Development of a Numerical Model for High-Temperature Shape Memory Alloys

Jonathan A. DeCastro
QSS Group, Inc., Cleveland, Ohio

Kevin J. Melcher and Ronald D. Noebe
Glenn Research Center, Cleveland, Ohio

Darrell J. Gaydosh
Ohio Aerospace Institute, Brook Park, Ohio

July 2006
Since its founding, NASA has been dedicated to the advancement of aeronautics and space science. The NASA Scientific and Technical Information (STI) program plays a key part in helping NASA maintain this important role.

The NASA STI Program operates under the auspices of the Agency Chief Information Officer. It collects, organizes, provides for archiving, and disseminates NASA's STI. The NASA STI program provides access to the NASA Aeronautics and Space Database and its public interface, the NASA Technical Reports Server, thus providing one of the largest collections of aeronautical and space science STI in the world. Results are published in both non-NASA channels and by NASA in the NASA STI Report Series, which includes the following report types:

• TECHNICAL PUBLICATION. Reports of completed research or a major significant phase of research that present the results of NASA programs and include extensive data or theoretical analysis. Includes compilations of significant scientific and technical data and information deemed to be of continuing reference value. NASA counterpart of peer-reviewed formal professional papers but has less stringent limitations on manuscript length and extent of graphic presentations.

• TECHNICAL MEMORANDUM. Scientific and technical findings that are preliminary or of specialized interest, e.g., quick release reports, working papers, and bibliographies that contain minimal annotation. Does not contain extensive analysis.

• CONTRACTOR REPORT. Scientific and technical findings by NASA-sponsored contractors and grantees.

• CONFERENCE PUBLICATION. Collected papers from scientific and technical conferences, symposia, seminars, or other meetings sponsored or cosponsored by NASA.

• SPECIAL PUBLICATION. Scientific, technical, or historical information from NASA programs, projects, and missions, often concerned with subjects having substantial public interest.

• TECHNICAL TRANSLATION. English-language translations of foreign scientific and technical material pertinent to NASA’s mission.

Specialized services also include creating custom thesauri, building customized databases, organizing and publishing research results.

For more information about the NASA STI program, see the following:

• Access the NASA STI program home page at http://www.sti.nasa.gov

• E-mail your question via the Internet to help@sti.nasa.gov

• Fax your question to the NASA STI Help Desk at 301–621–0134

• Telephone the NASA STI Help Desk at 301–621–0390

• Write to:
  NASA STI Help Desk
  NASA Center for AeroSpace Information
  7121 Standard Drive
  Hanover, MD 21076–1320
Development of a Numerical Model for High-Temperature Shape Memory Alloys

Jonathan A. DeCastro  
QSS Group, Inc., Cleveland, Ohio

Kevin J. Melcher and Ronald D. Noebe  
Glenn Research Center, Cleveland, Ohio

Darrell J. Gaydosh  
Ohio Aerospace Institute, Brook Park, Ohio

National Aeronautics and Space Administration

Glenn Research Center  
Cleveland, Ohio 44135

July 2006
Development of a Numerical Model for High-Temperature Shape Memory Alloys

Jonathan A. DeCastro
QSS Group, Inc.
Cleveland, Ohio 44135

Kevin J. Melcher and Ronald D. Noebe
National Aeronautics and Space Administration
Glenn Research Center
Cleveland, Ohio 44135

Darrell J. Gaydosh
Ohio Aerospace Institute
Brook Park, Ohio 44142

Abstract

A thermomechanical hysteresis model for a high-temperature shape memory alloy (HTSMA) actuator material is presented. The model is capable of predicting strain output of a tensile-loaded HTSMA when excited by arbitrary temperature-stress inputs for the purpose of actuator and controls design. Common quasi-static generalized Preisach hysteresis models available in the literature require large sets of experimental data for model identification at a particular operating point, and substantially more data for multiple operating points. The novel algorithm introduced here proposes an alternate approach to Preisach methods that is better suited for research-stage alloys, such as recently-developed HTSMAs, for which a complete database is not yet available. A detailed description of the minor loop hysteresis model is presented in this paper, as well as a methodology for determination of model parameters. The model is then qualitatively evaluated with respect to well-established Preisach properties and against a set of low-temperature cycled loading data using a modified form of the one-dimensional Brinson constitutive equation. The computationally efficient algorithm demonstrates adherence to Preisach properties and excellent agreement to the validation data set.

1. Introduction

In aeronautics propulsion, the transition to an all-electric engine is desirable from reliability and efficiency standpoints. By replacing conventional hydraulic actuation systems with smart material actuators in such components as valves, vanes, nozzles, it may be possible to realize substantial improvements in the overall size and weight of the engine, allowing higher efficiencies to be realized. Furthermore, the high energy output and bandwidth of such materials enable other lucrative control capabilities not previously within reach, such as active inlets, variable area nozzle control, and blade tip clearance control in the high-pressure turbine (refs. 1 to 4). In the active clearance control example, an array of actuators could be used to actively position several floating turbine shroud segments to within close proximity of the turbine blades using closed-loop control, which expand and contract during flight. In such an application, the smart material actuator must be capable of high-amplitudes of motion (~0.1 in.) amidst large bias loads (~2000 lbf) that can vary significantly during normal operation. Furthermore, such actuators must be capable of moderate bandwidths (on order of 1 Hz), be robust at elevated temperatures (i.e., long cycle life), consume minimal energy to operate, and be compact and light weight.
The particular class of smart materials known as shape memory alloys (SMAs) has emerged in several application areas including robotics, control of space-borne structures, and control of helicopter rotor blades (refs. 5 and 6). Some SMA materials, such as Nitinol, are capable of macroscopic strains up to 6 percent at stresses up to 500 MPa (ref. 7). In comparison with piezoceramics, for example, these capabilities amount to more than an order-of-magnitude higher energy output per unit density of material. Consequently, SMAs are well-suited for applications where size and weight are of high importance in actuator feasibility, as is the case with any propulsion system. Another advantage is that SMAs have similar manufacturability to conventional metals and, as such, are readily produced in structurally necessary forms: wires, rods, and sheets. Lastly, the unique thermal activation properties of SMAs give them a distinct advantage over other smart actuation methods because it is possible to circumvent potentially costly modifications to the propulsion power system that otherwise would be required (ref. 8). Because hot and cool air is often diverted from various stages in the engine in order to provide cooling to various high-temperature components, small portions of this secondary air may be actively modulated by way of a proportional valve. Therefore operation of the SMA actuator via active convective cooling may be possible in lieu of more conventional active resistance heating methods that require large power draws.

A functional drawback of SMAs is that the phase transitions that produce work are governed by hysteresis. Furthermore, with widely-available Nickel-Titanium (NiTi) alloys, phase transition temperatures and useable range of operation reside close to room temperature, rendering them infeasible for moderate- to high-temperature applications. To overcome these limitations, a Nickel-Platinum-Titanium (Ni20Pt30Ti50) high-temperature shape memory alloy (HTSMA) has been recently developed at NASA Glenn Research Center. This alloy exhibits work outputs comparable to conventional NiTi alloys, but with transformation temperatures greater than 250 °C (480 °F) and oxidation rates that are less than half that of typical NiTi SMA materials (fig. 1). This is a significant advancement over conventional SMA materials, enabling the possibility for actuation in an engine’s hot sections.

In partial alleviation of the hysteresis problem, HTSMA hysteresis deadband characteristics are narrower and hysteresis profiles shallower than many conventional NiTi alloys. Because it is not possible to fully eliminate inherent hysteresis, more sophisticated control strategies that guarantee stability, maximize HTSMA bandwidth, and maintain uniform response across the entire range of operation may still be required (refs 9 to 11). In any case, reliable development and evaluation of candidate actuator control systems usually require accurate models of the HTSMA. In addition, such models must be executable in a real-time environment in order for model-based control or estimation to be implemented.

Generally speaking, SMA models consist of two parts: a hysteresis model and a constitutive model. To accurately model hysteresis, minor hysteresis loops in the response trajectory must be accounted for, although inclusion of such detail substantially complicates the model and is an active area of research.
It has been shown that the generalized Preisach model of hysteresis is suited for describing SMA hysteresis and obeys certain empirically-verified properties (refs. 12 to 14). The usefulness of generalized Preisach models is fully exploited only when a full database of empirical response data is available, usually in the form of probability density functions (PDFs) obtained from a population of first-order descending (FOD) curves (ref. 15). In the case of the HTSMA, however, this database does not yet exist due to the time and costs associated with conducting high temperature experiments. Furthermore, accurate constitutive models are still needed for estimation of internal SMA variables, leading to potential inaccuracies in the results. Since continuous-function PDF approximants have been shown to represent SMA behavior quite well under quasi-static situations (ref. 16), it is reasonable to adopt this approach for the HTSMA application. With these approximants, algorithmic hysteresis models that are both conceptually simple and easy to implement can be used as long as they are capable of capturing minor loop hysteresis phenomena.

Madill and Wang (ref. 16) developed an efficient algorithm for SMA control simulations to compute martensite volume fraction, an internal state variable, as temperature varies between intermediate points. However, memory of the trajectory history was limited to only the prior hysteresis loop, not every prior dominant loop, and hence violates well-established Preisach properties. Efficient, yet accurate algorithmic hysteresis models that memorize multiple prior dominant loops have been developed for electromagnetic power circuits by Casoria, et al. (ref. 17). To the authors’ knowledge, however, such models have not been applied in a similar manner to shape memory alloys. Moreover, a thorough exposition of how these models were implemented has not been made widely available.

In this work, a numerical, algorithmic realization of HTSMA hysteresis is presented as an alternative to analytic generalized Preisach models based upon PDF data. In its simplicity, the algorithm offers an intuitive approach to SMA hysteresis in lieu of the ‘black box’ approach offered by generalized Preisach models. As such, one can directly define physical properties of the material or directly manipulate the constitutive equations as an alternative to obtaining sets of PDF identification data as new HTSMA materials become available. Although applicable to any of the several SMA constitutive models reported, the Brinson phenomenological constitutive model is chosen here because it possesses the fidelity to be applied across a wide range of stress-strain conditions (ref. 18).

The remainder of this paper is organized in the following manner. Section 3 describes the model architecture and development. In this section, the SMA constitutive equations are introduced, a modification to the transformation kinetics model is proposed, and incorporation of the minor-loop hysteresis algorithm is presented. Section 4 details experimental characterization of the HTSMA material and methodologies used in determining the necessary model parameters. The results of model validation are presented in section 5. Here, qualitative adherence to known Preisach properties and quantitative agreement to available experimental data is provided. A summary of the model’s capabilities and recommended future work is given in section 6.

### 2. Nomenclature

\(A\) \hspace{1cm} \text{region of austenite transformation} \\
\(A_f\) \hspace{1cm} \text{temperature at which austenite transformation is completed or finished} \\
\(A_s\) \hspace{1cm} \text{temperature at which austenite transformation starts} \\
\(C\) \hspace{1cm} \text{stress influence coefficient} \\
\(E\) \hspace{1cm} \text{modulus of elasticity} \\
\(F\) \hspace{1cm} \text{value of } \xi_n \text{ when } d\xi_n/dt \text{ changes sign} \\
\(F_{\text{min}}\) \hspace{1cm} \text{vector containing past minima} \\
\(F_{\text{max}}\) \hspace{1cm} \text{vector containing past maxima} \\
\(M\) \hspace{1cm} \text{region of martensite transformation}
$M_f$ temperature at which martensite transformation is completed or finished
$M_s$ temperature at which martensite transformation starts
$S$ region of stress-induced martensite transformation
$T$ temperature
$b$ kinetic curve bias
$dt$ differential element of time
$i$ integer
$f$ kinetic curve describing minor loop transformation
$k$ integer
$n$ integer
$\Theta$ coefficient of thermal expansion
$\Phi$ kinetic curve scaling factor
$\Omega$ phase transformation coefficient
$\varepsilon$ material strain
$\varepsilon_L$ maximum recoverable strain
$\xi_{\text{total}}$ total martensite fraction
$\xi_{S}$ stress-induced martensite fraction
$\xi_{T}$ temperature-induced martensite fraction
$\xi_{n}$ normalized martensite fraction
$\sigma$ applied stress
$\sigma_{cr}$ critical stress

Subscripts:
0 referring to an initial state
$A$ referring to austenite transformation
$Af$ referring to temperature at which austenite transformation finishes
$As$ referring to temperature at which austenite transformation starts
$M$ referring to martensite transformation
$Mf$ referring to temperature at which martensite transformation is completed or finished
$Ms$ referring to temperature at which martensite transformation starts
$S$ stress-induced component
$T$ temperature-induced component
$max$ maximum value
$min$ minimum value
$s$ starting point
$f$ finish or final point

Superscripts:
$A\rightarrow M$ transformation from austenite to martensite
$A\rightarrow S$ transformation from austenite to stress-induced martensite
$M\rightarrow A$ transformation from martensite to austenite
transformation from stress-induced martensite to austenite

critical state

3. HTSMA Phenomenological Model

SMAs generally exist in the austenite phase at high temperatures and the martensite phase at low temperatures. Functionally, drastic shape changes may arise when transitioning between these two crystallographic material phases. Large macroscopic work may accompany the resulting shape change when the parent austenite phase is higher in strength than the lower-temperature martensite phase. In such cases, the parent phase exhibits high-stiffness Hooke’s law behavior while the martensite exhibits a lower-stiffness, pseudoplastic response. Transition paths are completely reversible, but are dominated by major loop hysteresis as shown in figure 2(a). Minor loop hysteresis is prevalent when transition changes occur before transformation is complete, as shown in figure 2(b); and is a more difficult phenomenon to model. The specific parameters used to model the shape memory effect are discussed in detail later in this section.

The HTSMA model consists of three parts: a transformation kinetics model describing major loop hysteresis, a generalized Preisach hysteresis algorithm relating temperature and stress to martensite fractions, and a constitutive equation relating martensite fractions to strain output under tensile loading. The various elements of the model are implemented as shown in figure 3 using MATLAB/Simulink (The Mathworks, Inc.). By adopting a modular implementation, the hysteresis model can be coupled with other models, such as heat transfer models or higher-fidelity constitutive models, depending on the particular needs facing the designer.
Brinson Quasi-Static Constitutive Model

SMA actuators are typically implemented in a wire bundle arrangement due to ease of manufacture, fast heat transfer characteristics, failure redundancy, and conduciveness to resistive heating methods (ref. 19). A constitutive model based on uniaxial tensile loads representative of alloy wires is therefore an appropriate choice. Of the various phenomenological models available in the literature, none have broader applicability than the Brinson model (ref. 18) because of distinctions made between separate stress- and temperature-induced martensite fractions. This is especially important for predicting stress-related effects at low temperature.

The region where martensite transitions between its variants independent of temperature is denoted by $S$ in the stress-temperature phase plot in figure 4(a). As stress is applied through region $S$, the alloy transitions from temperature-induced martensite to stress-induced martensite, causing a pseudoplastic deformation and large macroscopic strains. The thermomechanical explanation of this effect is that it is possible for multiple randomly-oriented martensite variants to coexist when martensite is in the “temperature-induced” state. In contrast, when martensite becomes completely “stress-induced,” only the variant that corresponds to the direction of the applied stress is in equilibrium. Hence, the two-way SMA energy output increases substantially with applied stress. When stress is again relieved, the martensite remains in the stress-induced state and maintains an elongated shape. Strain restoration cannot be achieved unless the material is heated to austenite, a process commonly referred to as the shape memory effect. Regions where austenite transformation occurs, denoted $A$ in figure 4(a), and martensite occurs, denoted $M$, have an approximate linear dependence with applied stress. Variations in stress above $A_f$, but below the permanent martensite deformation temperature, result in complete strain recovery, known as the pseudoelastic effect.

The Brinson model describes the thermomechanical evolution by first separating total martensite fraction $\xi$ into a stress-induced martensite component $\xi_s$ and temperature-induced component $\xi_T$:

$$\xi = \xi_s + \xi_T$$  \hspace{1cm} (1)

The constitutive equation relating the two martensite states and the strain $\varepsilon$, stress $\sigma$, and temperature $T$ is as follows:

$$\sigma - \sigma_0 = E(\xi)\varepsilon - E(\xi_0)\varepsilon_0 + \Omega(\xi)\varepsilon_s - \Omega(\xi_0)\varepsilon_{s0} + \Theta(T - T_0).$$  \hspace{1cm} (2)
Here the subscript 0 denotes the initial states, while Θ is the coefficient of thermal expansion (CTE). In addition, the modulus of elasticity, \(E\), is non-constant and defined as,

\[E(\xi) = E_A + \xi(E_M - E_A),\]

with subscripts \(A\) and \(M\) referring to austenite and martensite, respectively. Furthermore, the phase transformation coefficient used in equation (2) is defined as,

\[\Omega(\xi) = -\varepsilon_L E(\xi),\]

where \(\varepsilon_L\) is the maximum recoverable strain.

The boundaries of the stress-induced martensite region \(S\) are defined by the starting critical stress \(\sigma_s^{cr}\) and the finish critical stress \(\sigma_f^{cr}\). Bounding stresses for the austenite region \(A\) may be defined by

\[\sigma_s^{cr} = C_A s(T - A_s)\]
\[\sigma_f^{cr} = C_A f(T - A_f)\]

Similarly, the martensite region \(M\) is bounded by

\[\sigma_s^{cr} = C_M s(T - M_s)\]
\[\sigma_f^{cr} = C_M f(T - M_f)\]

The crosshatched region in figure 4(a) represents the location where the two martensite regions coincide. Here \(A_s, A_f, M_s,\) and \(M_f\) are defined as the austenite start, austenite finish, martensite start, and martensite finish temperatures, respectively, evaluated at zero stress. The stress influence coefficients \(C_A, C_A, C_M,\) and \(C_M\) are the corresponding slopes of the transition temperature-stress boundaries. The hysteresis model describing the relationship between temperature and martensite fraction within the critical stress regions is outlined in the following sections.

**Modification to Original Brinson Model**

The cosine transformation kinetic approximant used to describe major loop behavior of both total and stress-induced martensite fraction was initially proposed by Liang and Rogers (ref. 20) and later modified by Brinson (ref. 21) to accommodate stress-induced martensite. For transitions from austenite to martensite, Brinson’s model describing stress-induced martensite fraction encounters discontinuities in situations when temperature decreases through \(M\), with \(\sigma_s^{cr} < \sigma < \sigma_f^{cr}\), as depicted by the path 1\(\rightarrow\)2 in figure 4(a). As the kinetic equations are written, transformation of \(\xi_S\) progresses from finite slope to zero slope at a point along the kinetic curve corresponding to the instantaneous stress at \(M\), instead of proceeding to the lower bound of the kinetic curve. To establish continuity, the relationship \(\xi_n = \xi_S / \xi\) is introduced and hysteresis is instead described by martensite fraction \(\xi\) and normalized martensite fraction \(\xi_n\). This substitution modifies the original phase transformation regions from those shown in figure 4(a) to those shown in figure 4(b). The modified Brinson model further enables \(M\) to be a linear function of stress (instead of stress-independent), permitting closer agreement between the model and data at stresses below \(\sigma_s^{cr}\).
Transformation from austenite to martensite is represented by the following:

\[
\xi_{A \rightarrow M}(\sigma, T) = \frac{1}{2} \cos \left[ \pi \frac{T - M_f - \sigma C_M^{-1}}{M_s - M_f + \sigma (C_M^{-1} - C_M)} \right] + \frac{1}{2},
\] (7a)

\[
\xi_{n \rightarrow A}(\sigma, T) = \frac{1}{2} \cos \left[ \pi \frac{\sigma - \sigma_{cr}^f}{\sigma_{cr} - \sigma_{cr}^f} \right] + \frac{1}{2}.
\] (7b)

The superscripts designate the direction of transformation, i.e., \(A \rightarrow M\) = austenite to martensite; \(A \rightarrow S\) = austenite to stress-induced martensite. The cross-hatched region in figure 4(b) represents the intersection of all non-saturated values of \(\xi_{A \rightarrow M}\) and all non-saturated values of \(\xi_{n \rightarrow S}\) (all nonzero values of \(\xi_{n \rightarrow S}\)). The only other notable difference between this model and past realizations is that independent stress influence coefficients are now applied to \(M_s\) and \(M_f\).

For transition from \(M \rightarrow A\) the transformation kinetics are represented by

\[
\xi_{M \rightarrow A}(\sigma, T) = \frac{1}{2} \cos \left[ \pi \frac{T - A_s - \sigma C_A^{-1}}{A_f - A_s + \sigma (C_A^{-1} - C_A)} \right] + \frac{1}{2},
\] (8a)

\[
\xi_{n \rightarrow S}(\sigma, T) = F \xi_{M \rightarrow A}(\sigma, T),
\] (8b)

where \(F\) represents the value of \(\xi_{n \rightarrow A}\) at the instant when \(d\xi_{n \rightarrow A}/dt\) changes sign. In figure 4(b), transformations in stress-induced martensite and total martensite to austenite are both represented by \(A\), as equation (8) reflects. Because \(\xi_{M \rightarrow A}\) and \(\xi_{n \rightarrow S}\) are coincident, \(F\) may be thought of as a scalar on martensite fraction in determining \(\xi_{n \rightarrow S}\). For implementation of the model, the basis variables of equations (7) and (8) are passed as inputs to the minor loop computation algorithm.

**Minor Loop Algorithm**

To illustrate the phenomenological behavior of minor hysteresis loops, consider trajectory 4-5 in figure 2(b). In this example, the trajectory initially follows the transition path that intersects point 3 as temperature decreases. After the intersection with 3, a new path is followed, intersecting point 1. Transition paths are dependent on prior extrema; however, as smaller loops are overtaken by larger ones, the smaller extrema are eliminated from memory. This is described succinctly by the *wiping out property* of the generalized Preisach model. Once the material reaches point 1 (also point 5), the double minor loop is closed and all extrema are eliminated from memory. An additional property, known as the minor loop *congruency property*, is required to complete the generalized Preisach model.

Within the Preisach modeling paradigm, it has been shown that reasonable approximations to the PDF can be obtained using information contained in major loop data (ref. 15). Therefore, monotonic variations in temperature during minor loop formation may be described by the major loop, which is scaled and translated based on the past history of temperature (or martensite fraction) extrema. Representing Preisach hysteresis during transition from martensite to austenite is accomplished in a piecewise manner that depends on the direction of transformation.
\[
f(\sigma, T) = \begin{cases} 
\max \left[ \Phi_{A\to M} f^{A\to M}(\sigma, T) + b_{A\to M} \right] & \text{for } df/dt > 0 \\
\min \left[ \Phi_{M\to A} f^{M\to A}(\sigma, T) + b_{M\to A} \right] & \text{for } df/dt < 0 
\end{cases}
\]

where the scalars and biases are

\[
\Phi_{A\to M} = \frac{F_{\min} - F_{\max}}{F_{\min}^{A\to M} - F_{\max}^{A\to M}}
\]

\[
\Phi_{M\to A} = \frac{F_{\min} - F_{\max}}{F_{\min}^{M\to A} - F_{\max}^{M\to A}}
\]

\[
b_{A\to M} = F_{\max} - \Phi_{A\to M} F_{\max}^{A\to M}
\]

and

\[
b_{M\to A} = F_{\min} - \Phi_{M\to A} F_{\min}^{M\to A}
\]

Here \( f \) is used to represent both \( \zeta \) and \( \zeta_n \), as algorithms are identical for both variables. It follows that \( f^{A\to M} \) and \( f^{M\to A} \) are the associated kinetic curves (for \( \zeta_n \), superscript \( M \) is replaced with \( S \)). \( F_{\min} \) contains the past \( n \) minima of \( f \) and \( F_{\max} \) contains the past \( n \) maxima of \( f \), collected at instants when \( f \) changes sign. \( F_{\min_{\text{max}}}^{A\to M} \) and \( F_{\max_{\text{max}}}^{M\to A} \) are analogous vectors corresponding to \( f^{A\to M} \) and \( f^{M\to A} \).

In order to satisfy the wiping out property, the hysteresis model must discard past extrema that are smaller in magnitude than the most recent value. Each extrema vector is subject to the following conditions:

- For \( i = 1, \ldots, n \), if \( df/dt > 0 \) and \( F_{\min}(i) > F_{\min}(1) \) (1 being the most recent), then
  i. Let \( F_{\min}(i) = F_{\min}(1) \)
  ii. Let \( F_{\max}(i) = \min_{k} F_{\max}(k) \) with \( k = 1, \ldots, n \) and \( k \neq i \)

  This modifier is applied to the remaining two extrema vector pairs \( F_{\min}^{A\to M} \), \( F_{\max}^{A\to M} \) and \( F_{\min}^{M\to A} \), \( F_{\max}^{M\to A} \).

- For \( i = 1, \ldots, n \), if \( df/dt < 0 \) and \( F_{\max}(i) < F_{\max}(1) \) (1 being the most recent), then
  i. Let \( F_{\max}(i) = F_{\max}(1) \)
  ii. Let \( F_{\min}(i) = \max_{k} F_{\min}(k) \) with \( k = 1, \ldots, n \) and \( k \neq i \)

  This modifier is applied to the remaining two extrema vector pairs \( F_{\max}^{A\to M} \), \( F_{\min}^{A\to M} \) and \( F_{\max}^{M\to A} \), \( F_{\min}^{M\to A} \).

In the simulation, the number of memorized loops \( n \) is set to 10 to ease computational efficiency. To eliminate the feedback path, the sign of \( df/dt \) for detection of the turn-around points is computed instead by testing the sign of \( df^{A\to M}/dt \) and \( df^{M\to A}/dt \). If either one changes sign, then an extremum has
been detected. The algorithm may encounter situations where small input variations or noise may cause oscillations in the extremum detector, therefore a threshold was incorporated in the detection scheme to reject these oscillations. It is known that hysteresis response approaches a linear characteristic with small driving amplitudes, thus thresholding will not compromise the model’s accuracy. Within the MATLAB/Simulink environment, the hysteresis model was implemented as an Embedded MATLAB Function, which was called once every time a martensite fraction extremum is detected.

4. HTSMA Characterization

Experimentally-identified properties of the HTSMA that were used to parameterize the present model are given in table 1. Additional data and explanation of the properties are provided in reference 0. The data was collected from several forms of a Ni₃₀Pt₂₀Ti₅₀ alloy that has been identified for high-temperature actuator applications and is undergoing additional development and optimization at NASA Glenn. It should be noted, however, that material properties are highly dependent on variations in thermomechanical processing and minor deviations in composition.

<table>
<thead>
<tr>
<th>Parameter</th>
<th>Value</th>
<th>Parameter</th>
<th>Value</th>
</tr>
</thead>
<tbody>
<tr>
<td>Martensite start temperature ( M_s )</td>
<td>510 °F</td>
<td>( M_s ) stress influence coefficient ( C_{M_s} )</td>
<td>0.83 ksi/°F</td>
</tr>
<tr>
<td>Martensite finish temperature ( M_f )</td>
<td>463 °F</td>
<td>( M_f ) stress influence coefficient ( C_{M_f} )</td>
<td>1.44 ksi/°F</td>
</tr>
<tr>
<td>Austenite start temperature ( A_s )</td>
<td>505 °F</td>
<td>( A_s ) stress influence coefficient ( C_{A_s} )</td>
<td>1.78 ksi/°F</td>
</tr>
<tr>
<td>Austenite finish temperature ( A_f )</td>
<td>527 °F</td>
<td>( A_f ) stress influence coefficient ( C_{A_f} )</td>
<td>0.83 ksi/°F</td>
</tr>
<tr>
<td>Start critical stress ( \sigma_{cr}^s )</td>
<td>33.4 ksi</td>
<td>Elastic modulus of martensite 200 °C ( E_{M_s} )</td>
<td>10152 ksi</td>
</tr>
<tr>
<td>Finish critical stress ( \sigma_{cr}^f )</td>
<td>128.6 ksi</td>
<td>Elastic modulus of austenite (300 to 600 °C) ( E_A )</td>
<td>12328 ksi</td>
</tr>
<tr>
<td>Max. recoverable strain ( \varepsilon_L )</td>
<td>3.3 %</td>
<td>Coefficient of thermal expansion ( \Theta )</td>
<td>1.02e-5 1/K</td>
</tr>
</tbody>
</table>

The Young’s moduli and room temperature tensile properties in table I were specifically obtained from measurements on a 0.060-in.-diameter rod that was produced by a multiple extrusion technique from a 1-in.-diameter cast and homogenized billet, resulting in a uniform, fine-grained material. The sample is representative of the material in wire form, which is the most likely product form for actuator applications. Difficulties in measuring material modulus by direct techniques, i.e., extensometry or strain gages, arise from the extreme dependence of modulus on strain. Instead, modulus measurements were determined on the as-processed rod by an impulse frequency excitation technique.

Values for critical stresses were determined in a similar manner to Epps and Chopra (ref. 23), at approximate points where the room temperature stress-strain characteristic deviates from linear elastic regimes that correspond, in theory, to the start and finish of the stress-induced martensite transformation \( (\xi_S = 0 \rightarrow 1) \). In practice, the start critical stress \( \sigma_{cr}^s \) is an approximation to the start of the nonlinear regime and the finish critical stress \( \sigma_{cr}^f \), coupled with the stress-free maximum recoverable strain \( \varepsilon_L \), are chosen to obtain adequate model agreement to the slope of the transition region below 3 percent strain. Based on experience with conventional shape memory alloys, the material will experience reduced fatigue life at strain levels above approximately 3 percent, therefore any application utilizing these alloys should prevent strains from exceeding those levels. Since there is no practical way of measuring stress-induced martensite fraction \( \xi_S \) to make correlation with \( \sigma_{cr}^f \) and \( \varepsilon_L \) possible, the model-calculated value of \( \xi_S \) loses physical meaning, even with a perfect model fit. If the model is indeed exercised within the valid strain range, the overall modeling accuracy will not be compromised.

The tensile curve used to determine these parameters is shown in figure 5 with the model fit overlaid on the plot. The sample used for the experiment was a thin rod that was heat-treated at 1112 °F (600 °C) to regain the shape memory properties that were suppressed due to the significant level of hot-working in
the alloy. The tensile properties of the heat-treated rod are very similar to those for the as-processed rod (provided in table I), except for an increase in the yield strength in the heat-treated material. The linear region on the right-hand side of the model curve illustrates the material’s actual behavior when the strain is relieved following an initial stress increase past $\sigma_{cr}^f$. Note that, by decreasing the stress to zero, the value of the modeling parameter $\varepsilon_L$ becomes apparent.

A plot of the phase transformation temperatures as a function of stress for the alloy is shown in figure 6, along with the fitted boundaries, using the values listed in table I. The data set was determined by measuring the strain-temperature response of the material under various levels of constant stress. These tests were performed on cast and homogenized ingots that were extruded at 2012 °F (1100 °C) at a reduction ratio of 7:1. This simple, high-temperature processing resulted in a very course grained
material. Coefficient of thermal expansion (CTE) was obtained via dilatometry testing using the same course grained rod. Figure 6 illustrates that the martensite and austenite transition temperatures have less overlap but tighter transformation bands at low stress, but larger overlap and wider transformation bands at larger stresses. Parameters in the linear stress-temperature relationships of equations (5) and (6) were adjusted to approximate the respective data for the transformation boundary regions. It should be noted that the zero-stress transition temperatures in the data were considered as outliers and were removed from the parameterization.

5. Model Evaluation

The goal of the evaluation is to validate the full model against both established hysteresis principles and a set of stress-cycled data. The modified Brinson major loop model was compared with responses with the original representation, while the minor loop hysteresis model was validated against well-established Preisach properties. The complete model was then validated against a set of stress-cycled experimental data. In each of the following simulations, the CTE was set to zero and the simulation increment was set to 1 ms.

Major Loop Simulation Tests

Major loop evaluation was conducted by cycling the internal temperature between $M_f$ and $A_f$ at several constant stress conditions. Figure 7 shows the result of the evaluation at the four stress conditions given in table II. It should be noted that test points 2 to 4 are considered to be in the extrapolated range of the phase plot in figure 6 and test points 3 and 4 are considered to be outside the valid strain range in figure 5. The figure highlights dependence on peak-to-peak strain on stress-induced martensite fraction. At test point 1, stress-induced martensite remains close to zero throughout the cycle, since the alloy stress is below the start critical stress $\sigma_{cr}$. The resulting peak-to-peak strain in figure 7, or equivalently, the difference between martensite strain and austenite strain, is 0.05 percent. At test points 2, 3, and 4, stress-induced martensite peaks at 0.33, 0.82, and 0.98, respectively, causing reversible peak-to-peak strains to grow accordingly. Due to the dependence of stress on transformation temperatures, the hysteresis curves shift to higher temperatures and decrease in slope as applied stress increases. Discontinuity in $\xi_S$ response from the original Brinson model at the stress-independent intersection of $M$ and $S$ (at ~290 °C) is eliminated in the modified version by addition of the new variable $\xi_n$, as can be noted in the figure.

<table>
<thead>
<tr>
<th>Test point</th>
<th>Description</th>
<th>Stress, ksi</th>
</tr>
</thead>
<tbody>
<tr>
<td>1</td>
<td>Stress below critical region, $\sigma &lt; \sigma_{cr}$</td>
<td>30</td>
</tr>
<tr>
<td>2</td>
<td>Stress within critical region, $\sigma_{cr} &lt; \sigma &lt; \sigma_f$</td>
<td>70</td>
</tr>
<tr>
<td>3</td>
<td>Stress within critical region, $\sigma_f &lt; \sigma &lt; \sigma_{cr}$</td>
<td>100</td>
</tr>
<tr>
<td>4</td>
<td>Stress above critical region, $\sigma &gt; \sigma_f$</td>
<td>130</td>
</tr>
</tbody>
</table>
High Temperature Minor Loop Simulation Tests

The next set of tests focused on evaluating minor hysteresis loop prediction. For all temperature-cycled evaluation tests, a constant stress from test point 1 was used. The first two tests aim to verify adherence to Preisach hysteresis that is described by the following two properties: minor loop congruency and wiping out of past extrema. The second two tests illustrate the prediction capability given more complex loading profiles.

The congruency property test in figure 8 was generated from a temperature input that produces two opposing loops. The first loop occurs during transition along the $M \to A$ kinetic curve and the second occurs along the $A \to M$ curve with the same temperature reversal points as the first, designated by paths 2-3-4 and 6-7-8, respectively. It is evident from figure 8 that the correct hysteresis trajectory is developed, as the two minor loops appear to be congruent in shape. Nevertheless, the scaling factors $\Phi_{M\to A}$ and $\Phi_{A\to M}$ are slightly different between the two minor loop segments, resulting in a small discrepancy in the sizes of the two minor loops. It is important to note that, although equation (9) scales minor loop segments from major loop bases, the algorithm does not guarantee congruency of two vertically-shifted loops. To
investigate congruency adherence further, a profile of minor loop congruency inaccuracies across the range of operation was generated by performing the congruent loop test at 1000 randomly-generated stresses and temperatures. Figure 9 shows differences between the minor loop scaling factors for the first and second loop as functions of minor loop size. The plot also shows differences between minor loop sizes between the two loops versus minor loop size for the same set of tests. Results indicate that scaling factor errors are large for small magnitude loops, but become smaller as magnitudes increase. A similar trend can be seen in the magnitude errors, however, large errors persist over a wider range of loop magnitudes. The pseudorandom nature of the errors confirms that, on average, the congruency property is satisfied. Based on the data it is hypothesized that truncation errors in the denominator of $\Phi$ in equation (9) cause differences in $\Phi$ to grow as minor loops become small. One fix to resolve this issue is to implement an amplitude threshold to prevent truncation errors from dominating the response with small loops.

In the wiping out test, shown in figure 10, a past extrema pair that makes up a minor loop must be “wiped out” when one of the two points is overtaken. The property is clearly demonstrated in the figure, as the first two decaying minor loops (path 2-3-4-5-6-7) are replicated one-for-one with the second two loops (7-3-4-5-6-7). For verification, investigating the sequence shows that both extrema from the innermost loop are indeed eliminated from memory as the second half of the trajectory begins, resulting
in reconstruction of the larger minor loop. It is thus concluded that the algorithm correctly adheres to the wiping out property.

Results of the multiple decaying minor loop test is shown in figure 11. These tests were performed to verify the algorithm’s ability to construct minor loops given a slightly more complex input sequence. In figure 11(a), after the three decaying minor loops are developed, the martensite fraction trajectory intersects the previous three extrema as temperature is decreased in the final monotonic input segment. After each intersection, the associated extrema pair is eliminated from memory as it approaches the
martensitic kinetic curve. After the third of the extrema pair is eliminated, the martensitic kinetic curve is traced exactly.

This test also sheds light on how minor loop behavior in any hysteretic system can impact control design by virtue of the system being non-stationary. As driving amplitude decays over time, the resulting minor loops converge to a small slope. In the worst-case, if the trajectory evolves to a steady-state that is in the flat region of one of the kinetic curves (i.e., either between \( M_s \) and \( A_f \) or \( M_f \) and \( A_s \)), then this slope may approach zero. With such a severe change in input-output sensitivity, it is obvious that control stability and robustness can be significantly compromised as steady-state is reached. Such concerns have been echoed by other researchers (ref. 16), who have focused on development of robust SMA control techniques. Having multiple minor loop fidelity, the present model can capture these important effects, and hence predict stability given any compensative action.

A similar set of tests is shown in figure 11(b), except the minor loops were generated by stress cycling. As with temperature variation tests, the input-output profile follows partial recovery paths up to the last segment, during which time the trajectory eliminates the previous minor loop extrema after they are intersected. At each stress relief segment, the trajectory follows linear elastic theory described by \( E_M \) until stress drops below \( \sigma_{cr}^s \), at which point strain partially recovers.

### Low Temperature Stress-Cycled Minor Loop Tests

A cycled stress input test was conducted to examine the stress-strain effects at temperatures below \( M_f \) and to validate the one-way hysteresis of the model. As shown in the room temperature results in figure 12, the strain follows an irreversible path during the stress relaxation phase of the cycle, returning to linear elasticity following the turn-around points. Hysteresis loops that appear in the data during the stress relief cycle are believed to be a result of partial recovery of \( \xi_S \) as stress reaches zero, i.e., a partial two-way effect. Nevertheless, it is encouraging that the stress-strain output upon stress reversal initially follows a slope very close to \( E_M \), as confirmed by moduli taken at the initial portions of the reversal curves in table III. It should be noted that, because the data was taken at room temperature, the average modulus more closely approximates modulus data taken on room temperature samples, a slightly larger value than the \( E_M \) at 200 °C used in the model.

<table>
<thead>
<tr>
<th>Cycle number</th>
<th>Modulus (decreasing), ksi</th>
<th>Modulus (increasing), ksi</th>
</tr>
</thead>
<tbody>
<tr>
<td>1</td>
<td>13437</td>
<td>13048</td>
</tr>
<tr>
<td>2</td>
<td>11034</td>
<td>12054</td>
</tr>
<tr>
<td>3</td>
<td>10616</td>
<td>12001</td>
</tr>
<tr>
<td>4</td>
<td>10267</td>
<td>11199</td>
</tr>
<tr>
<td>5</td>
<td>11202</td>
<td>10655</td>
</tr>
<tr>
<td>6</td>
<td>11894</td>
<td>10945</td>
</tr>
<tr>
<td>Average modulus</td>
<td></td>
<td>11529</td>
</tr>
<tr>
<td>Mart. modulus at room temperature</td>
<td></td>
<td>12000</td>
</tr>
</tbody>
</table>

In all of the phenomenological models, the two-way recovery critical stresses are in the material’s compression region. The Tanaka model (ref. 24), upon which the Liang & Rogers and Brinson models are based, uses exponential relationships that are capable of preserving the continuous nature of phase transformations. However the latter’s models have replaced the exponential with a piecewise continuous cosine relationship to avoid numerical convergence problems of the exponential. As shown in figure 13, the piecewise nature of the Liang & Rogers and Brinson models require that the material exhibit linear elasticity between \( \sigma_{cr}^s \) and \( \sigma_{cr}^{s'} \), therefore permitting exclusion of \( \sigma_{cr}^{s'} \) from the model. Without
substantial modification, these models are incapable of predicting these partial recovery loops due to this oversimplification of linear behavior. Although problems with the piecewise model has been acknowledged and addressed in the literature (as it also occurs with NiTi alloys) (refs. 7, 12, and 18), as yet there are no reported solutions to the issue.

6. Conclusions

In this paper, a simple, computationally efficient numerical model of a shape memory alloy material is developed. The model consists of a constitutive model with transformation kinetics and a novel minor loop hysteresis model. In the transformation kinetics relationships, Brinson’s formulation was modified from original form to remove discontinuities in the response. For the hysteresis model, an algorithmic representation of hysteresis is proposed as an alternative to generalized Preisach methods that require a
database of first order descending curves to properly identify the hysteresis distribution profile. As such, the present model is conducive to simplified identification methods for new shape memory alloy (SMA) materials such as the recently-developed high-temperature SMA. The hysteresis algorithm is implemented as a modular component so that any supplemental models (i.e., heat transfer, transformation kinetics, constitutive equations) may be independently modified or freely interchanged.

The parameterized model was empirically verified to obey two established Preisach properties: the minor loop congruency property and the wiping out property. These tests were sufficient to validate the behavior of the hysteresis model itself in comparison with existing quasi-static Preisach models. More complex loading profiles were also evaluated, with the outcome that the model is well-suited for controls analysis and evaluation. Low-temperature stress cycling was validated against a set of experimental data, demonstrating agreement with the model over the operating range of interest.

As a result of this thorough evaluation of the model, insight can be gained as to the future work that may be pursued:

1. modification of amplitude thresholding to prevent noise from producing excessive errors in scaling factor;
2. inclusion of partial recovery minor loops, not captured by the present model; and
3. thorough validation against an exhaustive data set of a “benchmark” high-temperature alloy.

The present model is well adept for prototype design and evaluation of high-temperature shape memory alloy (HTSMA) actuators. Pending experimental validation, the model may also be used for proportional control system design, as a result of the inclusion of path-dependent minor hysteresis loops. It is envisaged that the model can be applied to simulate the behavior of a broad range of passive or active actuation devices utilizing HTSMA or other research-stage materials.

References

# Development of a Numerical Model for High-Temperature Shape Memory Alloys

Jonathan A. DeCastro, Kevin J. Melcher, Ronald D. Noebe, and Darrell J. Gaydosh

A thermomechanical hysteresis model for a high-temperature shape memory alloy (HTSMA) actuator material is presented. The model is capable of predicting strain output of a tensile-loaded HTSMA when excited by arbitrary temperature-stress inputs for the purpose of actuator and controls design. Common quasi-static generalized Preisach hysteresis models available in the literature require large sets of experimental data for model identification at a particular operating point, and substantially more data for multiple operating points. The novel algorithm introduced here proposes an alternate approach to Preisach methods that is better suited for research-stage alloys, such as recently-developed HTSMAs, for which a complete database is not yet available. A detailed description of the minor loop hysteresis model is presented in this paper, as well as a methodology for determination of model parameters. The model is then qualitatively evaluated with respect to well-established Preisach properties and against a set of low-temperature cycled loading data using a modified form of the one-dimensional Brinson constitutive equation. The computationally efficient algorithm demonstrates adherence to Preisach properties and excellent agreement to the validation data set.