118 | 0.095 | 0.133 | 0.205 |
| firm | RMSE | 2.911 | 2.891 | 2.476 | 4.203 |
| | MAE | 1.871 | 1.684 | 1.625 | 2.707 |
| | NRMSE | 0.112 | 0.063 | 0.113 | 0.103 |
| extrafirm | RMSE | 5.746 | 4.877 | 5.200 | 4.644 |
| | MAE | 3.400 | 3.386 | 3.659 | 3.922 |
| | NRMSE | 0.089 | 0.147 | 0.080 | 0.148 |

| unified | error | slow | medium | fast | test |
|---|---|---|---|---|---|
| silken | RMSE | 1.473 | 2.318 | 2.643 | 1.633 |
| | MAE | 1.182 | 1.893 | 2.150 | 1.435 |
| | NRMSE | 0.274 | 0.742 | 0.762 | 0.277 |
| firm | RMSE | 2.867 | 3.140 | 3.056 | 3.880 |
| | MAE | 2.131 | 2.231 | 2.321 | 2.728 |
| | NRMSE | 0.111 | 0.068 | 0.141 | 0.095 |
| extrafirm | RMSE | 7.632 | 4.030 | 5.405 | 4.080 |
| | MAE | 4.982 | 3.290 | 3.642 | 3.297 |
| | NRMSE | 0.118 | 0.122 | 0.083 | 0.131 |

## 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.** Figure B.17 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 $\mathsf{R}^2 = 0.7449$.

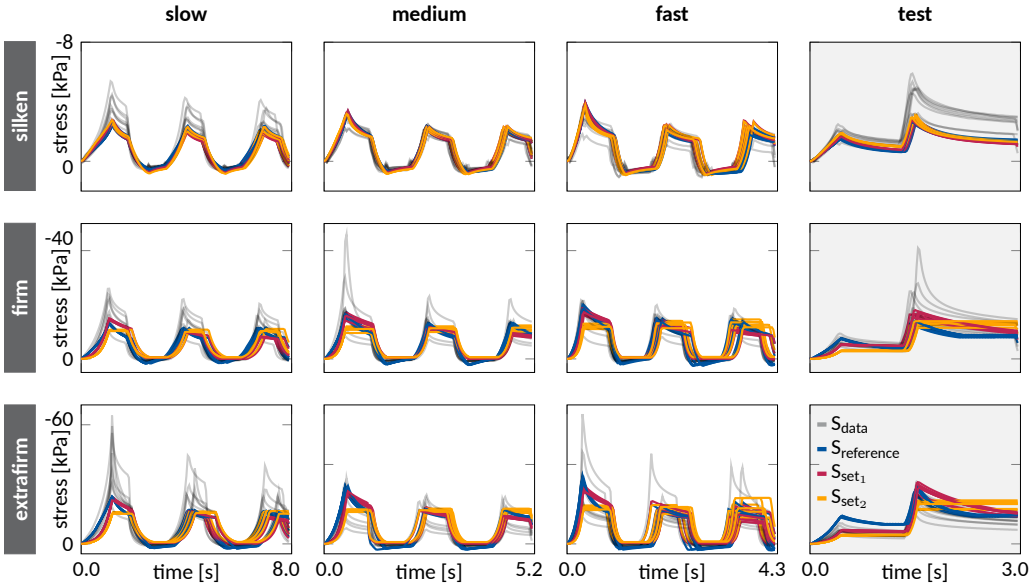


Figure B.18: **Preliminary sensitivity analysis for varying activation functions.** Stress histories $S(t)$ for silken (top), firm (middle), and extrafirm (bottom) tofu. Experimental data are plotted in gray; reference data from Figure 10 with trainable parameter $p$ are plotted in blue; predictions with fixed parameters $p = 1$ and $p = 2$ are shown in red as set 1 and in orange as set 2.

The finer {30, 16, 12} network results in a reduced goodness of fit of $\mathsf{R}^2 = 0.7085$, associated with the stress predictions in red. The coarser {4, 4, 2} network results in a similarly reduced goodness of fit of $\mathsf{R}^2 = 0.7091$, 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.** Figure B.18 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 $\mathsf{R}^2 = 0.7449$. Fixing the powers to $p = 1$ and $p = 2$ reveals similar trends and a notable reduction of the goodness of fit to $\mathsf{R}^2 = 0.7212$ for $p = 1$, associated with the red curves, and even further to $\mathsf{R}^2 = 0.6586$ 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 [47, 48, 49] and dissipative [28, 30, 31] networks similar to Figures 3 and 4 that have demonstrated high robustness and reproducibility.

**Regularization parameters.** Figure B.19 summarizes our preliminary sensitivity analysis for varying regularization coefficients $\alpha_1$ and $\alpha_2$ 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 $\mathsf{R}^2 = 0.7449$ Higher regularization coefficients of $\alpha_1 = \alpha_2 = 0.001$ result in a reduced goodness of fit of $\mathsf{R}^2 = 0.6556$ 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 $\mathsf{R}^2 = 0.7510$ as the reference model, with stress predictions, shown in orange, that predict similar features as the reference model, shown in blue.

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.

*Acknowledgements*

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.

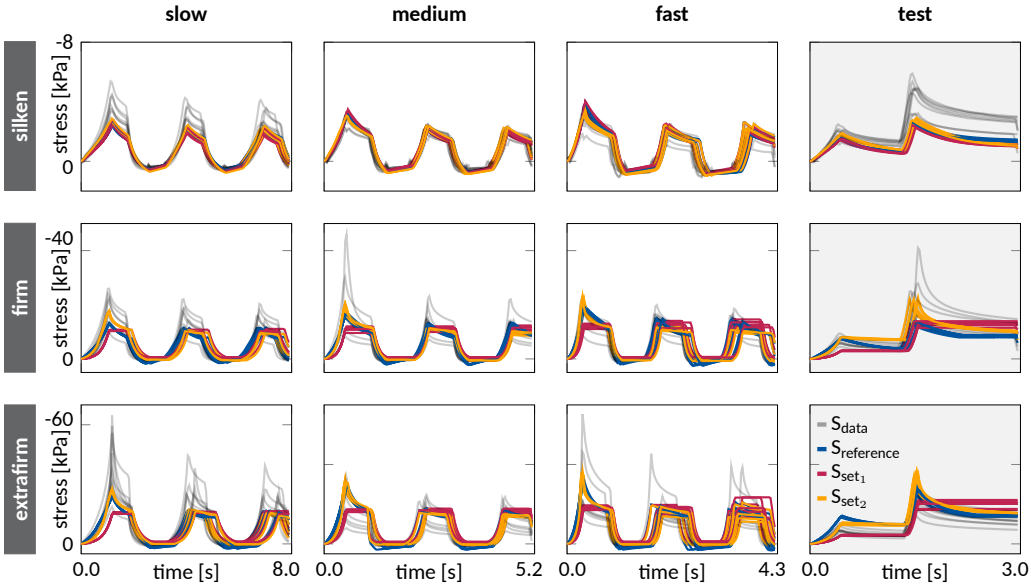


Figure B.19: **Preliminary sensitivity analysis for varying regularization parameters.** Stress histories $S(t)$ for silken (top), firm (middle), and extrafirm (bottom) tofu. Experimental data are plotted in gray; reference data from Figure 10 with regularization parameters $\alpha_1 = \alpha_2 = 0.0001$ 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.

*CRediT authorship contribution statement*

BB: Conceptualization, Methodology, Software, Formal analysis, Data Curation, Investigation, Validation, Writing Original Draft, Writing Review and Editing. JWS: Funding acquisition, Writing Review and Editing. HH: Conceptualization, Methodology, Software, Formal analysis, Validation, Writing Review and Editing. EK Conceptualization, Resources, Writing Review and Editing.

*Data availability*

Our source code and examples are available at https://doi.org/10.5281/zenodo.16993236.

## References


[1] Abdolazizi, K.P., Linka, K., Cyron, C.J., 2024. Viscoelastic Constitutive Artificial Neural Networks (VCANNs) – a framework for data-driven anisotropic nonlinear finite viscoelasticity. J. Comput. Phys. 499, 112704.

[2] 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. PMLR 70, 146–155.

[3] 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.

[4] Boes, B., Simon, J.W., Holthusen, H., 2024. Accounting for plasticity: An extension of inelastic constitutive artificial neural networks. Eur. J. Mech. A/Solids. 117. 105998.

[5] Bourne, M.C., 1968. Texture profile of ripening pears. J. Food Sci. 33 (2), 223–226.

[6] Bowen, R.M., 1976. Continuum Physics, Chapter Theory of Mixtures, Volume 3, 1-127, Academic Press, New York.

[7] Bowen, R.M., 1980. Incompressible porous media models by use of the theory of mixtures. Int. J. Eng. Sci. 18 (9), 1129–1148.

[8] Bruhns, O.T., 2024. Some remarks on Lode angle and Lode parameter. Proc. Appl. Math. Mech. 24: e202400004.

[9] Brunton, S.L., Proctor, J.P., Kutz, J.N., 2016. Discovering governing equations from data by sparse identification of nonlinear dynamical systems. Proc. Nat. Acad. Sci. 113:3932-3937.

[10] Budday, S., Sommer, G., Haybaeck, J., Steinmann, P., Holzapfel, G.A., Kuhl, E., 2017. Rheological characterization of human brain tissue. Acta Biomat. 60, 315–329.

[11] Budday, S., Sommer, G., Holzapfel, G.A., Steinmann, P., Kuhl, E., 2017. Viscoelastic parameter identification of human brain tissue. J. Mech. Beh. Biomed. Mat. 74, 463-476.

[12] Budday, S., Oveart, T.C., Holzapfel, G.A., 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.

[13] Chen, C.C., Hsieh, J.F., Kuo, M.I., 2023. Insight into the processing, gelation and functional components of tofu: a review. Processes 11 (1), 202.

[14] Chen, Y., Shimizu, N., Kimura, T., 2002. Texture evaluation of tofu (soybean curd) be viscoelastic tests. J. Japanese Soc. Agricult. Mach. 64 (6), 137-144.

[15] 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.

[16] Coulomb, C.A., 1776. Essai sur une application des règles des maximis et minimis à quelques problèmes de statique relatifs à l'architecture. Mémoires de l'Académie Royale des Sciences. 7, 343–387.
[17] Datta, B., Buehler, M.J., Chow, Y., Gligoric, K., Jurafsky, D., Kaplan, D.L., Ledesma-Amaro, R., Del Missier, G., Neidhardt, L., Pichara, K., Sanchez-Lengeling, B., Schlangen, M., St. Pierre, S., Tagkopoulos, I., Thomas, A., Watson, N., Kuhl, E., 2025. AI for sustainable future foods. arXiv, doi:10.48550/arXiv.2509.21556.
[18] de Boer, R., Ehlers, W., 1990. The development of the concept of effective stresses. Acta Mech. 83, 77-92.
[19] de Boer, R., 2020. Theory of Porous Media: Highlights in Historical Development and Current State. Springer, Berlin.
[20] Dekking, F.M., 2005. A modern introduction to probability and statistics: understanding why and how. Springer Science & Business Media.
[21] Drucker, D.C., Prager, W., 1952. Soil mechanics and plstic analysis or limit design. Quart. Appl. Math. 10, 157-165.
[22] Dunne, R.A., Darwin, E.C., Perez Medina, V.A., Levenston, M.E., St. Pierre, S.R., Kuhl, E., 2025. Texture profile analysis and rheology of plant-based and animal meat. Food Res. Int. 205, 115876.
[23] Ehlers, W., 1993. Constitutive equations for granular materials in geomechanical context. In: Continuum Mechanics in Environmental Sciences and Geophysics. Springer, 313–402.
[24] Flaschel, M., Kumar, S., De Lorenzis, L., 2021. Unsupervised discovery of interpretable hyperelastic constitutive laws. Comput. Methods Appl. Mech. Eng. 381, 113852.
[25] Franceschini, G., Bigoni, D. Regitnig, P., Holzapfel, G.A., 2006. Brain tissue deforms similarly to filled elastomers and follows consolidation theory. J. Mech. Phys. Solids 54, 2592–2620.
[26] Friedman, H.H., Whitney, J.E., Szczesniak, A.S., 1963. The texturometer – a new instrument for objective texture measurement. J. Food Sci. 28 (4), 390–396.
[27] Holmes, M.H., Mow, V.C., 1990. The nonlinear characteristics of soft gels and hydrated connective tissues in ultrafiltration. J. Biomech. 23: 1145-1156.
[28] Holthusen, H., Lamm, L., Brepols, T., Reese, S., Kuhl, E., 2024. Theory and implementation of inelastic constitutive artificial neural networks. Comput. Methods Appl. Mech. Eng. 428, 117063.
[29] Holthusen, H., Lamm, L., Brepols, T., Reese, S., Kuhl, E., 2024. Polyconvex inelastic constitutive artificial neural networks. PAMM 24 (3), e202400032.
[30] Holthusen, H., Brepols, T., Linka, K., Kuhl, E., 2025. Automated model discovery for tensional homeostasis: constitutive machine learning in growth and remodeling. Comput. Biol. Med. 186, 109691.
[31] Holthusen, H., Linka, K., Kuhl, E., Brepols, T., 2025. A generalized dual potential for inelastic constitutive artificial neural networks: a JAX implementation at finite strains. J. Mech. Phys. Solids. 206: 106337.
[32] Holthusen, H., Kuhl, E., 2026. A complement to neural networks for anisotropic inelasticity at finite strains. Comput. Methods Appl. Mech. Eng. 450, 118612.
[33] Holzapfel, G.A., 2000. Nonlinear Solid Mechanics. John Wiley & Sons, New York.
[34] 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.
[35] Hudobivnik, B., Korelc, J., 2016. Closed-form representation of matrix functions in the formulation of nonlinear material models. Finite Elem. Anal. Des. 111, 19–32.
[36] Kainz, M.P., Greiner, A. Hinrichsen, J., Kolb, D., Comellas, Steinmann, P., Budday, S., Terzano, M., Holzapfel, G.A., 2023. Poro-viscoelastic material parameter identification of brain-tissue mimicking hydrogels. Front. Bioeng. Biotech. 11, 1143304.
[37] Kainz, M.P., Polz, M., Terzano, M., Holzapfel, G.A., 2025. Biomechanical and biological characterization of brain tissue-inspired hydrogels: Focus on porous material properties for tissue engineering applications. Proceedings of the 12th European Solid Mechanics Conference ESMC 2025, Lyon, France.
[38] Kalina, K.A., 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.
[39] Kalina, K.A., 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.
[40] Klein, D.K., Roth, F.J., Valizadeh, I., Weeger, O., 2023. Parametrized polyconvex hyperelasticity with physics-augmented neural networks. Data-Centric Eng. 4, e25.
[41] Kohyama, K., Sano, Y., Doi, E., 1995. Rheological characteristics and gelation mechanism of tofu (soybean curd). J. Agric. Food Chem. 43 (7), 1808–1812.
[42] 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.
[43] Kuhl, E., 2025. AI for Food: Accelerating and democratizing discovery and innovation. npj Sci. Food. 9: 82.
[44] Lee, Y.C., Rosenau, J.R., Peleg, M., 1983. Rheological characterization of tofu. J. Texture Studies. 14 (2), 143-154.
[45] 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.
[46] Linka, K., Hillgärtner, M., Abdolazizi, K.P., Aydin, R.C., Itskov, M., Cyron, C.J., 2021. Constitutive artificial neural networks: a fast and general approach to predictive data-driven constitutive modeling by deep learning. J. Comput. Phys. 429, 110010.
[47] Linka, K., Kuhl, E., 2023. A new family of constitutive artificial neural networks towards automated model discovery. Comput. Methods Appl. Mech. Eng. 403, 115731.
[48] Linka, K., St. Pierre, S.R., Kuhl, E., 2023. Automated model discovery for human brain using constitutive artificial neural networks. Acta Biomater. 160, 134–151.
[49] Linka, K., Tepole, A.B., Holzapfel, G.A., Kuhl, E., 2023. Automated model discovery for skin: discovering the best model, data, and experiment. Comput. Methods Appl. Mech. Eng. 410, 116007.
[50] 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.
[51] Lode, W. 1925. Versuche, über den Einfluß der mittleren Hauptspannung auf die Fließgrenze. Z. Ang. Math. Mech. 5: 142-144.
[52] Lucey, J.A., 2002. Formation and physical properties of milk protein gels. J. Dairy Sci. 85: 281-294.
[53] Martonová, D., Peirlinck, M., Linka, K., Holzapfel, G.A., 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.
[54] Masi, F., Stefanou, I., Vannucci, P., Maffi-Berthier, V., 2021. Thermodynamics-based artificial neural networks for constitutive modeling.

J. Mech. Phys. Solids 147, 104277.
[55] MATLAB, 2022. MATLAB version 9.13.0 (R2022b). The MathWorks, Inc., Natick, Massachusetts. https://www.mathworks.com.
[56] McCulloch, J.A., St. Pierre, S.R., Linka, K., Kuhl, E., 2024. On sparse regression, Lp-regularization, and automated model discovery. Int. J. Num. Meth. Eng. 125: e7481.
[57] von Mises, R., 1913. Mechanik der festen Körper im plastisch deformablen Zustand. Nachr. Königl. Ges. Wiss. Göttingen, Math. Phys. Kl. 582-592, Weidmann.
[58] Mohr, O. 1900. Über die Darstellung des Spannungszustandes und des Bruchverhaltens fester Körper. Zeitschrift des Vereines Deutscher Ingenieure. 44, 1524–1527, 1572–1577.
[59] Mow, C.V., Holmes, M.H., Lai, W.M., 1984. Fluid transport and mechanical properties of articular cartilage: a review. J. Biomechanics 17, 377-394.
[60] Paredes, J., Cortizo-Lacalle, D., Imaz, A.M., Aldazabal, J., Vila, M., 2022. Application of texture analysis methods for the characterization of cultured meat. Sci. Rep. 12 (1), 77785.
[61] Peirlinck, M., Linka, K., Kuhl, E., 2025. Atrial constitutive neural networks. In: Proc. Int. Conf. Functional Imaging and Modeling of the Heart. Springer, 249–259.
[62] Poore, J., Nemecek, T., 2018. Reducing food's environmental impacts through producers and consumers. Science 360, 987–992.
[63] Qin, P., Wang, T., Luo, Y., 2022. A review on plant-based proteins from soybean: Health benefits and soy product development. J. Agricult. Food Ress 7, 100265.
[64] Rao, M.A., Steffe, J.F., 1992. Viscoelastic Properties of Foods Elsevier Applied Science, London.
[65] Reese, S., Govindjee, S., 1998. A theory of finite viscoelasticity and numerical aspects. Int. J. Solids Structures. 35, 3455-3482.
[66] Rizzo, G., Baroni, L., 2018. Review: Soy, soy foods and their role in vegetarian diets. Nutrients, 10, 43.
[67] Schreuders, F.K.G., Schlangen, M., Kyriakopoulou, K., Boom, R.M., van der Goot, A.J., 2021. Texture methods for evaluating meat and meat analogue structures: a review. Food Control 127, 108103.
[68] Shurtleff, W., Aoyagi, A., 1975. The Book of Tofu: Protein Source of the Future. Soyinfo Center, Lafayette.
[69] Simo, J.C., 1987. On a fully three-dimensional finite-strain viscoelastic damage model: Formulation and computational aspects. Comput. Methods Appl. Mech. Eng. 60, 153-173.
[70] St. Pierre, S.R., Rajasekharan, D., Darwin, E.C., Linka, K., Levenston, M.E., Kuhl, E., 2023. Discovering the mechanics of artificial and real meat. Comput. Methods Appl. Mech. Eng. 415, 116236.
[71] St. Pierre, S.R., Darwin, E.C., Adil, D., Aviles, M.C., Date, A., Dunne, R.A., Lall, Y., Parra Vallecillo, M., Perez Medina, V.A., Linka, K., Kuhl, E., 2024. The mechanical and sensory signature of plant-based and animal meat. npj Sci. Food 8 (1), 94.
[72] Tac, V., Gardner, C., Kuhl, E., 2026. Generative Artificial Intelligence creates delicious, sustainable, and nutritious burgers. arXiv. doi:10.48550/arXiv.2602.03092.
[73] Tresca, H. 1864. Mémoire sur l'écoulement des corps solides soumis à de fortes pressions. Comptes Rendus de l'Académie des Sciences. 59, 754–758.
[74] Tsai, S.J., Lan, C.Y., Kao, C.S., Chen, S.C., 1981. Studies on the yield and quality characteristics of tofu. J. Food Sci. 46 (6), 1234–1737.
[75] Tukey, J.W., 1977. Exploratory data analysis. Springer, 1–688.
[76] van der Velden, T., Boes, B., Brepols, T., Kuhl, E., Holthusen, H., 2025. A note on constitutive artificial neural networks for finite strain plasticity. SSRN preprint 5588650.
[77] Vervenne, T., Peirlinck, M., Famaey, N., Kuhl, E., 2025. Constitutive neural networks for main pulmonary arteries: discovering the undiscovered. Biomech. Model. Mechanobiol. 24: 615-634.
[78] Vervenne, T., St. Pierre, S.R., Famaey, N., Kuhl, E., 2025. Probing mycelium mechanics and taste: The moist and fibrous signature of fungi steak. Acta Biomat. 202, 341-351.
[79] Vladimirov, I.N., Pietryga, M.P., Reese, S., 2008. On the modeling of non-linear kinematic hardening at finite strains with application to springback – comparison of time integration algorithms. Int. J. Numer. Methods Eng. 75 (1), 1–28.
[80] Watanabe, T., 1997. The Science of Tofu. Food Journal Co., Kyoto.
[81] 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.
[82] Wang, L.M., Linka, K., Kuhl, E., 2023. Automated model discovery for muscle using constitutive recurrent neural networks. J. Mech. Behav. Biomed. Mat. 145: 106021.
[83] Yuan, Y., Zhang, Q., Lin, S., Li, J., 2025. Water: The soul of hydrogels. Prog. Mat. Sci. 148: 101378.
[84] Zheng, L., Regenstein, J.M., 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.