**Abstract**

Our tool provides an intuitive workflow that imports and processes raw unfiltered shape memory alloy mechanical (tensile/compression), thermal (DSC), or thermomechanical (tensile/compression with environmental chamber) data to produce customizable figures and systematically derived material data. This toolset can extract data from multiple inputs such as tensile test data and external thermocouples and automatically synchronize them onto the same time series. With raw force and displacement data, the SMA REACT can calculate strains and stresses based on various sample geometries. Coupling temperature, stress, and strain data, this tool can apply customizable filters and remove systematic errors within the dataset, periodically prompting the user for filter approval. The program then produces various figures to help visualize the complex shape memory alloy material behavior.

**Introduction**

The fundamental understanding and industrial application of shape memory alloys (SMAs) continues to grow. Add a few sentences here about why SMAs are used (high work energy density). Specifically highlight the actuator properties, because that's what the calibration routine can do right now.

The process for developing and integrating a shape memory alloy engineering system can be divided into six main steps (detailed graphically in figure 1). First, based on system requirements, material requirements (i.e., stiffness, actuation strain, transformation temperatures) are derived. Then, material scientists formulate potential compositions, either based on *a priori* knowledge or existing tools (cite Othmane). With a number of candidate compositions, the material is processed via vacuum induction melting (VIM), vacuum arc reacting (VAR), or another technique; for a thorough description of each processing method, the curious reader is directed to the review papers by these authors. Thermomechanical test specimens are then produced to characterize the material response for the particular loading condition in question (i.e., tension, compression, or torsion). At this stage, the material scientists typically assess the extent at which the shape memory alloy meets the requirements, and then either try another composition or processing technique or provide the processed material to an engineering design team.

****

Figure 1: The typical SMA development process involves many discrete steps. This work provides an easy constitutive model calibration tool, the Rendering of Experimental Analysis and Calibration Tool, to enable SMA component design.

For proper/rigorous system design, the full thermomechanical constitutive response must be captured via a material model. The design team calibrates a constitutive model (examples of which include the Brinson and Lagoudas models) to best fit the thermomechanical characterization data in the operating regime of interest. Finally, which a calibrated constitutive model, the greater engineering system can be designed, taking into account the intricacies of the SMA in question (variable actuation strain with respect to stress, etc.).

Each stage of the development process detailed above requires significant time and effort, but the greater SMA community has developed tools to speed certain development stages. NASA and TAMU (cite Othmane/Karaman) have developed user-friendly tools and methods to understand the composition-processing-property relationships in SMAs, enabling quick discovery of new alloys. ASMADA focuses on identifying heating and cooling cycles of SMAs and extracting ASTM standard SMA material properties. Other SMA-focused tools also exist, such as SMAnalytics, which is a comprehensive analyzer that does.... Further, there is XXX which attempts to identify SMA material properties based on the graphical interpretation of data, not the data itself (check SMST 2022). Add a note about the TAMU UMAT and others here (Cite Hartl, Lai, etc.) also add a note about standards for thermomechanical characterization (and perhaps add it to the figure?).

Recently, significant momentum towards standardizing SMA test methods and exploring the full composition – processing – property space has collected/been undertaken/can be seen. There exist robust tools to graphically interpret the relationship between composition, processing, and material properties, developed and maintained by NASA Glenn. Automated extraction of ASTM standard properties can be performed via an open-source GUI. However, despite many tools for model calibration of standard and even superelastic SMA constitutive models available in open-source and commercial software, no such analog exists for SMA actuation models.

However, a laborious workflow of experimental analysis and constitutive model calibration is still commonly required for rigorous SMA characterization. This is especially difficult for newcomers to the field or those engaged in multi-disciplinary efforts. The various external state variables that govern shape memory material behavior (i.e., temperature, stress, strain) often require conglomeration of multiple instruments to properly record sufficient data and can result in inefficient use of time when synchronizing various datasets from different instruments. Further, when such data has been properly compiled, there are myriad additional methods of arriving at an accurate constitutive model to describe the specific SMA material behavior.

In this work, we detail a streamlined open-source tool to help both material scientists and design engineers analyze their thermomechanical data and calibrate an appropriate SMA constitutive model. This tool is deemed REACT, for the Rendering of Experimental Analysis and Calibration Tool. Our tool provides an intuitive workflow that imports and processes raw unfiltered shape memory alloy mechanical (tensile/compression), thermal (DSC), or thermomechanical (tensile/compression with environmental chamber) data to produce customizable figures and systematically derived material data (depicted schematically in Figure 2). SMA REACT consists of two graphical user interfaces (GUIs), written in python and available on GitHub under the GNU General Public License. These two GUIs accomplish separate essential tasks in the SMA development process: data pre-processing and filtering, and material model calibration.

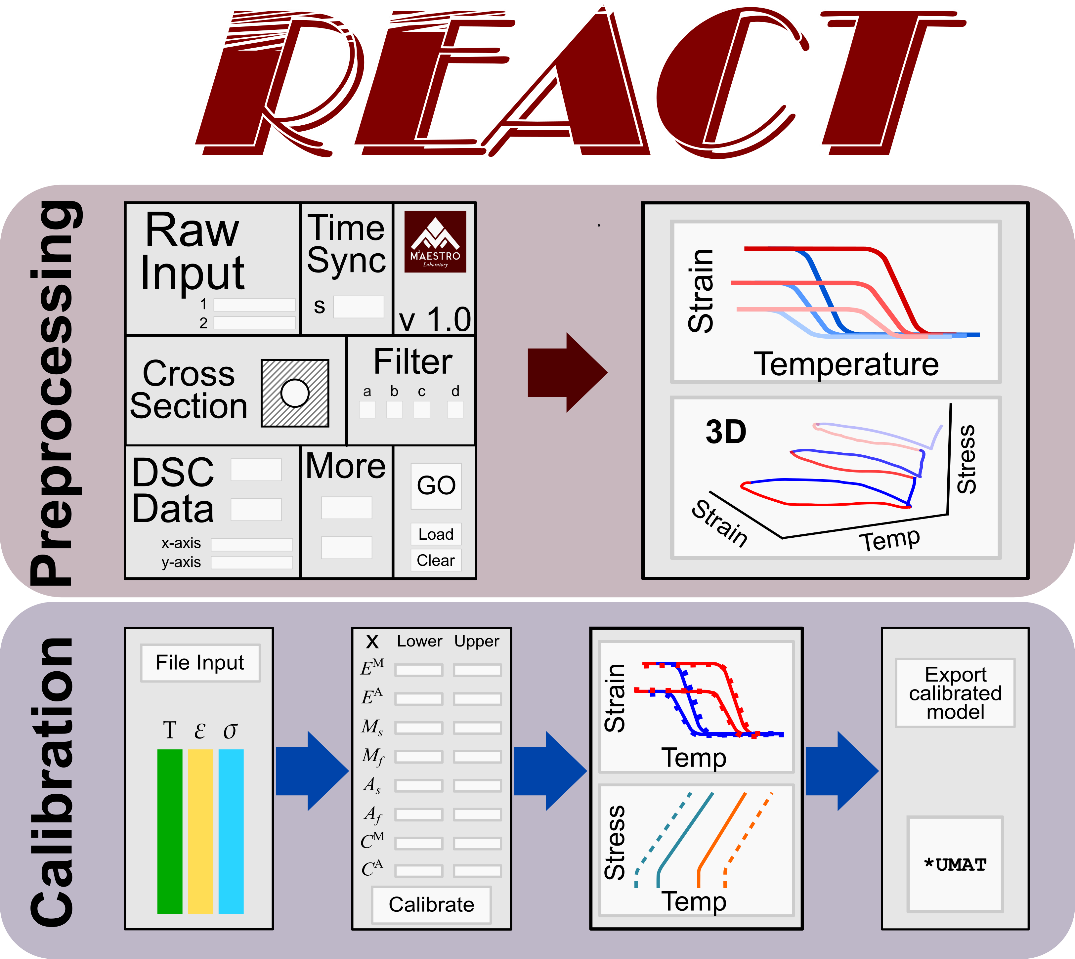


Figure 2: SMA-REACT allows the user to load their own data, specify known model parameters, and find an optimal calibration that best approximates experimental response.

In this work, we describe a new open-source GUI for constitutive model calibration of SMA actuators. We hope to provide a vital link between materials scientists and SMA design engineers via an accessible software, written in python but requiring no programming experience. We focus on the temperature-driven Lagoudas 1-D constitutive model, but the methods and accompanying software described herein can be easily extended to consider other constitutive models, higher dimensional models (e.g., 3D models with anisotropic effects), and different loading modes (e.g., superelasticity).

The pre-processing GUI extracts data from multiple inputs such as tensile test data and external thermocouples and automatically synchronize them onto the same time series. With raw force and displacement data, the SMA REACT can calculate strains and stresses based on various sample geometries. Coupling temperature, stress, and strain data, this tool can apply customizable filters and remove systematic errors within the dataset, periodically prompting the user for filter approval. The program then produces various figures to help visualize the complex shape memory alloy material behavior.

Furthermore, the calibration GUI finds the best fit of constitutive model parameters (martensitic elastic modulus, austenite start temperature, etc.) given filtered and synchronized experimental data. Following the thermodynamically consistent model derived by Lagoudas, et al., the developed calibration routine leverages global optimization strategies to minimize error between model prediction and experimental data. The tool described herein enables the user to customize the optimization routine as well as the model parameters to be optimized (e.g., bounds and free variables). Outputs from the calibration routine include a set of model parameters to be used in future analyses (i.e., material properties for FEA) and a thermodynamically consistent phase diagram based on calibrated model parameters. The current workflow attempts to minimize tribal knowledge contained within the SMA constitutive modeling community by demystifying processes used for calibration. We hope this tool can provide an efficient workflow and salient guidance to others in the shape memory alloy community for years to come.

**Method description**

**Data-preprocessing**

**Model calibration**

SMA Constitutive Model Calibration, or Parameter Identification, describes the process of finding the set of model parameters (Martensite Start Temperature, Maximum transformation strain, etc.) that best fit material experimental data in the mode of operation relevant for the engineering component of interest (e.g., tension, compression, torsion, or a combination thereof). Mathematically, calibration is the process of minimizing error between constitutive model predictions and experimental data subject to physical constraints (conservation laws, known bounds for material properties, etc.) by varying model parameters.

An appropriately calibrated constitutive model is essential for design of complex systems with SMAs. The inherent thermomechanical coupling present in SMAs makes it important to understand how the SMA component will behave when installed in the system and subject to relevant loading conditions. Understanding the strain recovery behavior of SMAs, especially in the presence of minor loops, is crucial to designing the entirety of the engineering component. The inherent complexity of Shape Memory Alloys is an opportunity to design more space- and weight-efficient assemblies, but a challenge to accurately design these systems to perform as intended. For many applications, selecting a particular SMA component based on transformation temperature and maximum transformation strain is insufficient; the transformation temperatures and actuation strain in the *operating stress regime* must be well characterized and predictable. Constitutive model calibration is a vital link for designing and validating SMA performance.

Historically, SMA model calibration has been performed analytically, based on analysts’ best-guesses of appropriate properties, and via numerical optimization. Closed-form analytical results/expressions can be derived for simple models (e.g., cite cite cite) when a deterministic amount of data is available. However, when the operating range of the SMA spans many stress regions and requires many (>3) experimental tests, these analytical methods become overdetermined. Commonly, an appropriate calibrated model is determined based on rules of thumb and guess-and-check methods that rely on design engineer intuition and many manual iterations. More recently, many groups have adopted numerical optimization to find the combination of model parameters that best fit experimental data. These approaches help to speed the process, but exist as purpose-built codes and are themselves steeped in institutional knowledge that is difficult to transfer from analyst to analyst. Most existing calibration methods, while sufficient to produce an adequate fit, have limited applicability outside the authors’ specific application or research group. The rules of thumb and tribal knowledge required to produce a consistent and accurate calibration lead to difficulties reproducing similar calibrations on different material systems, and the required effort for new analysts to learn the ``tricks of the trade'' limits the effectiveness and transferability of these methods.

For this work, due to the inherent interdependence of so many material properties, and assuming that the driving factor for calibration is proper fit of experimental data, we can approach the calibration problem as a numerical optimization problem. For constant force thermal cycle experiments, strain is measured as a function of temperature at certain stress levels. We seek a set of material properties that best matches the experimental data over this set of experiments. This can be represented as a mathematical optimization problem.

Because of the aforementioned material property interdependence, we use a hybrid optimization scheme to best balance global searches with local optima; when multiple experiments are conducted, this optimization problem is overdetermined so there may exist many local optima. Hybrid optimization consists of two main stages: global optimization followed by a local search on the best set of design variables that the global optimization found. The global optimization searches the entire space and hopefully finds the small region where the best solution lies. Then, based on that point, a gradient-based optimization is implemented to find the mathematical optimum point in that smaller subset of the design space. Hybrid optimization has a better chance of finding the true optimum because the preliminary global optimization acts as a “Greedy design of experiments” – it selectively samples points based on knowledge gained in the past.

Our tool leverages the genetic algorithm NSGA-II for the global search and then SLSQP implemented in SciPy for the local search, although the tool is modular and other optimization algorithms can be easily inserted in lieu of the ones discussed herein. The population size for NSGA-II is commonly set to 100 and the genetic algorithm is typically run for at least 10 generations, while SLSQP is set to run for approximately 100 maximum iterations. All optimization parameters are modifiable in the GUI.

**One-dimensional Lagoudas SMA Constitutive Model**

The Lagoudas shape memory alloy constitutive model uses the Gibbs' free energy to derive a thermodynamically consistent relationship between stress and strain. In this work, we will discuss the commonly used temperature- and strain-driven implementation of this model for wider applicability in standard finite element suites. From the perspective of model calibration, seventeen unique but dependent model parameters must be found to best match experimental data. In this section, we will omit a full model derivation, but rather discuss the mathematical foundation of the model and then highlight the seventeen model parameters that need calibrated and their effects on constitutive behavior.

Table 1:The one-dimensional reduction of the Lagoudas SMA constitutive model requires calibration of 17 unique but dependent parameters.

|  |  |  |
| --- | --- | --- |
| **Parameter** | **Mathematical Symbol** | **Units (SI)** |
| **Thermoelastic properties** |  |  |
| Elastic moduli |  | Pa |
| Coefficient of thermal expansion |  | 1/K |
| **Transformation properties** |  |  |
| Transformation temperatures |  | K |
| Stress-influence coefficients |  | Pa/K |
| **Transformation strain properties** |  |  |
| Minimum transformation strain |  | mm/mm |
| Maximum transformation strain |  | mm/mm |
| Critical stress at which transformation strain manifests |  | Pa |
| Transformation strain rise time |  | 1/Pa |
| **Smooth hardening properties** |  |  |
| Smooth hardening coefficients |  | - |

Alternatively, we can discuss the model from the lens of how the calibrated parameters affect constitutive response. In shape memory alloy constitutive modeling, three distinct plots are crucial to understand:

1. The phase diagram (Figure 3a), which describes the surfaces that define forward and reverse transformation in stress-temperature space. Zero-stress transformation temperatures can be found by inspection, and the stress-influence coefficients are defined as the slopes of the forward and reverse transformation surfaces at the *calibration stress*. The calibration stress is *a priori* defined by the designer, and common practice dictates selecting a value close to the design working stress of the material. Get a citation on this. Talk to Hartl?
2. Constant-force thermal cycles (Figure 3b), which describe material behavior in strain-temperature response and inform designers of actuator-like behavior. The elastic moduli, transformation strain properties, and thermal expansion coefficient can be extracted from this data. Additionally, this plot can give intuition about the smooth hardening coefficients. Note that the transformation strain properties are not simply the strain in martensite minus the strain in austenite; this is a measure of total strain, and the transformation strain must be calculated by applying Hooke's law.
3. The current transformation strain as a function of applied stress (Figure BLANK), which is crucial to understand if the given material exhibits sufficient transformation strain at the design stress. We show a special case herein, where and are both nonzero. However, in most cases, one or both of these material properties are zero. The exponential coefficient k denotes the ``rise time'' or how quickly the transformation strain increases from to . Cite some old work that backs this up. Book? My papers? Bertagne? Leal?

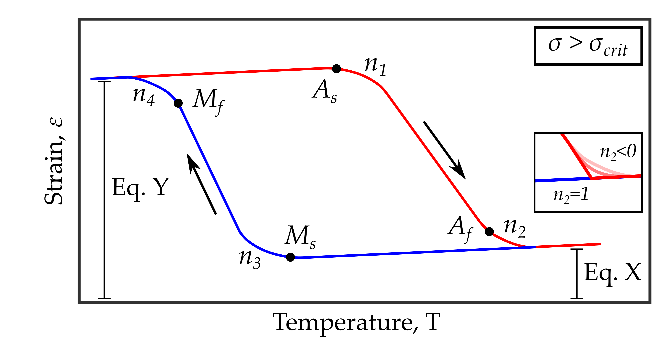
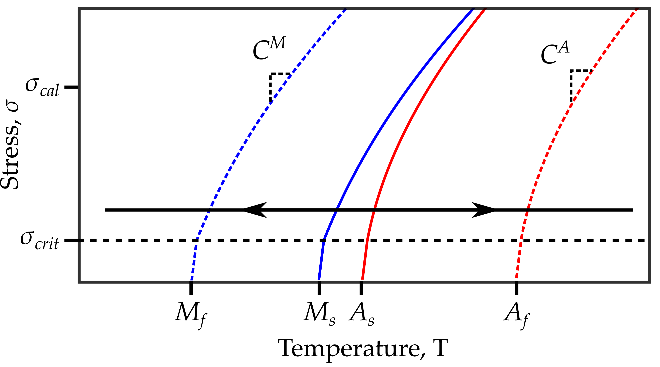


Figure 3: Due to the interdependence of model parameters, numerical optimization is required for a robust fit of experimental data.

As mentioned earlier, the seventeen material properties that define shape memory alloy constitutive response are unique but interdependent. For example, a change in smooth hardening coefficient will result in a change in the corresponding zero-stress transformation temperature. This is because the model definition of the transformation temperature is based on the point at which transformation either initiates or stops and not the tangent to both lines. Add a note about how the definition of these properties is different than how the ASTM standards define them. Herein lies a crucial nuance of calibrating the Lagoudas constitutive model: the model defines the transformation temperatures as the point at which transformation begins (i.e., the state where the transformation criteria are activated), rather than the tangent. In past work, manually updating smooth hardness coefficients, transformation temperatures, and stress-influence temperatures to best fit experimental data has been the most time-intensive part of calibration.

Many other material properties are interrelated as well; changing transformation strain properties will change both the strain-temperature response and the shape of the transformation surfaces. Additionally, while the stress-influence coefficients are single numbers for each phase, they are only one part of the mathematical expression to define the transformation surface in stress-temperature space. For these reasons, to ensure a robust fit of experimental data (which herein we assume consists of many strain-temperature cycles at various stress levels), numerical optimization is needed.

**Calibration via numerical optimization**

The current implementation of the tool, in a GUI-based format, allows the SMA designer to specify both optimization parameters and material property bounds and values. If certain properties are known a priori (e.g., the Young’s moduli from tensile tests), these values can be defined and the optimization will minimize error between model prediction and experiment by varying all other material properties. In this way, our GUI allows the designer greater flexibility than previous methods, but the typical iterative method can still be used. The only difference is the speed in which different combinations of material properties and their respective bounds can be tested.

Each calibration routine can be executed in less than 10 minutes, depending on the size of the optimization, and the results are easily digestible for those who are not innately familiar with the Lagoudas SMA constitutive model. In this way, our tool provides a high-throughput, low-barrier-to-entry calibration method that we hope to increase use of SMAs in practice.

**Post-processing and outputs**

**Implementation example**

**Experimental data**

****

Figure 4: To demonstrate the utility of SMA-REACT, we will calibrate a constitutive model to fit published experimental data [1].

To calibrate an accurate SMA constitutive model to capture actuator behavior, *n* isobaric (constant force thermal cycling) tests, where *n* is preferably greater than 4, are required. Each test requires stress-strain-temperature histories.

We use experimental data of a Ni50.5Ti27.2Hf22.3 alloy from Bigelow et al [1]. This dataset is chosen because NiTiHf is a relevant material system, with many members of the SMA community exploring Hf additions for high-temperature performance (cite a handful of NiTiHf review papers). Additionally, the quality and quantity of data (six different constant force thermal cycling tests were performed), non-zero coefficients of thermal expansion, and nonlinear relationship between applied stress and transformation strain make this data set an ideal candidate to demonstrate the utility of our calibration tool.

**Conventional calibration procedure**

Calibration of the 17 unknown parameters that define the Lagoudas SMA constitutive model can be calibrated without a global optimization strategy by estimating parameter groups (e.g., transformation temperatures, thermoelastic properties, etc.) sequentially. However, due to the nonlinearities present in the model with respect to the current transformation strain (H\_cur) and smooth hardening coefficients (n\_i), optimization, or nonlinear curve fitting, is still required. Herein, we will discuss one method to find these parameters without using a global optimization strategy, as this method can serve as a good baseline with which to compare the optimized calibration.

First, transformation temperatures for each tested stress level can be estimated via the tangent method or similar. If a “zero-stress” isobaric test (i.e., 7 MPa or lower) was performed, the transformation temperatures found for this test can be taken as the zero-stress transformation temperatures etc. Otherwise, each zero-stress transformation temperature can be found via the x-intercept of a linear regression of the transformation temperatures as a function of stress. This estimate is equivalent to a Lagoudas model calibration with smooth hardening parameters set to . The average slope of the start and finish transformation surfaces for martensite and austenite for a specified stress range about the user-determined *calibration stress* can be taken as the stress-influence coefficients ( and ). Note that the stress-influence coefficients should not be derived from the average slope from estimated transformation temperatures at all stress levels; most shape memory alloys exhibit a nonlinear change in transformation temperature with respect to stress (cite Othmane), and the Lagoudas model compensates for this via the transformation surfaces (), where the stress-influence coefficients at the calibration stress are a contributing factor.

With transformation temperatures and stress-influence coefficients estimated, thermoelastic properties and transformation strain properties can be calculated. Austenite elastic modulus can be found by extracting the total strains and a temperature well above at each tested stress level. Then, by designating this temperature such that , Hooke’s law becomes:

Austenite elastic modulus is the best-fit linear coefficient from this equation.

At this point, the analyst has two choices in terms of calculating the rest of the thermoelastic and transformation strain properties. First, the coefficient of thermal expansion can be calculated separately by extracting the total strain at another temperature :

Alternatively, coefficient of thermal expansion can be found concurrently with the other thermoelastic and transformation strain properties via a nonlinear system of equations. At a temperature , Hooke’s Law can be written as:

where

In this equation, there are six unknowns: and . Ideally, to calibrate these six unknowns, one will have performed six or more constant-force thermal cycle tests. However, for shape memory materials that do not exhibit two-way shape memory effect, both and can be set to zero, reducing the number of required tests to four. If the coefficient of thermal expansion was calibrated based on elastic response, the other five parameters are calibrated in the same way as described above.

Both of these approaches to calculate the remaining thermoelastic properties and transformation strain properties may introduce modeling errors. If , the strain due to thermal expansion will be incorrectly predicted across the tested temperature range. However, this is a limitation of the one-dimensional reduction of the Lagoudas constitutive model; assuming thermal expansion is invariant of material phase allows for the use of simpler nonlinear solution methods (i.e., Convex Cutting Plane). Additionally, the nonlinearities associated with the current transformation strain make the transformation strain very sensitive to change in model parameters (in particular the rise time ). This sensitivity requires the analyst to try a range of starting values for before a nonlinear curve fitting routine converges to the global optimum. Typical values for range from and for materials with small and large changes in transformation strain as a function of stress, respectively.

At this point, all material properties are estimated; to fully capture the true strain-temperature response, iterative calibration of each smooth hardening coefficient is necessary until a satisfactory fit is accomplished. However, due to the interdependencies highlighted earlier, each change of smooth hardening coefficient will need to be accompanied by a change in the associated transformation temperature and perhaps stress-influence coefficient (i.e., a change in will need to be accompanied by a change in and . Without global numerical optimization, the best model calibration will be found via manual changes and the analyst’s intuition.

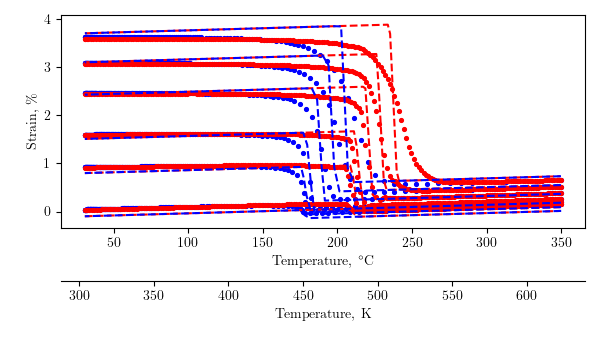


Figure 5: Conventional analytical/numerical calibration techniques produce a passable solution, but rely on user iterations to fine-tune model response.

The resulting conventional calibration, with , is shown in Figure 5. The mean squared error between model and experiment is 3.13% (see equation BLANK), which for many applications could suffice. Transformation temperatures and transformation strain properties are well-captured, as the model bisects the experimental curves as a function of stress (i.e., the martensite elastic response is underpredicted at low levels of applied stress but over approximated at high levels of applied stress). However, there are a number of areas that could be improved. First, the coefficient of thermal expansion is clearly too high (see the elastic response in Martensite). Additionally, this calibration routine still requires a nonlinear curve fitting procedure to find the transformation strain properties, as well as requiring more intuition about the relationship between model parameters and constitutive response.

**Calibration via numerical optimization**

The conventional calibration shown above is used as a starting point for a global numerical calibration using SMA-REACT. Because the current optimization strategy includes a preliminary genetic alogirhtm followed by a gradient-based algorithm, the previously found model parameters were used to determine bounds for each model parameter (i.e., the conventional calibration estimated austenitic elastic modulus as BLANK, so the bounds were set to BLANK and BLANK). This allows the optimizer to start in the neighborhood of feasible solutions, but gives it freedom to explore for a better performing result.

Then, based on the values to which the optimization converged, the parameters that converged to the bounds were further inspected, bounds were widened, further improving the calibration accuracy. This process of inspecting the converged results and comparing to the optimization bounds was repeated three times until each parameter converged to a value well within the set bounds. This indicates that a local optima is found, and with a large enough initial population in the genetic algorithm, we are confident that this is near the globally optimal calibration for this model formulation.

Table 2: The SMA-REACT tool allows further refinement of the calibrated solution.

|  |  |  |
| --- | --- | --- |
| **Calibration Number** | **Mean squared error** | **Notes** |
| 1 | 3.13% | Analytical calibration, . |
| 2 | 2.09% | Numerical calibration with bounds around analytical values |
| 3 | 1.57% | Widened bounds on and . |
| 4 | 1.46% | Widened bounds on and . Fixed and . |
| 5 | 1.43% | Widened bounds on . Fixed everything but transformation temperatures and . |

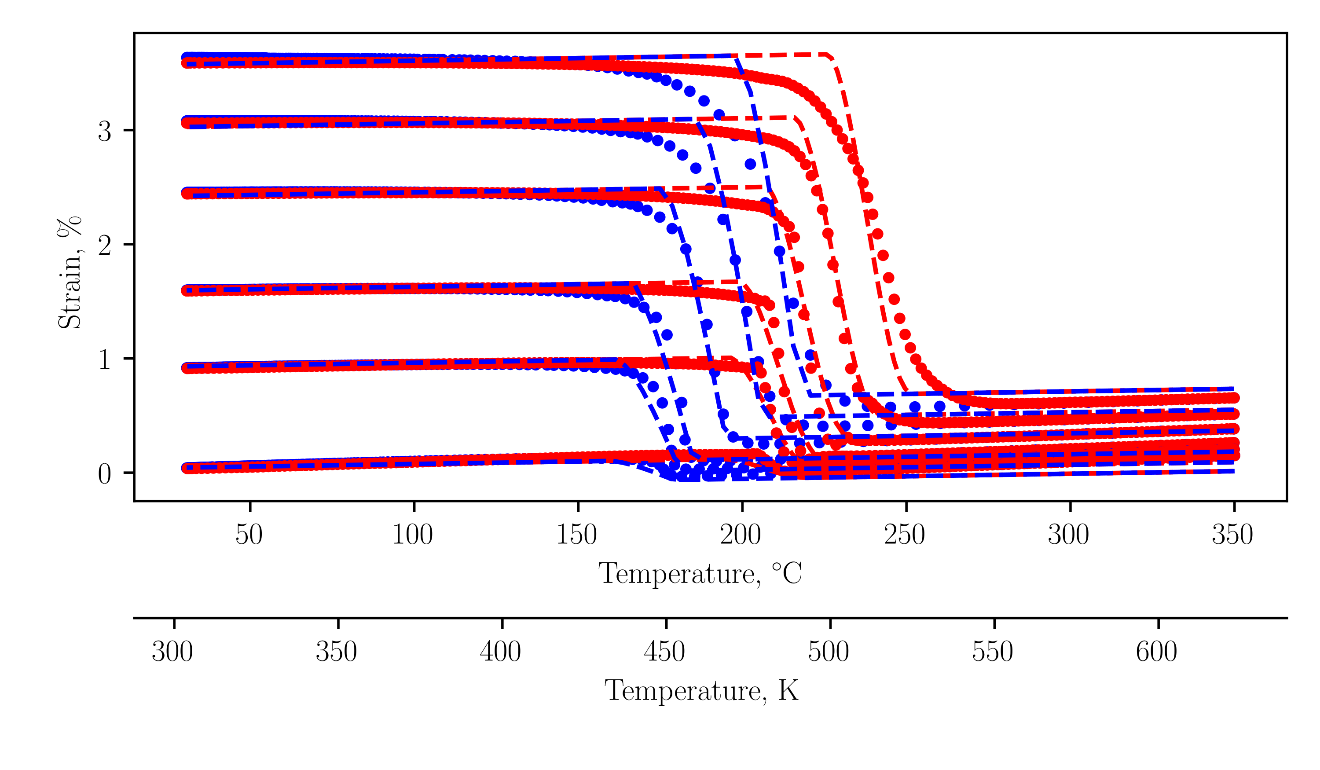


Figure 6: The final calibration agrees with the experimental data to within 1.5% mean squared error.

The final numerical calibration is depicted in Figure 6. The model predicts the elastic response in martensite almost perfectly, which signifies that both the martensitic elastic modulus and transformation strain properties are well calibrated. Additionally, transformation temperatures show good agreement at low levels of applied stress. At higher levels of applied stress, the model-predicted transformation starts to overshoot the experimental data then predict a sharper minor loop. This is because the transformation temperatures are not a linear function of stress (i.e., the stress-influence coefficients are not constant, see figure BLANK in Bigelow (cite)), and because the smoothness of transformation initiation is not constant with stress (compare the 100 MPa transformation into austenite with the analogous location at 300 MPa). This calibration is a perfect example of the utility of numerical optimization; the optimizer finds the best global fit of data, especially with regards to the austenite transformation temperatures. For lower stresses, is too early, and is too late. At intermediate stresses, like 100 and 200 MPa, the transformation temperatures are almost perfect. Then, at 300 MPa, is too late and is too early. This could be better fit at the relevant stresses by biasing the solution to prioritize fitting certain stress levels (see my paper) or by simply calibrating the model at the stress levels that matter most.

However, this calibration is not perfect, mainly due to model deficiencies. The coefficient of thermal expansion is not constant for austenite and martensite. Clearly, the coefficient of thermal expansion in austenite is larger than that in martensite. This is a model deficiency because the current model uses a convex cutting plane assumption for numerical integration and could be improved in future work.

Regardless, these five optimizations improved calibration accuracy by over 50% compared to the conventional calibration, and were accomplished in less than an hour on a lightweight laptop with a low-performance processor (Intel Core m3-6Y30 CPU @ 0.90 GHz with 4 Gb RAM). Herein lies the main contribution of this work: this calibration routine can be performed by general analysts, designers, or material scientists, without the need for exotic hardware or python programming experience.

**Conclusions and further refinements**