PAMM

WILEY-VCH

SPECIAL ISSUE 95th Annual Meeting of the International Association of Applied Mathematics and Mechanics (GAMM)

PAMM Proceedings in Applied Mathematics and Mechanics

RESEARCH ARTICLE OPEN ACCESS

Effective Material Modeling of Parameterizable Viscoelastic Shell Structures with Artificial Neural Networks

Jeremy Geiger | Werner Wagner | Steffen Freitag

Institute for Structural Analysis, Karlsruhe Institute of Technology (KIT), Karlsruhe, Germany

Correspondence: Jeremy Geiger (jeremy.geiger@kit.edu)

Received: 23 June 2025 | **Revised:** 20 October 2025 | **Accepted:** 24 October 2025

ABSTRACT

In this contribution, we seek to reduce the computational cost of structural simulations by taking advantage of a consistent homogenization scheme for shell representative volume elements to train an artificial neural network (ANN) material model for approximating the effective strain–stress relation. As ANNs can capture inelastic material behavior, this work focuses on the application to high-dimensional strain–stress relations for shell structures with underlying viscoelastic microstructures. We demonstrate that a small database comprising uniaxial synthetic material tests on a representative microstructure, in combination with derivative information, is sufficient to train a feasible ANN material model. Furthermore, we explore the limits of the approximation capabilities of the ANN by including non-strain-related parameters of the microstructure—such as volume fractions—as inputs to the material model. Studies include comparisons with full-scale and multiscale models, highlighting computational efficiency and practical feasibility in application to real-world engineering problems involving complex microstructures.

1 | Introduction

The growing demand for high-performance materials has led to significant advancements in engineering composite structures, such as laminates, fiber-reinforced materials, and novel developments such as voided slabs for reinforced concrete. Although these materials optimize structural performance by taking advantage of the unique properties of the different constituents, developing effective material models based on classical phenomenological approaches remains challenging. Full 3D solutions, on the other hand, under consideration of the complex microstructures, quickly encounter the bottleneck of a high computational effort.

As an alternative, multiscale techniques based on consistent homogenization schemes, which effectively capture the microstructural constitutive response at each macroscopic integration point, have been employed. Although this approach is highly feasible for linear elastic structures—where the micro problem only must be solved once—they become computationally expensive for inelastic microstructures due to repeated microscale evaluations. To overcome this limitation, one can train an artificial neural network (ANN) material model on the effective strain–stress data, which will be the approach adopted in the present work. Building on early works in neural network-based material modeling [1–3], many advances have been made in recent decades, particularly in modeling inelastic material behavior [4, 5] and integrating ANNs within multiscale frameworks [6, 7]. Recent efforts have focused on including physical requirements, that is, thermodynamic consistency, either in a weak sense via penalization [8, 9] or in a strong sense by embedding the constraints directly into the ANN architecture [10–12]. For an overview on neural network material modeling, the reader is referred to [13].

In this work, we train an ANN on effective viscoelastic strain–stress paths obtained from a shell representative volume element (RVE) to gain significant computational advantages for structural simulations. Besides the time and history dependence of the material, the ANN additionally learns the effective response of the material under consideration of varying geometric features—specifically, the radius of a void within the matrix. We show that the proposed approach yields accurate results at both the material point and structural level. Furthermore, we present a simple sampling strategy that enables efficient training of the ANN, even in the high-dimensional input–output space.

2 | Numerical Homogenization for Shell RVEs

For the derivation of the effective strain–stress relations for shell RVEs, we introduce the shell's reference surface Ω_0 occupying the three-dimensional Euclidean space $\mathcal{B}$ with shell height h in the reference configuration. Displacement boundary conditions are prescribed on Γ_u and stress boundary conditions on Γ_σ, with $\Gamma = \Gamma_\sigma \cup \Gamma_u$ being the boundary of the reference surface Ω_0 and $\Gamma_\sigma \cap \Gamma_u = \emptyset$. At each material point $\mathbf{X}$, we identify the general non-linear Green–Lagrangian vector of shell strains $\boldsymbol{\varepsilon}$ and the work conjugate 2. Piola–Kirchhoff shell stress resultants

$$\boldsymbol{\varepsilon}(\mathbf{X}, \mathbf{v}, t) = [\varepsilon_{11}, \varepsilon_{22}, 2\varepsilon_{12}, \kappa_{11}, \kappa_{22}, 2\kappa_{12}, \gamma_1, \gamma_2]^T,$$
$$\boldsymbol{\sigma}(\mathbf{X}, \mathbf{v}, t) = [n_{11}, n_{22}, n_{12}, m_{11}, m_{22}, m_{12}, q_1, q_2]^T, \tag{1}$$

with $\mathbf{v}$ summarizing the shell's degrees of freedom and $t \in I \subset \mathbb{R}$ denoting the time in a closed interval I. The distinction between covariant and contravariant basis vectors is neglected, since a local Cartesian basis is introduced within the finite element formulation. For further details on the kinematic assumptions, the reader is referred to [14]. Within the scope of this work, all derivations are made with limitations to small strains, but large displacements. Without further specification of the underlying constitutive relation, the material matrix reads

$$\mathbf{D} := \frac{\partial \boldsymbol{\sigma}}{\partial \boldsymbol{\varepsilon}}, \quad \mathbf{D} = \begin{bmatrix} \mathbf{D}_m & \mathbf{D}_{mb} & \mathbf{0} \\ \mathbf{D}_{mb} & \mathbf{D}_b & \mathbf{0} \\ \mathbf{0} & \mathbf{0} & \mathbf{D}_s \end{bmatrix},$$
$$(..)_m = \text{membrane}, (..)_b = \text{bending}, (..)_s = \text{shear}, \tag{2}$$

assuming $\mathbf{D}_{ms}, \mathbf{D}_{bs} = \mathbf{0}$. To derive the macroscopic constitutive relation $\boldsymbol{\sigma}(\boldsymbol{\varepsilon}, t)$, an RVE is introduced at each material point $\mathbf{X}$ in the form of a three-dimensional continuum $\mathcal{B}^m$, see Figure 1. The RVE is described by a orthonormal coordinate system, where x, y refer to the in-plane and z to basis vectors in thickness direction. The domain $\mathcal{B}^m$ is bounded within

$$-\frac{l_x}{2} \leq x \leq \frac{l_x}{2}, \quad -\frac{l_y}{2} \leq y \leq \frac{l_y}{2}, \quad h^- \leq z \leq h^+, \tag{3}$$

where it follows that $h^- = -h/2$ and $h^+ = +h/2$, if the midsurface is chosen as reference surface. Further, we introduce

$$\mathbf{E} = [E_{11}, E_{22}, E_{33}, 2E_{12}, 2E_{13}, 2E_{23}]^T,$$
$$\mathbf{S} = [S_{11}, S_{22}, S_{33}, S_{12}, S_{13}, S_{23}]^T, \tag{4}$$

where $\mathbf{E}$ denotes the Green–Lagrangian continuum strain measure and $\mathbf{S}$ the 2. Piola–Kirchhoff stress in vector notation. Assuming the constitutive relation of different phases of the RVE can inhibit viscoelastic properties, we define

$$\mathbf{S} := \mathbf{S}(\mathbf{E}, \boldsymbol{\xi}, t), \quad \mathbf{C} := \frac{\mathrm{d}\mathbf{S}(\mathbf{E}, \boldsymbol{\xi}, t)}{\mathrm{d}\mathbf{E}} \tag{5}$$

with $\boldsymbol{\xi}$ being an internal variable and $\mathbf{C}$ the material tangent. For the coupling of the two scales, a modified form of the Hill–Mandel condition, demanding equivalence of virtual work on both scales

$$\delta\boldsymbol{\varepsilon}^T \boldsymbol{\sigma} = \frac{1}{A_0} \int_{\mathcal{B}^m} \delta\mathbf{E}^T \mathbf{S} \, \mathrm{d}V \tag{6}$$

is used. The average is hereby taken only over the undeformed reference area of the RVE $A_0 = l_x l_y$. From Equation (6), admissible periodic boundary conditions for the RVE

$$\mathbf{u}^+ = \mathbf{u}^- + \mathbf{A}(\Delta x, \Delta y, z)\, \boldsymbol{\varepsilon},$$
$$\mathbf{A} = \begin{bmatrix} \Delta x & 0 & \frac{1}{2}\Delta y & \Delta x\, z & 0 & \frac{1}{2}\Delta y\, z & 0 & 0 \\ 0 & \Delta y & \frac{1}{2}\Delta x & 0 & \Delta y\, z & \frac{1}{2}\Delta x\, z & 0 & 0 \\ 0 & 0 & 0 & 0 & 0 & 0 & \Delta x & \Delta y \end{bmatrix} \tag{7}$$

can be derived, where $(\cdot)^+$ and $(\cdot)^-$ denote opposing lateral surfaces perpendicular to the xy-plane. Furthermore, additional constraints are used to ensure the independence of the effective shear stiffness from the geometry of the RVE and to prevent rigid body deformations, for details see [15, 16]. To obtain the effective quantities from the RVE, we linearize both sides of Equation (6), yielding

$$\delta\boldsymbol{\varepsilon}^T(\mathbf{D}\Delta\boldsymbol{\varepsilon} + \boldsymbol{\sigma}) = \frac{1}{A_0} \int_{\mathcal{B}^m} \delta\mathbf{E}^T \mathbf{C} \Delta\mathbf{E} + \Delta\delta\mathbf{E}^T \mathbf{S} + \delta\mathbf{E}^T \mathbf{S} \, \mathrm{d}V\,. \tag{8}$$

Inserting a standard displacement ansatz into the right-hand side of Equation (8) for the virtual and linearized quantities results in the discretized global system of equations

$$\delta\boldsymbol{\varepsilon}^T(\mathbf{D}\Delta\boldsymbol{\varepsilon} + \boldsymbol{\sigma}) = \frac{1}{A_0} \begin{bmatrix} \delta\mathbf{V} \\ \delta\boldsymbol{\varepsilon} \end{bmatrix}^T \left\{ \begin{bmatrix} \mathbf{K}_{11} & \mathbf{K}_{12} \\ \mathbf{K}_{21} & \mathbf{K}_{22} \end{bmatrix} \begin{bmatrix} \Delta\mathbf{V} \\ \Delta\boldsymbol{\varepsilon} \end{bmatrix} + \begin{bmatrix} \mathbf{F}_1 \\ \mathbf{F}_2 \end{bmatrix} \right\}, \tag{9}$$

where $\mathbf{V}, \mathbf{K}$, and $\mathbf{F}$ refer to the assembled global vectors of generalized displacements, the stiffness matrix, and the residual vector, respectively. $\Delta\mathbf{V}$ can be eliminated from the system of equations via static condensation, and by comparison of coefficients the effective shell stress resultants and the homogenized material matrix can be extracted

$$\boldsymbol{\sigma} = \frac{1}{A_0}(\mathbf{F}_2 - \mathbf{K}_{21}\mathbf{K}_{11}^{-1}\mathbf{F}_1), \quad \mathbf{D} = \frac{1}{A_0}(\mathbf{K}_{22} - \mathbf{K}_{21}\mathbf{K}_{11}^{-1}\mathbf{K}_{12})\,. \tag{10}$$

The described framework for calculating the desired quantities for a given prescribed strain vector at time t will be used in the following section to generate a database for training an artificial neural network as a parameterizable effective viscoelastic material model.

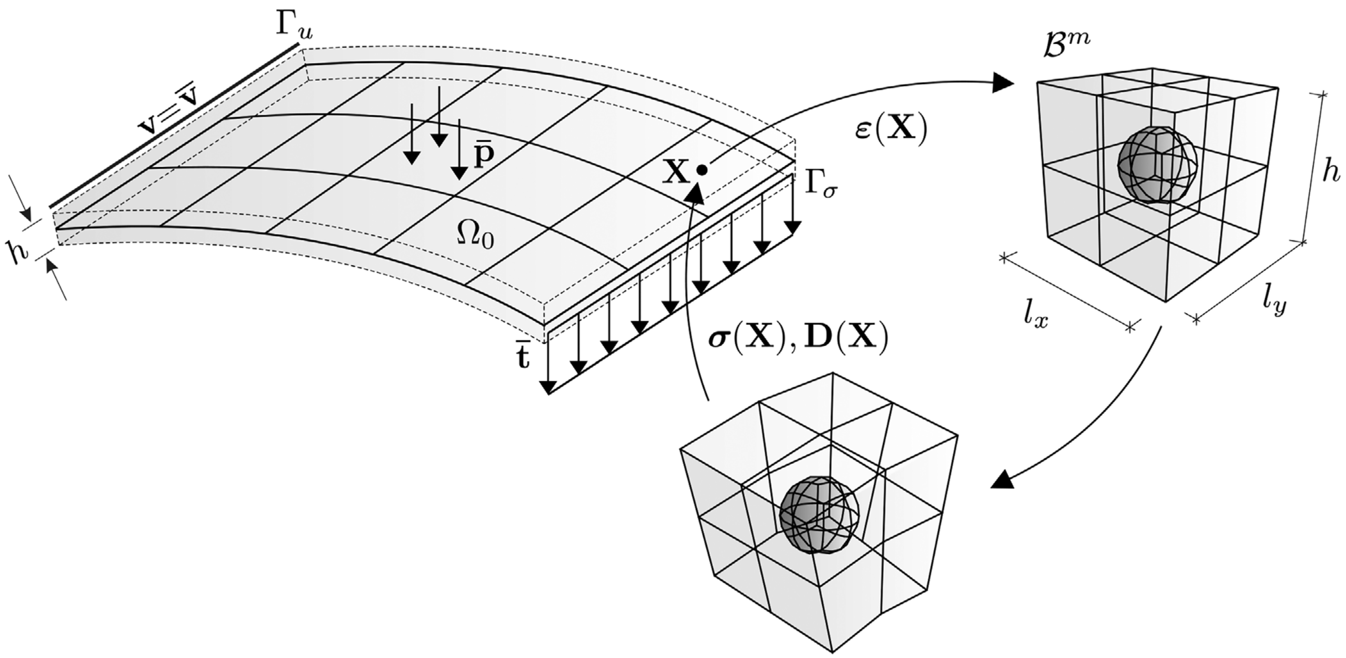


FIGURE 1 | Schematic representation of the homogenization procedure for shell representative volume elements (RVEs).

3 | Parameterizable Viscoelastic ANN Material Model

The following section is devoted to deriving an ANN material model for replacing the time consuming microscopic finite element calculations for the effective quantities. The authors have previously shown in [17], that ANNs are generally capable of approximating linear viscoelastic material behavior in the high-dimensional stress–strain space of shell structures. Here, we further enhance the potential of the model for parameterizable RVEs, including non-strain related input variables, for example, volume fractions of inclusions.

3.1 | Feedforward Neural Network Definitions and Mapping

We define a feedforward ANN as an unambiguous mapping

$$f(\mathbf{x};\mathbf{w}) : \mathbf{x} \mapsto \mathbf{z} \tag{11}$$

with the input vector $\mathbf{x}$, the output vector $\mathbf{z}$, and the free parameters $\mathbf{w}$. For parameterizable viscoelastic material behavior, we define suitable input and output vectors

$$\mathbf{x} := \begin{bmatrix} \Delta\boldsymbol{\varepsilon}_{n+1} \\ \boldsymbol{\varepsilon}_n \\ \boldsymbol{\sigma}_n \\ \Delta t_{n+1} \\ \mathbf{p} \end{bmatrix}, \quad \mathbf{z} := \boldsymbol{\sigma} \in \mathbb{R}^8 \tag{12}$$

with the subscripts $n+1$ referring to the current and n to the last time step, respectively. $\boldsymbol{\sigma}_n$ can be interpreted as an internal variable accounting for the loading history and Δt_{n+1} is the current time increment. The vector $\mathbf{p} \in \mathbb{R}^p$ contains the non-strain related parameters, leading to $\mathbf{x} \in \mathbb{R}^{25+p}$. Consequently, the ANN material matrix follows

$$\mathbf{D}^{ANN} = \frac{\partial \boldsymbol{\sigma}_{n+1}}{\partial \boldsymbol{\varepsilon}_{n+1}} = \frac{\partial \boldsymbol{\sigma}_{n+1}}{\partial \Delta\boldsymbol{\varepsilon}_{n+1}} \underbrace{\frac{\partial \Delta\boldsymbol{\varepsilon}_{n+1}}{\partial \boldsymbol{\varepsilon}_{n+1}}}_{\mathbf{1}} = \frac{\partial \boldsymbol{\sigma}_{n+1}}{\partial \Delta\boldsymbol{\varepsilon}_{n+1}}. \tag{13}$$

Since $\mathbf{D}^{ANN}$ depends on the current strain increment, the viscoelastic ANN material model is by construction non-linear, in contrary to linear viscoelastic models, where the tangent only depends on the current time increment. We further note that the ANN architecture used is not thermodynamically consistent. An approach, where this physical requirement is enforced in a weak sense can be found in [9] or by construction in [12]. As error function, an enhanced version of the classic mean squared error is used

$$E(\mathbf{x}_k, \mathbf{x}_k^C; \mathbf{w}) = \frac{1}{2P} \sum_{k=1}^{P} \sum_{j=1}^{8} (z_j(\mathbf{x}_k;\mathbf{w}) - t_j(\mathbf{x}_k))^2 + \frac{\lambda}{2P^C} \sum_{k=1}^{P^C} \sum_{j=1}^{8} \sum_{i=1}^{8} \left(\frac{\partial z_j(\mathbf{x}_k^C;\mathbf{w})}{\partial x_{\Delta\varepsilon_i}} - D_{ji}(\mathbf{x}_k^C) \right)^2. \tag{14}$$

The first term is formulated with respect to the primary outputs of the neural network, where t_j is the reference stress solution for a sample k and P is the number of samples in a given data set. The second term is enforced by the penalty parameter λ, where we demand the derivative of the ANN to be equal to the material matrix obtained from the homogenization scheme on a given set of constraint samples $\mathbf{x}_k^C \in P^C$, also known as Sobolev training [18]. The search for a suitable set of weights, which minimizes $E(\mathbf{x}_k, \mathbf{x}_k^C; \mathbf{w})$ is done with the Levenberg–Marquardt method [19]. All details on the implementation can be found in [17].

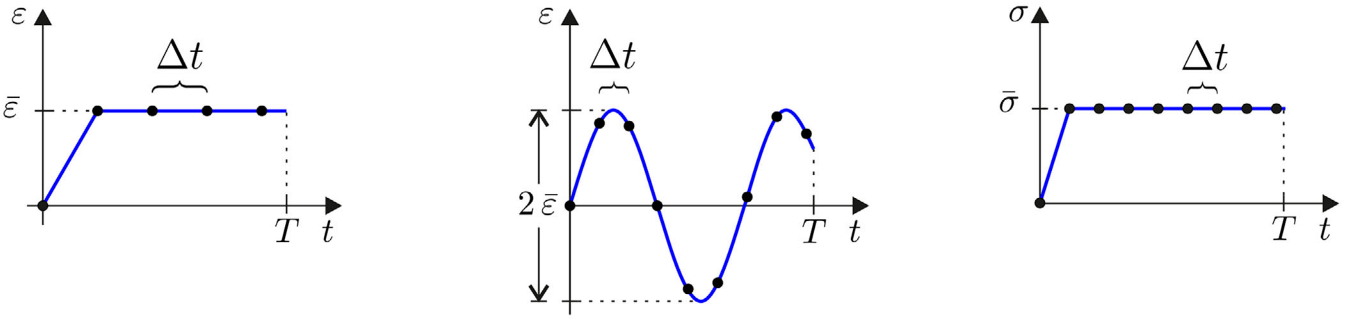


FIGURE 2 | 1D representation of the strain–time and stress–time paths used for the synthetic material tests on the representative volume element (RVE).

3.2 | Generation of the Synthetic Material Test Database

Having defined the ANN architecture, the next step is generating a suitable database. In analogy to experimental testing procedures for viscoelastic materials, the RVE will be subjected to relaxation tests, cyclic loading and creep tests, see Figure 2. Although the former two are done strain controlled with prescribed strain vectors $\bar{\boldsymbol{\varepsilon}}$, the latter is done stress controlled with the prescribed stress vector $\bar{\boldsymbol{\sigma}}$. Additionally, the tests are done in a pseudo-uniaxial sense, further replicating experimental setups, where only the stress in the prescribed test direction is unequal to zero. This, in combination with the constraint with respect to the material matrix, leads to good generalization capabilities for multiaxial and unknown loading paths, for example, unloading, see [17]. Each path $\mathcal{P}_\alpha$ therefore contains the following information:

$$\mathcal{P}_\alpha = \{\dots, (\boldsymbol{\varepsilon}_n, \boldsymbol{\sigma}_n, t_n, \mathbf{p}, \mathbf{D}), (\boldsymbol{\varepsilon}_{n+1}, \boldsymbol{\sigma}_{n+1}, t_{n+1}, \mathbf{p}, \mathbf{D}), \dots\}_\alpha, \quad \alpha = 1, \dots, N_{paths}, \tag{15}$$

where N_{paths} refers to the maximum number of time paths generated. $\mathbf{p}$ is constant for a given path α as well as the material matrix $\mathbf{D}$, since the time increment for a path is constant as well. N_{paths} is usually a multiple of 24, resulting from the number of strain components times the three different synthetic material tests. We limit our material models to the small strain assumption $||\mathbf{E}|| < 0.03$ from which one can derive the limits for the curvatures

$$E_{ij} = \varepsilon_{ij} + z_{max}\,\kappa_{ij}, \quad i, j = 1, 2 \tag{16}$$

considering the relation between the shell and continuum strains. Within a preprocessing, the limits for the shell stress resultants $\boldsymbol{\sigma}$ are determined, such that the resulting maximum creep deformation does not exceed the small strain assumption. Creep and relaxation tests on the RVE are always done until equilibrium is reached, and cyclic loading is at maximum for two periods to avoid unnecessary sample repetition. From the time T it takes to reach equilibrium, adequate limits for the chosen time step size Δt are chosen, which seem suitable for engineering application. Currently, a minimum of 20 and a maximum of 80 samples are taken to discretize a time path to determine adequate boundaries for Δt. After sampling, each time path is transformed to the incremental form given in Equation (12).

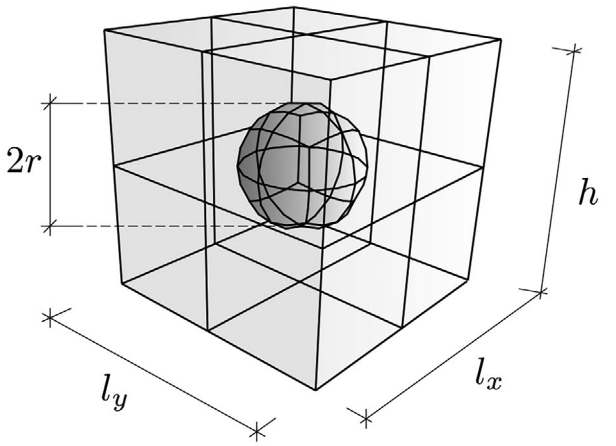
3.3 | Investigations at Material Point Level

In the following section, the potential of the ANN for approximating effective viscoelastic behavior will be investigated. Therefore, an RVE consisting of a viscoelastic matrix and an elastic spherical inclusion is introduced; see Figure 3. The RVE is discretized with 104 27-node solid elements and periodic boundary conditions, where five additional Lagrange parameters are set. Although the length, width, and height of the RVE are constant, the radius of the elastic inclusion is defined in an interval $r \in [7.5\,cm, 12.5\,cm]$ and will be considered as an additional input parameter for the ANN. For the matrix, a generalized Maxwell model with two Maxwell elements of the form

$$E(t) = E_\infty + \sum_{i=1}^{2} E_i e^{-t/\tau_i}, \quad E_0 = E_\infty + \sum_{i=1}^{2} E_i, \quad \mu_i = \frac{E_i}{E_0} \tag{17}$$

is used to describe the viscoelastic behavior for deviatoric parts of the strain measure. For details on the numerical implementation, see, for example, [20]. The material parameters for both the inclusion and the matrix can be found in Table 1. They are chosen in such a way that the matrix could be seen as a very simple model for the elastic creep behavior of concrete with a soft embedded inclusion, which could, for example, be lost formwork. The size of the radius does not affect the time scale of interest, but only the initial and equilibrium response of the material. For the training of the ANN, ten uniaxial creep, cyclic, and relaxation tests are performed on the RVE, resulting in 240 paths and 9199 data samples. All tests are done over a duration of $1 \cdot 10^6$ s ($\approx$ 11.5 days) with $\Delta t \in [12\,500\,\mathrm{s}, 50\,000\,\mathrm{s}]$ resulting in a maximum of 80 and a minimum of 20 samples per path. From Equation (16), with $z_{max} = \pm h/2 = \pm 15$ cm, one can derive the limits for the sampling space under the assumption that $\varepsilon_{ij} = 0$ for uniaxial testing, which here results in $\kappa_{ij} \in [-2 \cdot 10^{-3}, 2 \cdot 10^{-3}]$. It has shown that reducing the curvatures by an order of magnitude 10 improves the ANN training, as it avoids unnecessary large bending moments compared to the remaining shell stress resultants. With this consideration, for the creep tests, the limits of the prescribed shell stress results are

$$|n_{ii}| \leq 1000\,\frac{kN}{cm}, \quad |n_{ij}| \leq 400\,\frac{kN}{cm}, \quad |\kappa_{ii}| \leq 600\,\frac{kNcm}{cm}, \quad |\kappa_{ij}| \leq 250\,\frac{kNcm}{cm}, \quad |q_i| \leq 300\,\frac{kN}{cm}, \quad i \neq j, \tag{18}$$



l_x	$30\ cm$
l_y	$30\ cm$
h	$30\ cm$
r	$[7.5,12.5] cm$

FIGURE 3 | Representative volume element (RVE) consisting of a matrix and spherical inclusion with variable radius.

TABLE 1 | Material parameters for the elastic inclusion and viscoelastic matrix of the reinforced plate.

	Initial modulus E_0	Poisson ratio ν_0	Maxwell term i	Shear parameter μ_i	Relaxation time τ_i
	(kN/cm^2)	$(-)$		$(-)$	(s)
Inclusion	500	0.20	—	—	—
Matrix	3000	0.20	1	0.25	10 000
			2	0.25	100 000

TABLE 2 | Overview of generated test paths on the representative volume element (RVE); for the sake of readability units are neglected.

No.	Test type	Test method	Prescribed strain	Maximum strain values	Δt	T	r
1	Tension	Biaxial	$\varepsilon_{11}, \varepsilon_{22}$	0.02, 0.02	17 500	250 000	7.5
2	Tension	Biaxial	$\varepsilon_{11}, \varepsilon_{22}$	0.02, 0.02	47 500	2 000 000	7.5
3	Tension	Biaxial	κ_{11}, κ_{22}	$1 \cdot 10^{-4}, 1 \cdot 10^{-4}$	17 500	250 000	10
4	Tension	Biaxial	κ_{11}, κ_{22}	$1 \cdot 10^{-4}, 1 \cdot 10^{-4}$	47 500	2 000 000	10
5	Tension	Biaxial	γ_1, γ_2	0.01,0.01	17 500	250 000	12.5
6	Tension	Biaxial	γ_1, γ_2	0.01,0.01	47 500	2 000 000	12.5

to not exceed the small strain assumption for the resulting creep strains at equilibrium. In addition to the data samples, 1000 random constraint samples are chosen where the material matrix constraint is enforced. The neural network consists of two hidden layers with 80 neurons per layer. This architecture was determined in a preprocessing stage based on preliminary studies and has shown to yield accurate and stable results across the considered cases. As an activation function, a hyperbolic tangent is used and the ANN training was terminated after $\approx$ 100 epochs, as convergence in the error measure was reached with the Levenberg–Marquardt method. To evaluate the generalization capabilities of the neural network, six biaxial tension tests are performed on the RVE. These tests are done at two different strain rates for the two prescribed strain components and include both loading and unloading phases. Variations in the inclusion geometry are introduced by considering three different radii. Additionally, tests are carried out with either prescribed membrane strains, curvatures, or shear strains. Lastly, smaller time steps will be used for the tests under higher strain rates and larger time steps for the tests under vanishing strain rates. The data for the test paths can be found in Table 2.

The results of the tests with the ANN material model are depicted in Figure 4 together with the reference solution. The rate dependency is clearly visible in the strain–stress hysteresis, exposing dissipation effects under higher strain rates (left column) and convergence toward perfectly elastic material response for vanishing strain rates (right column). Overall, the ANN provides satisfying results for all tests, considering that the loading and unloading paths and the combination of strain values are not included in the training data and furthermore time steps and values for the radii of the inclusion at the sampling space limits are used. The slightly greater deviation from the reference solution for vanishing strain rates is due to the under representation of the purely elastic training samples. The creep and relaxation tests are cut off at the equilibrium response of the material, which is the point where the material is purely elastic. More accurate results can be achieved by extending the duration of the tests or,

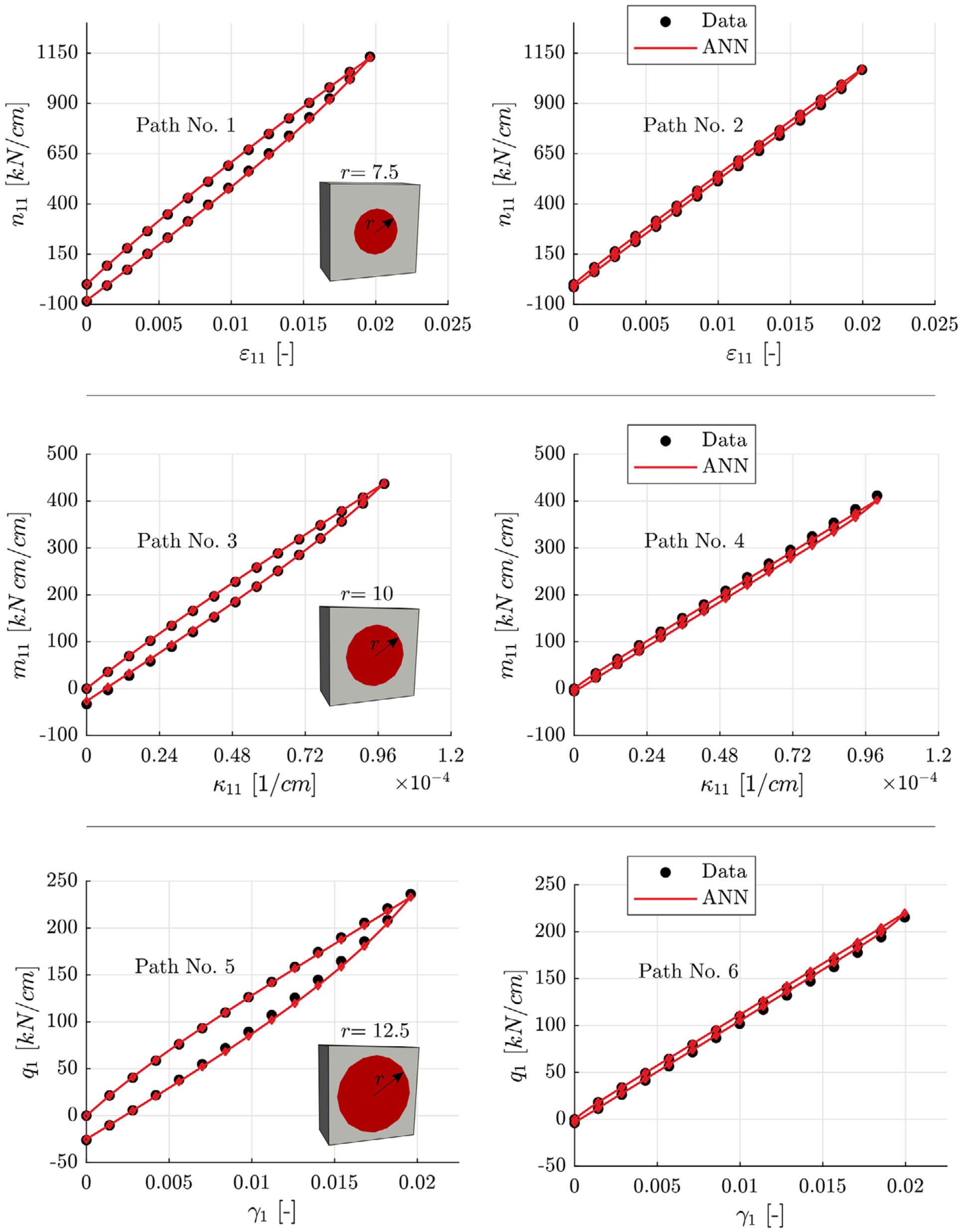


FIGURE 4 | Results of the artificial neural network (ANN) for the biaxial tension tests at different strain rates and varying inclusion radius.

of course, simply including purely elastic samples directly into the training data, for example, by conducting tension/compression tests under vanishing strain rates. This is no problem in general, it simply increases the number of samples. Overall, the ANN is well suited for approximating effective viscoelastic behavior under additional consideration of varying geometric parameters of the underlying RVE.

4 | Structural Simulation Using the ANN Material Model

In this section, the ANN will be used as effective parameterizable viscoelastic material model for structural finite element simulations with shell elements. The results will be compared to full scale 3D solutions as well as results from an FE^2 (shell + RVE) computation as described in Section 2. For the system, see Figure 5, one fourth of a quadratic plate will be modeled with appropriate boundary conditions on the symmetry axes. Simply supported boundary conditions ($w = 0$) are set at the outer boundaries, where in the 3D case only the midside nodes at $z = 0$ are fixed. For the shell model [16], either an FE^2 solution with the RVE from Section 3.3 is used at each of the 2×2 integration points per finite element, or the ANN material model. The 3D solution consists of 15×15 spherical inclusions embedded in the matrix, so essentially 225 RVEs, which are discretized with 27-node solid elements, as described in Section 3.3. The shell on the macroscale is discretized with 10×10 elements. For both solutions, convergence studies have been done in advance. The system is subjected to a surface load—at $z = +h/2$ for the 3D model—which is applied in the first time step and held constant in order to investigate the creep deformation under varying radius of the inclusion. Although the FE^2 and 3D models require new RVEs or, respectively, a new full-scale mesh to be modeled, the

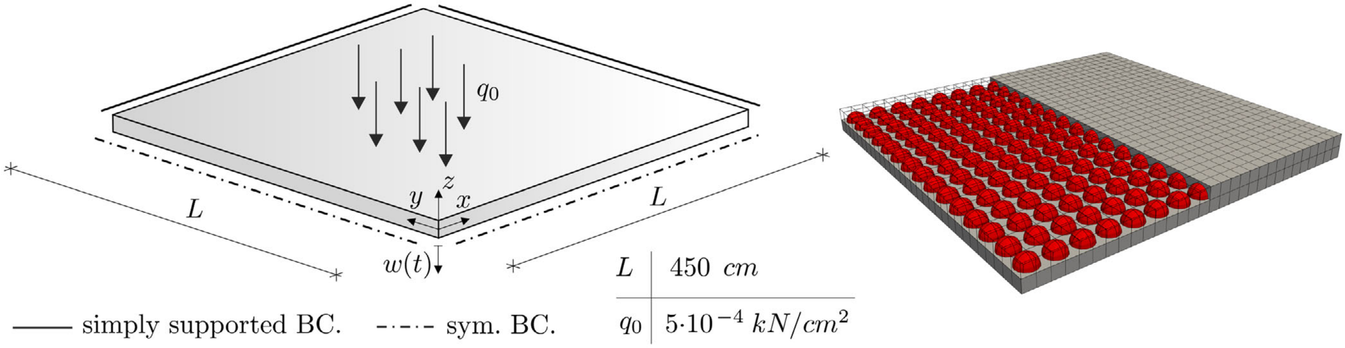


FIGURE 5 | Concrete plate with spherical inclusions as lost formwork. System and boundary conditions (left) and full 3D model (right).

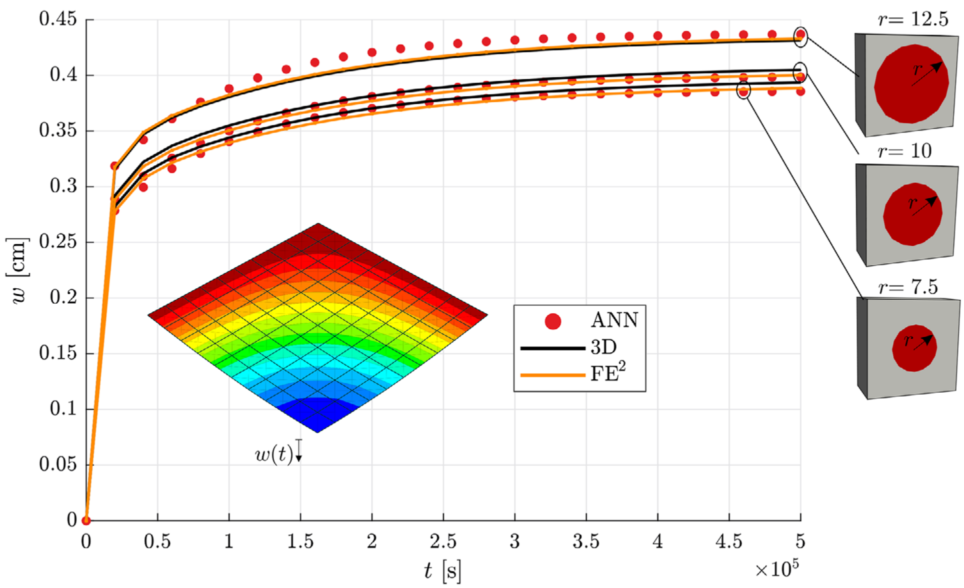


FIGURE 6 | Creep deformation of the three different models under varying inclusion radius.

shell + ANN material model only needs to be evaluated for the desired radius. The quantity of interest is the deformation $w(t)$ at the corner node ($x = y = z = 0$). The load is be applied over $5 \cdot 10^5\, s$ ($\approx$ 6 days) and calculations are done with $\Delta t = 20\,000\, s$ ($\approx$ 5.5 h).

The results for the corner displacement for all three models and radii $r = 7.5, 10, 12.5$ cm are shown in Figure 6. It can be seen that similar displacements are obtained with the three models. As expected, with increasing radius, the creep deformation of the plate increases non-linearly, since the inclusion is softer than the surrounding matrix. This leads to relative differences in the deformation between the minimum and maximum radius of about 10%–15%, depending on the underlying model. The ANN shows some deviation in the displacements for the first time steps and slightly overestimates the creep behavior for $r =$ 12.5 cm before reaching the elastic equilibrium, but provides very satisfying results overall. When it comes to computation times and considering the ANN as reference, the relative difference between the models is factor 10 for the 3D and 35 for the FE^2 solution. It should be considered though, that the ANN is a nonlinear model and therefore makes about five Newton iterations necessary to reach equilibrium, while the other two models are linear. Furthermore, with increased complexity of the microstructure, the computational advance of the ANN approach increases significantly, which has been shown in [17] for fiber reinforced structures. Lastly, the computational cost can furthermore be reduced for solutions computed with the shell model when using a formulation with a one-point quadrature and stabilization as proposed in [21].

5 | Conclusion

In the present work, it has been shown that feedforward neural networks are capable of approximating effective viscoelastic material behavior under additional consideration of varying geometric parameters of the underlying RVE. The ANN was trained with only uniaxial creep and relaxation, as well as cyclic loading data. In combination with Sobolev training, this provides a feasible strategy for the training process and allows for a simple generation of a surrogate model. Tests at the material point level have shown good generalization performance of the model, even for arbitrary unknown loading paths. The ANN is capable of distinguishing between elastic and viscoelastic behavior for different strain rates and provides good results even close to the sampling space limits. Lastly, one significantly reduces the computational time for simulations on the structural level in comparison to 3D or FE^2 solutions, which can be highly advantageous in the design process of materials with complex underlying microstructures.

Acknowledgments

Open access funding enabled and organized by Projekt DEAL.

References

1. J. Ghaboussi, X. Wu, and G. Kaklauskas, "Neural Network Material Modelling," *Statyba* 5, no. 4 (1999): 250–257.

2. M. Lefik and B. Schrefler, "Artificial Neural Network as an Incremental Non-Linear Constitutive Model for a Finite Element Code," *Computer Methods in Applied Mechanics and Engineering* 192, no. 28–30 (2003): 3265–3283.

3. Y. M. A. Hashash, S. Jung, and J. Ghaboussi, "Numerical Implementation of a Neural Network Based Material Model in Finite Element Analysis," *International Journal for Numerical Methods in Engineering* 59, no. 7 (2004): 989–1005.

4. S. Jung and J. Ghaboussi, "Neural Network Constitutive Model for Rate-Dependent Materials," *Computers & Structures* 84, no. 15–16 (2006): 955–963.

5. N. N. Vlassis and W. C. Sun, "Sobolev Training of Thermodynamic-Informed Neural Networks for Interpretable Elasto-Plasticity Models with Level Set Hardening," *Computer Methods in Applied Mechanics and Engineering* 377 (2021): 113695.

6. J. F. Unger and C. Könke, "Neural Networks as Material Models Within a Multiscale Approach," *Computers & Structures* 87, no. 19–20 (2009): 1177–1186.

7. R. Hambli, "Application of Neural Networks and Finite Element Computation for Multiscale Simulation of Bone Remodeling," *Journal of Biomechanical Engineering* 132, no. 11 (2010): 114502.

8. P. Weber, J. Geiger, and W. Wagner, "Constrained Neural Network Training and Its Application to Hyperelastic Material Modeling," *Computational Mechanics* 68, no. 5 (2021): 1179–1204.

9. P. Weber, W. Wagner, and S. Freitag, "Physically Enhanced Training for Modeling Rate-Independent Plasticity with Feedforward Neural Networks," *Computational Mechanics* 72, no. 4 (2023): 827–857.

10. D. K. Klein, M. Fernández, R. J. Martin, P. Neff, and O. Weeger, "Polyconvex Anisotropic Hyperelasticity with Neural Networks," *Journal of the Mechanics and Physics of Solids* 159 (2022): 104703.

11. K. A. Kalina, L. Linden, J. Brummund, and M. Kästner, "FE^{ANN}: An Efficient Data-Driven Multiscale Approach Based on Physics-Constrained Neural Networks and Automated Data Mining," *Computational Mechanics* 71, no. 5 (2023): 827–851.

12. M. Rosenkranz, K. A. Kalina, J. Brummund, W. C. Sun, and M. Kästner, "Viscoelasticty with Physics-Augmented Neural Networks: Model Formulation and Training Methods Without Prescribed Internal Variables," *Computational Mechanics* 74 (2024): 1279–1301.

13. J. Dornheim, L. Morand, H. J. Nallani, and D. Helm, "Neural networks for constitutive modeling: from universal function approximators to advanced models and the integration of physics," *Archives of Computational Methods in Engineering* 31 (2024): 1097–1127.

14. W. Wagner and F. Gruttmann, "A Robust Non-Linear Mixed Hybrid Quadrilateral Shell Element based on the Hu-Washizu Functional," *International Journal for Numerical Methods in Engineering* 64, no. 5 (2005): 635–666.

15. S. Klarmann, F. Gruttmann, and S. Klinkel, "Homogenization Assumptions for Coupled Multiscale Analysis of Structural Elements: Beam Kinematics," *Computational Mechanics* 65, no. 3 (2019): 635–661.

16. F. Gruttmann and W. Wagner, "A FE2 Shell Model with Periodic Boundary Conditions for Thin and Thick Shells," *International Journal for Numerical Methods in Engineering* 125, no. 11 (2024): e7433.

17. J. Geiger, W. Wagner, and S. Freitag, "Multiscale Modeling of Viscoelastic Shell Structures with Artificial Neural Networks," *Computational Mechanics* 76 (2025): 541–565.

18. W. M. Czarnecki, S. Osindero, M. Jaderberg, G. Swirszcz, and R. Pascanu, "Sobolev Training for Neural Networks," *Advances in Neural Information Processing Systems* 30 (2017): 4278–4287.

19. M. Hagan and M. Menhaj, "Training Feedforward Networks with the Marquardt Algorithm," *IEEE Transactions on Neural Networks and Learning Systems* 5, no. 6 (1994): 989–993.

20. J. C. Simo and T. J. R. Hughes, "Computational Inelasticity," in *Interdisciplinary Applied Mathematics* (Springer 1998).

21. F. Gruttmann and W. Wagner, "A Nonlinear 4-Node Shell Element with One Point Quadrature and Stabilization Based on a Hu–Washizu Variational Formulation," *Computational Mechanics* 76 (2025): 613–623.