

On the Integration of Multiple Load Protocols for Parameter Identification and Calibration of Superelastic Shape Memory Alloy Models Using Artificial Neural Networks

Niklas Lenzen, Florian Kolisch, Sven Klinkel

Abstract. In this study, we present an Artificial Neural Network (ANN) architecture that efficiently identifies all model parameters, while considering multiple load protocols simultaneously. Our approach also addresses the ill-posed nature of the parameter identification problem by implementing a dual-network architecture for training, improving the accuracy of parameter identification. The first part of the dual-network maps stress responses of multiple load protocols to parameters. The second part is a surrogate of the constitutive model and reconstructs the stress responses using the calculated parameters from the first part. To generate training data, we use the constitutive model to simulate stress responses for multiple load protocols with parameters sampled by Latin hypercube sampling. Once trained, the proposed ANN can predict suitable parameter sets for arbitrary compositions and elements of superelastic shape memory alloys (SMA), provided that the parameters are within the sampled space. We validate the proposed method by comparing the numerical results with cyclic stress–strain test data. In conclusion, our method offers an efficient and user-friendly tool for robust parameter identification of superelastic SMA models. It easily integrates multiple load protocols and eliminates the need for expert knowledge and additional training.

Key words: Parameter Identification, Constitutive Modeling, Shape Memory Alloys, Artificial Neural Networks, Inverse Modeling

1 INTRODUCTION

Due to their superior hysteretic energy dissipation and deformation capabilities, superelastic shape memory alloy (SMA) elements are increasingly being incorporated into vibration control devices for civil engineering structures. For the design of SMA-based control devices, accurate constitutive models are required. These models should particularly be able to precisely replicate strain rate-dependent responses such as the uniaxial model proposed by Auricchio et al. [1].

Niklas Lenzen & Florian Kolisch & Sven Klinkel
Chair of Structural Analysis and Dynamics, Department of Civil Engineering, RWTH Aachen University, Mies-van-der-Rohe-Str. 1, 52074 Aachen, Germany, e-mail: lenzen@lbb.rwth-aachen.de

Besides mathematical modeling of SMA, data-driven response modeling techniques, such as *Artificial Neural Networks* (ANNs), have been developed as these can directly learn patterns from data. In their pioneering work, Ozbulut et al. developed an Adaptive Neuro-Fuzzy Inference System to map strain and strain rate inputs to stress responses of SMAs [2]. More recently, Weber et al. successfully modeled the responses of SMA fibers using a *Feedforward Neural Network* (FNN) [3]. To enhance the generalizability, the FNN was trained with a physical constraint. Besides, more advanced architectures, such as *Deep Operator Networks* [4], can be applied for response modeling of SMAs [5]. However, ANN-based response modeling typically requires large amounts of data, which is often not available in desired quality. Thus, in case of low data availability, mathematical modeling approaches are advantageous due to their thermodynamic-consistent framework.

A critical step towards accurate response modeling using mathematical models is the identification of model parameters. Physical parameters, such as the Young's modulus, are typically determined directly through physical experiments, cf. [6]. However, constitutive models also rely on parameters critical to internal calculations, whose relationship with observed responses is often intricate and challenging to interpret. While established calibration methods for SMA models, such as by Meraghni et al. [7], have demonstrated strong capabilities, the evolution of ANNs has brought techniques that offer greater efficiency, especially when modeling highly non-linear relations. Despite their strengths, conventional methods often demand extensive expert knowledge of both the constitutive models and the optimization algorithms, along with access to specialized experimental equipment. This becomes particularly challenging when modeling material responses under dynamic loading conditions, where the large number of model parameters often leads to an ill-posed identification problem.

To efficiently solve these issues, ANNs offer great potential. For SMA models, Helm proposed an FNN-based approach to identify material parameters from stress inputs [8]. As this approach only considered quasi-static loads, several advances were made to apply dynamic loads [9, 10, 11]. However, these methods are specifically designed to predict parameter sets for a specific SMA composition.

In this study, we present an ANN-based method that predicts parameter sets for various SMA compositions without any adjustments of the ANN's trainable parameters. The proposed method enables the ANN to learn the mapping between stress time histories and parameter sets across a broad parameter space, ensuring adaptability and generalizability. Once trained, it provides predictions in near real-time and eliminates the need of expert knowledge for the calibration of SMA models. To ensure robust response modeling across different load scenarios, we integrate multiple load protocols as inputs to the ANN. Moreover, to overcome the ill-posed nature of the identification process, we implement a *dual-network* architecture for the training process, as introduced in [12, 13]. This architecture is inspired by *Autoencoders* [14].

The structure of the paper is as follows. Section 2 introduces the constitutive model used in this work and outlines our ANN-based approach for parameter identification. The results are presented in Section 3, where we examine the training accuracy and demonstrate the application of our method to two distinct SMA compositions. Section 4 summarizes our key findings.

2 METHODOLOGY

2.1 Constitutive model

To demonstrate the performance of the ANN-based parameter identification method, we use the uniaxial constitutive model by Auricchio et al. [1]. The model predicts the stress σ , material temperature T , and martensite volume fraction ξ as functions of the applied strain ε and strain rate $\dot{\varepsilon}$. In this work, the analysis is restricted to wires undergoing uniaxial tensile loading. Accordingly, the free energy is formulated as

$$\psi = [(u_A - T\eta_A) - \xi(\Delta u - T\Delta\eta)] + C \left[(T - T_0) - T \ln \frac{T}{T_0} \right] + \frac{1}{2} E (\varepsilon - \varepsilon_l \xi)^2, \quad (1)$$

where u_A is the internal energy of the austenite phase; η_A is the initial entropy of the austenite phase; Δu is the internal energy difference between the austenite and martensite phases; $\Delta\eta$ is the entropy difference between the austenite and martensite phases; C is the heat capacity; T_0 is the ambient/reference temperature; E is the Young's modulus; and ε_l is the maximum inelastic strain at $\xi = 100\%$.

The stress is derived from the free energy as

$$\sigma = \frac{\partial \psi}{\partial \varepsilon} = E (\varepsilon - \varepsilon_l \xi), \quad (2)$$

where the Young's modulus is defined according to Auricchio and Sacco [15] as a function of the martensite volume fraction to account for a transition between the two phases. It is expressed as

$$E = \frac{E_A E_M}{E_M + \xi(E_A - E_M)}, \quad (3)$$

where E_A and E_M are constants indicating the respective elastic moduli of the austenite and martensite phases. The heat equation is derived from the *first law of thermodynamics* as

$$\begin{aligned} C\dot{T} &= -\frac{\partial \psi}{\partial \xi} \dot{\xi} + T \frac{\partial^2 \psi}{\partial T \partial \varepsilon} \dot{\varepsilon} + T \frac{\partial^2 \psi}{\partial T \partial \xi} \dot{\xi} - \gamma(T - T_0) \\ &= \sigma \varepsilon_l \dot{\xi} + \Delta u \dot{\xi} - \gamma(T - T_0), \end{aligned} \quad (4)$$

where γ is the heat convection coefficient. To model the evolution of the martensite fraction ξ , the following first-order differential equations are employed:

$$\text{A} \rightarrow \text{M}: \quad \dot{\xi} = \beta^{AM} (1 - \xi) \frac{\dot{F}}{(F - R_f^{AM})^2} \mathcal{H}^{AM}, \quad (5)$$

$$\text{M} \rightarrow \text{A}: \quad \dot{\xi} = \beta^{MA} \xi \frac{\dot{F}}{(F - R_f^{MA})^2} \mathcal{H}^{MA}, \quad (6)$$

where F is the driving force; $R_f^{AM/MA}$ are the thermo-coupled limit stress levels; $\beta^{AM/MA}$ are the speed parameters controlling the rate effects on the phase transformation; and $\mathcal{H}^{AM/MA}$ are defined as the activation factors. These are expressed as

$$\mathcal{H}^{AM} = \begin{cases} 1 & \text{when } \dot{F} > 0 \text{ and } R_s^{AM} < F < R_f^{AM}, \\ 0 & \text{otherwise,} \end{cases} \quad (7)$$

$$\mathcal{H}^{MA} = \begin{cases} 1 & \text{when } \dot{F} < 0 \text{ and } R_f^{MA} < F < R_s^{MA}, \\ 0 & \text{otherwise,} \end{cases} \quad (8)$$

with

$$F = \sigma - T \frac{\Delta\eta}{\varepsilon_l}, \quad (9)$$

$$R_{s/f}^{AM/MA} = \sigma_{s/f}^{AM/MA} - T_0 \frac{\Delta\eta}{\varepsilon_l}, \quad (10)$$

where $\sigma_{s/f}^{AM/MA}$ are the critical stress levels at which phase transformations start/finish.

Accurate identification of these parameters is essential to reliably capture the dynamic response of the material. The large number of model parameters and their direct influence on each other highlights the ill-posed nature of the identification problem. In total, 13 parameters have to be determined.

2.2 ANN-based parameter identification

The constitutive model is represented as $\mathcal{M}((\varepsilon, \dot{\varepsilon}), \mathbf{p})$, where ε and $\dot{\varepsilon}$ denote the respective time histories of strain and strain rate. The parameter vector $\mathbf{p}$ contains all material model parameters. Our ANN-based method is designed to integrate multiple stress responses from cyclic tensile stress–strain experiments in the identification process. In this study, to effectively integrate multiple load protocols, we choose *Convolutional Neural Networks* (CNNs). Alternatively, FNNs would also be applicable. The CNN-based method is illustrated in Figure 1.

In a first step, training data is generated using the constitutive model. Therefore, I combinations of parameters $\mathbf{p}_i, \forall i \in \{1 : I\}$ are sampled within a predefined parameter space by the *Latin hypercube sampling* (LHS) [16]. The LHS is selected due to its effectiveness in achieving a balanced exploration of the parameter space, ensuring that samples are both representative and uniformly distributed across defined intervals. To cover all characteristics of the desired SMA specimen, such as strain rate effects and thermomechanical coupling, L load protocols, represented by strain time histories $\varepsilon_j, \forall j \in \{1 : L\}$, are considered. The load protocols should be chosen, such that all desired material phenomena are covered. Thus, we recommend to consider different strain rate loadings and strain amplitudes for the identification process. For training data generation, each parameter sample $\mathbf{p}_i$ is then applied to the constitutive model to generate stress responses $\boldsymbol{\sigma}$ for each of the L load protocols. As a result, the training data set $\mathcal{D} = \{(\mathbf{p}_i, [\boldsymbol{\sigma}_1 \boldsymbol{\sigma}_2 \dots \boldsymbol{\sigma}_L]_i), \forall i \in \{1 : I\}\}$ is obtained. To improve training efficiency, the

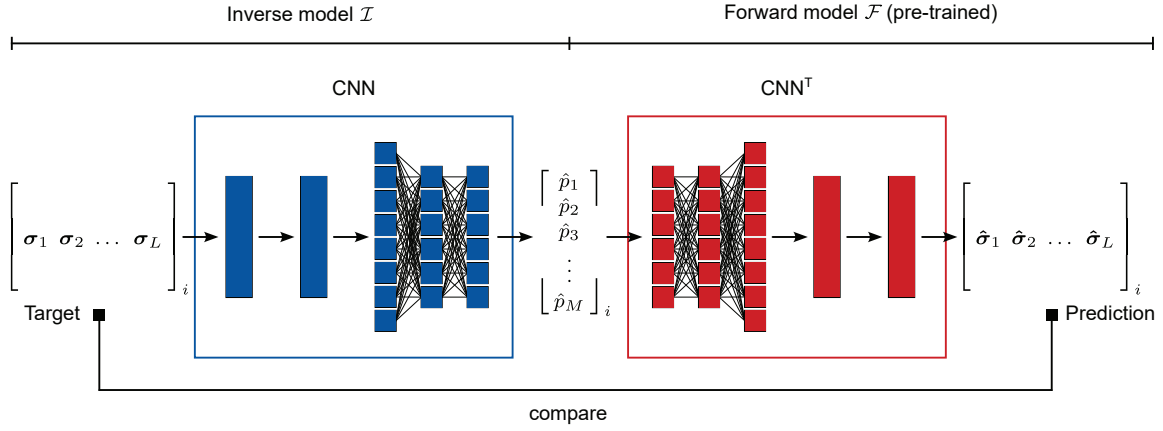


Figure 1: Training phase: The inverse model $\mathcal{I}$ is trained to map stress responses σ from L load protocols to parameters $\hat{\mathbf{p}}$. A pretrained forward model $\mathcal{F}$ is attached to reconstruct the stress responses σ from the parameter prediction $\hat{\mathbf{p}}$ of the inverse model $\mathcal{I}$. The predicted stress responses $\hat{\sigma}$ are compared with the target stress responses σ .

data set is normalized using *min-max* normalization. For greater stability against measurement error, *Gaussian noise* is added to the stress responses, with the mean and variance chosen based on the magnitude of the noise in the experimental test data.

Given the inherently ill-posed nature of the inverse problem, directly training an CNN to infer parameters from stress responses often leads to suboptimal results. This is primarily because the CNN may inadvertently penalize valid parameter sets. To mitigate this issue, a dual-network architecture is applied, cf. Figure 1. The training process is structured into two distinct phases. In the initial phase, a forward model $\mathcal{F}$, governed by a set of trainable parameters ω , is optimized to predict stress responses from given parameters, i.e., $\mathcal{F}_\omega : \mathbf{p}_i \rightarrow [\sigma_1 \ \sigma_2 \ \dots \ \sigma_L]_i$. The forward model is trained by minimizing a loss function defined as the mean squared error (MSE) between predicted $\mathcal{F}_\omega [\mathbf{p}_i]$ and target stress responses $[\sigma_1 \ \sigma_2 \ \dots \ \sigma_L]_i$ with respect to the CNN-parameters ω , i.e.,

$$\mathcal{F}_\omega \leftarrow \min_{\omega} \frac{1}{I} \sum_{i=1}^I \|\mathcal{F}_\omega [\mathbf{p}_i] - [\sigma_1 \ \sigma_2 \ \dots \ \sigma_L]_i\|^2. \quad (11)$$

Following the training of the forward model, an inverse model $\mathcal{I}$, parameterized by weights and biases θ , is introduced. This model aims to infer the underlying parameters from observed stress responses, i.e., $\mathcal{I}_\theta : [\sigma_1 \ \sigma_2 \ \dots \ \sigma_L]_i \rightarrow \mathbf{p}_i$. During the training of the inverse model, the previously trained forward model is employed to reconstruct stress responses using the parameter estimates produced by the inverse network, cf. [12, 13]. Within this architecture, the forward model effectively serves as a differentiable surrogate for the constitutive model. Its integration into the backpropagation process allows for efficient gradient computation via automatic differentiation. As a result, the inverse model is optimized by minimizing the MSE between the reconstructed $\mathcal{F}_\omega [\mathcal{I}_\theta [[\sigma_1 \ \sigma_2 \ \dots \ \sigma_L]_i]]$ and target stress responses $[\sigma_1 \ \sigma_2 \ \dots \ \sigma_L]_i$ with respect to the CNN-parameters θ , i.e.,

$$\mathcal{I}_\theta \leftarrow \min_{\theta} \frac{1}{I} \sum_{i=1}^I \|\mathcal{F}_\omega [\mathcal{I}_\theta [[\sigma_1 \sigma_2 \dots \sigma_L]_i]] - [\sigma_1 \sigma_2 \dots \sigma_L]_i\|^2. \quad (12)$$

During the training of the inverse model, the CNN-parameters ω of the forward model are held fixed. This ensures that only discrepancies between the reconstructed and target stress responses contribute to the loss, thereby guiding the inverse model's learning process. The dual-network architecture simultaneously reduces dependence on expert knowledge, both in terms of constitutive model formulation and the selection of parameter spaces. By focusing solely on reproducing the target stress responses, the training scheme guarantees that the resulting model effectively captures the material behavior, even if the inferred parameters lack direct physical interpretation. Pretraining of the forward model is essential to enable the inverse model to reliably predict the parameters. Without this step, the inverse model lacks a well-defined mapping between the parameters and the corresponding stress responses, resulting in arbitrary or unstable outputs.

Several additional components are integral to the training process. The hidden layers utilize the *ReLU* activation function in conjunction with the *He initialization* strategy to promote efficient gradient flow and weight scaling [17]. For the inverse model, a sigmoid activation function is applied in the output layer to constrain outputs within the $[0, 1]$ interval, thereby ensuring compatibility with the predefined parameter space. This layer is initialized using the *Glorot initializer* to maintain balanced activation and gradient flow during training [18].

3 RESULTS

3.1 Training

First, we generate training data by using the constitutive model. Therefore, we sample $I = 100 \cdot 10^3$ parameter combinations using the LHS. Depending on the predefined parameter space, the amount of training samples for successful training can vary. Table 1 reports the chosen upper and lower boundaries of the parameter space. To cover a variety of SMA compositions, we choose a broad parameter space. In total, 13 parameters are considered within the identification process. The ambient temperature used in the constitutive model is set to $T_0 = 296.2$ K. To generate stress responses, we use $L = 9$ load protocols, which are reported in Table 2. Each load protocol consists of $J = 1000$ time instants, which are sampled for a time step size $\Delta t = 0.001$ s. For training data generation, we feed the load protocols and the parameter samples through the constitutive model. As a result, we obtain data pairs consisting of parameter samples $\mathbf{p}$ and stress responses σ , such that the training data set $\mathcal{D} = \{(\mathbf{p}_i, [\sigma_1 \sigma_2 \dots \sigma_9]_i), \forall i \in \{1 : 100 \cdot 10^3\}\}$. Moreover, we use additional $\tilde{I} = 20 \cdot 10^3$ data pairs for validation. Before training, the data set is normalized using *min-max normalization*. Furthermore, to enhance the robustness of the CNN prediction, we add *Gaussian noise* with a zero mean and a variance of $1 \cdot 10^{-3}$. This is particularly important as we will be using noisy experimental data for testing.

Table 1: Sample space of the searched parameters. The lower and upper boundaries of the sample space are denoted by min and max.

Parameter		min	max	Unit
Young's moduli	E_A	25,000	40,000	MPa
	E_M	10,000	20,000	MPa
Max. inelastic strain at $\xi = 1$	ε_l	2.0	6.0	%
Critical stress levels	σ_s^{AM}	70	300	MPa
	σ_f^{AM}	350	650	MPa
	σ_s^{MA}	200	350	MPa
	σ_f^{MA}	70	200	MPa
Heat capacity	C	0.1	10.0	MPa K ⁻¹
Heat convection coefficient	γ	0	10	-
Internal energy difference	Δu	0	50	MPa
Entropy difference	$\Delta \eta$	0.0	1.0	MPa K ⁻¹
Speed parameters	β^{AM}	1	100	-
	β^{MA}	1	100	-

Table 2: Load protocols used for training. Harmonic signals are employed: $\varepsilon = \varepsilon_0 \sin(\Omega t)$, where ε_0 is the strain amplitude and Ω is the excitation frequency. The strain rate amplitude is computed by $\dot{\varepsilon}_0 = \varepsilon_0 \Omega$.

	LP1	LP2	LP3	LP4	LP5	LP6	LP7	LP8	LP9
ε_0 [%]	2	2	2	4	4	4	6	6	6
$\dot{\varepsilon}_0$ [%/s]	6.28	12.57	25.13	12.57	25.13	50.27	18.85	37.70	75.40

The hyperparameters of the dual-network architecture have been optimized by testing various configurations. Table 3 lists all network hyperparameters. To stabilize training and enhance the generalizability, we use *batch normalization* [19]. We trained the forward and inverse models for 2000 epochs with a batch size of $n_b = 1000$. To efficiently optimize the trainable parameters, the *Adam optimizer* [20] is employed with a learning rate of $\eta = 1 \cdot 10^{-3}$.

To determine the accuracy, we calculate the *coefficient of determination* R^2 , where we compare the target data of the validation data set with the CNN predictions. Both the forward and inverse models show great accuracy with $R_{\mathcal{F}}^2 = 99.90\%$ and $R_{\mathcal{I}}^2 = 99.93\%$.

3.2 Experimental validation

To validate the proposed method, we identify constitutive model parameters for two Nitinol wire compositions:

- Sample 1: Ni-55.8%-Ti-43.55% manufactured by SAES Getters S.p.A.

Table 3: Hyperparameters of the inverse $\mathcal{I}$ and forward $\mathcal{F}$ models.

	No. filters	Kernel size	Stride	Padding	No. neurons	Activation
$\mathcal{I}$	2×16	(9,16)	(2,1)	<i>same</i>	$4 \times 50, 1 \times 13$	<i>ReLU</i>
$\mathcal{F}$	$1 \times 16, 1 \times 1$	(9,16)	(2,1)	<i>same</i>	$1 \times 13, 4 \times 50$	<i>ReLU</i>

Table 4: Parameter sets $\hat{\mathbf{p}}$ identified from measured cyclic stress responses of the SMA wire samples using the proposed method.

Parameter		Sample 1	Sample 2	Unit
Young’s moduli	E_A	24,089	29,980	MPa
	E_M	10,922	12,564	MPa
Max. inelastic strain at $\xi = 1$	ε_l	3.1	2.7	%
Critical stress levels	σ_s^{AM}	203	189	MPa
	σ_f^{AM}	352	370	MPa
	σ_s^{MA}	307	321	MPa
	σ_f^{MA}	171	84	MPa
	C	6.0	3.3	MPa K ⁻¹
Heat capacity	C	6.0	3.3	MPa K ⁻¹
Heat convection coefficient	γ	0.12	5.1	-
Internal energy difference	Δu	49.20	48.25	MPa
Entropy difference	$\Delta \eta$	0.29	0.19	MPa K ⁻¹
Speed parameters	β^{AM}	9.9	6.2	-
	β^{MA}	76.1	70.4	-

- Sample 2: Ni-55.9%-Ti-43.95% manufactured by Baoji Hanz Metal Material Co.

For test data generation, we conducted cyclic tensile stress–strain experiments on wire specimens of length $l = 150$ mm and diameter $d = 0.2$ mm at ambient temperature $T_0 = 296.2$ K, which is consistent with the conditions assumed for training data generation. Moreover, to induce phase transformation at lower strain amplitudes, a pre-stress of $\sigma_0 = 134.9$ MPa was applied. More details on the test setup can be found in [11]. For accurate parameter identification, we measure the experimental stress responses for the same load protocols as used for training, cf. Table 2.

To identify the parameters, the experimental test data is fed into the trained inverse model $\mathcal{I}$ for both alloy compositions separately. The identified parameter sets are listed in Table 4.

Furthermore, we compare the experimental stress–strain data with the estimations calculated by the material model using the identified parameter sets from Table 4. The results for both alloy compositions are plotted in Figures 2 and 3. The calculated responses accurately reconstruct the experimental data, demonstrating the capabilities of our method. Moreover, to evaluate the accuracy of the estimations, we compute the coefficient of determination R^2 . The accuracy for

Sample 1 is greater than 88.56%, while Sample 2 reaches an accuracy of greater than 94.80%. It is important to note that the accuracy of our method is limited by the precision of the material model. This can be observed, for example, in the residual deformations that are not considered by the material model.

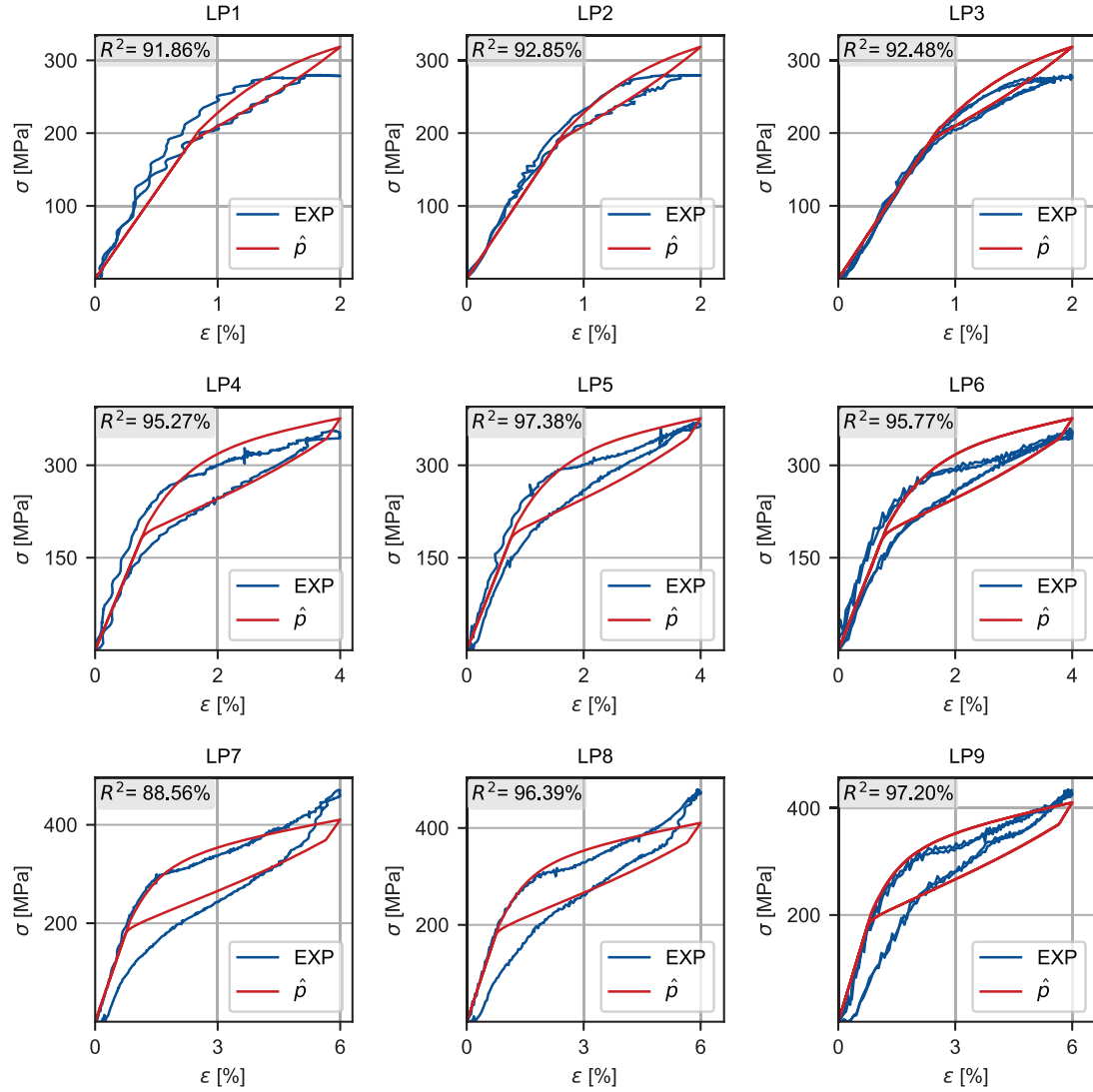


Figure 2: Sample 1: Comparison of the experimentally determined stress–strain responses (EXP) with predictions computed by the constitutive model using parameter set $\hat{p}$, cf. Table 4.

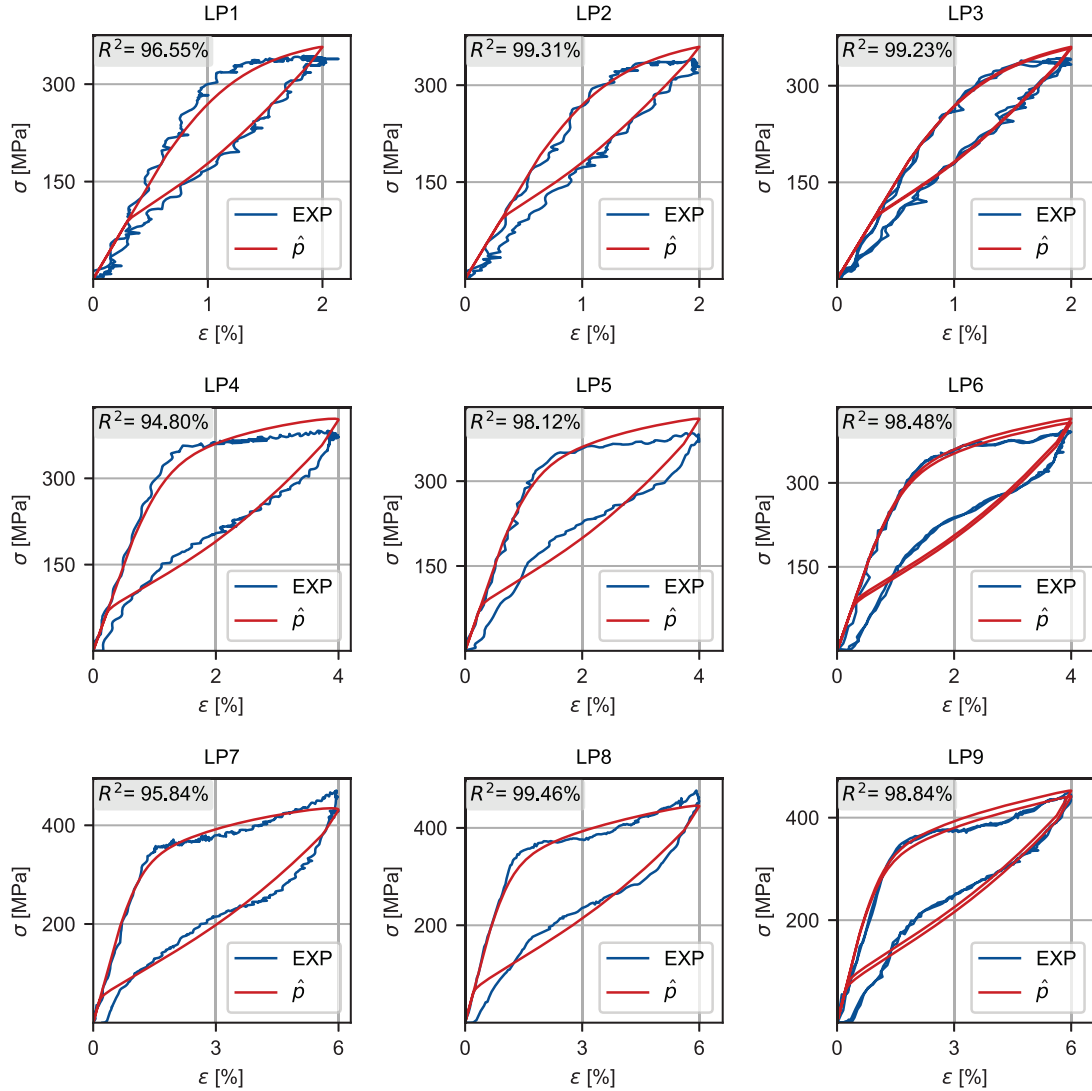


Figure 3: Sample 2: Comparison of the experimentally determined stress–strain responses (EXP) with predictions computed by the constitutive model using parameter set $\hat{p}$, cf. Table 4.

4 CONCLUSIONS

This paper presented an ANN-based parameter identification method for uniaxial superelastic SMA models subjected to multiple dynamic load protocols. An efficient integration of multiple load protocols has been developed for the parameter identification of constitutive SMA models, ensuring accurate and robust response modeling across a variety of load scenarios. In this study, we used CNNs to efficiently integrate multiple load protocols. To address the ill-posed

nature of the inverse problem, a dual-network architecture was implemented, significantly enhancing the reliability of parameter identification. The CNN prediction accuracy demonstrates the effectiveness of the dual-network architecture, reaching values above $R^2 = 99.90\%$. Moreover, the trained CNN enables near real-time prediction of suitable parameter sets for different SMA compositions, provided the desired parameters lie within the sampled space. In this study, we considered two different SMA compositions for parameter identification. The results underline the high precision of our method, achieving prediction accuracy above $R^2 = 88.56\%$ and $R^2 = 94.80\%$, respectively.

Looking ahead, as the presented method is limited by the accuracy of the constitutive model, further research into more precise material models is required. Consequently, to improve the prediction accuracy of the presented method, additional load protocols should be included in the ANN-based parameter identification process. These additional load protocols should cover all effects present in the material.

REFERENCES

- [1] F. Auricchio, D. Fugazza, and R. Desroches, “Rate-dependent thermo-mechanical modelling of superelastic shape-memory alloys for seismic applications,” *Journal of Intelligent Material Systems and Structures*, vol. 19, no. 1, pp. 47–61, Feb. 2007.
- [2] O. E. Ozbulut, P. N. Roschke, P. Y. Lin, and C. H. Loh, “GA-based optimum design of a shape memory alloy device for seismic response mitigation,” *Smart Materials and Structures*, vol. 19, no. 6, p. 065004, Apr. 2010.
- [3] P. Weber, W. Wagner, and S. Freitag, “ANN material modeling for SMA fibers enhanced with a physical constraint and its application to FE computations,” *Computational Mechanics*, Jan. 2025.
- [4] L. Lu, P. Jin, G. Pang, Z. Zhang, and G. E. Karniadakis, “Learning nonlinear operators via deeponet based on the universal approximation theorem of operators,” *Nature Machine Intelligence*, vol. 3, no. 3, pp. 218–229, Mar. 2021.
- [5] N. Lenzen and O. Altay, “Physics-informed deep operator network for predicting martensite evolution in superelastic shape memory alloys through cyclic tensile tests,” *Smart Materials and Structures*, vol. 33, no. 6, p. 065039, May 2024.
- [6] F. Auricchio, A. Coda, A. Reali, and M. Urbano, “SMA numerical modeling versus experimental results: Parameter identification and model prediction capabilities,” *Journal of Materials Engineering and Performance*, vol. 18, no. 5–6, pp. 649–654, Aug. 2009.
- [7] F. Meraghni, Y. Chemisky, B. Piotrowski, R. Echchorfi, N. Bourgeois, and E. Patoor, “Parameter identification of a thermodynamic model for superelastic shape memory alloys using analytical calculation of the sensitivity matrix,” *European Journal of Mechanics - A/Solids*, vol. 45, pp. 226–237, May 2014.

- [8] D. Helm, “Pseudoelasticity: experimental observations, thermomechanical modeling, and identification of the material parameters,” in *Smart Structures and Materials 2004: Active Materials: Behavior and Mechanics*, D. C. Lagoudas, Ed. SPIE, Jul. 2004.
- [9] N. Lenzen and O. Altay, “Machine learning enhanced dynamic response modelling of superelastic shape memory alloy wires,” *Materials*, vol. 15, no. 1, p. 304, Jan. 2022.
- [10] N. Lenzen, S. Klinkel, and O. Altay, “Parameter identification and modeling of superelastic shape memory alloy wires subjected to dynamic loads,” in *10th ECCOMAS Thematic Conference on Smart Structures and Materials*, ser. SMART 2023. Dept. of Mechanical Engineering & Aeronautics University of Patras, 2023, pp. 1768–1777.
- [11] N. Lenzen and O. Altay, “Feedforward neural network-assisted parameter identification and tuning for uniaxial superelastic shape memory alloy models under dynamic loads,” *Journal of Vibration and Control*, Jun. 2024.
- [12] Y. Long, J. Ren, Y. Li, and H. Chen, “Inverse design of photonic topological state via machine learning,” *Applied Physics Letters*, vol. 114, no. 18, May 2019.
- [13] S. Kumar, S. Tan, L. Zheng, and D. M. Kochmann, “Inverse-designed spinodoid metamaterials,” *npj Computational Materials*, vol. 6, no. 1, Jun. 2020.
- [14] G. E. Hinton and R. R. Salakhutdinov, “Reducing the dimensionality of data with neural networks,” *Science*, vol. 313, no. 5786, pp. 504–507, Jul. 2006.
- [15] F. Auricchio and E. Sacco, “A one-dimensional model for superelastic shape-memory alloys with different elastic properties between austenite and martensite,” *International Journal of Non-Linear Mechanics*, vol. 32, no. 6, pp. 1101–1114, Nov. 1997.
- [16] M. D. McKay, R. J. Beckman, and W. J. Conover, “A comparison of three methods for selecting values of input variables in the analysis of output from a computer code,” *Technometrics*, vol. 42, no. 1, pp. 55–61, Feb. 2000.
- [17] K. He, X. Zhang, S. Ren, and J. Sun, “Delving deep into rectifiers: Surpassing human-level performance on imagenet classification,” in *Proceedings of the IEEE International Conference on Computer Vision (ICCV)*. IEEE, December 2015.
- [18] X. Glorot and Y. Bengio, “Understanding the difficulty of training deep feedforward neural networks,” in *Proceedings of the thirteenth international conference on artificial intelligence and statistics*. JMLR Workshop and Conference Proceedings, 2010, pp. 249–256.
- [19] S. Ioffe and C. Szegedy, “Batch normalization: Accelerating deep network training by reducing internal covariate shift,” 2015.
- [20] D. P. Kingma and J. Ba, “Adam: A method for stochastic optimization,” 2014.