AI Models for Prediction of Displacement and Temperature in Shape Memory Alloy (SMA) Wire

Akshay Krishna Sheshadri[1,4 c)], Poonam Jangid[1, b)], Samarth Singh[2, a)], H N Bhargaw[3], B A Botre[1], S A Akbar[1] and SAR Hasmi[3]

[1]*CSIR-Central Electronics Engineering Research Institute, CEERI, Pilani, Rajasthan-333031, India*
[2]*CSIR-Central Electronics Engineering Research Institute, Naraina, New Delhi, Delhi-110028, India*
[3]*CSIR-Advanced Materials and Processes Research Institute, Bhopal, Madhya Pradesh-462026, India*
[4]*Birla Institute of Technology and Science, Pilani, Rajasthan-333031, India*

[a)] samarth@ceeri.res.in
[b)] poonam.jangid01@gmail.com
[c)] akshay.s.krishna97@gmail.com

Abstract. Shape Memory Alloys (SMAs) are a unique class of smart materials that have the ability to recover their shape on temperature stimuli. During this transformation, hysteresis and non-linear behavior can be observed and open-loop control design is inadequate for tracking control of these actuators. This is a major setback for the design and development of any SMA device. The main hurdle is that this nonlinearity can't be modeled effectively even by 3rd-degree differential equations. Additionally, the apparatus used for measurement of strain recovery during SMA actuation includes linear variable differential transducer (LVDT), which is very bulky and expensive and doesn't let us fully utilize the potential of SMA applications in miniaturized devices. This research work presents a method to eliminate the bulky position sensor by introducing an Artificial Neural Networks (ANN) to compensate for the non-linearity. Various researchers have attempted to model the behavior of SMAs using ANN techniques but these models had a high RMS error. In this paper, we develop a more complex neural network to model SMA's behavior. We model SMA's behavior with (i) displacement prediction (ii) temperature prediction. The results of our research not only demonstrate the effectiveness of Neural Networks for prediction of displacement and temperature of the SMA but also show how the proposed architectures have a much lesser error as compared to earlier models and are much more effective at modeling SMAs.

Shape Memory Alloys (SMAs) are a unique class of smart materials that have the ability to recover their shape on temperature stimuli. This Shape Memory Effect (SME) occurs due to temperature stimuli and stress-dependent shifts which causes a change in the material's crystalline structure between two different phases - martensite (low-temperature phase) and austenite (high-temperature phase). During the phase transformation, hysteresis and non-linear behavior can be observed which are a major setback for any device design and development. Therefore, most SMA applications are based on only two-states transformation - maintaining a total austenite phase by continuous heating or maintaining a total martensite phase by continuous cooling. Due to the huge advantages provided by SMA, the Control of SMA's phase transformation phenomenon is essential for its usage in many areas of aerospace, biomedical, automobile, vibration-dampers, etc. The main hurdle is that the nonlinearity can't be modeled effectively even by 3rd-degree differential equations. Additionally, the apparatus used for measurement of strain recovery during SMA actuation is includes linear variable differential transducer (LVDT), which is too very bulky and expensive and doesn't fully utilize the potential of SMA applications in miniaturized devices. This research work presents a method to eliminate the bulky position sensor i.e. LVDT by introducing an artificial neural network (ANN). This elimination will also reduce the total cost of an SMA system. Various researchers have attempted to model the behavior of SMAs using ANN techniques but these models still had a high RMS error. In this paper, we develop a complex ANN model to predict SMAs behavior and estimate strain recovery based on identified input parameters.

Current is used to produce the Joule heating effect to raise the temperature which triggers the phase transformation in the SMA wire. The ANN model tries to model the hysteresis (and non-linearity) between the Resistance difference and the displacement of the SMA wire by taking resistance difference (dR) as input and predicting the displacement of the SMA wire. We have also developed another Neural Network model to predict the temperature of the SMA wire with the current running through it and the resistance as the input. The results of our research not only demonstrate the effectiveness of the Neural Networks for prediction of displacement and temperature of the SMAs but also show how the proposed architectures have a much lesser error as compared to earlier models and are much more effective at modeling SMAs.

INTRODUCTION

Shape Memory Alloys (SMAs) are a unique class of smart materials that have the ability to return to their predetermined shape on temperature stimuli. When the alloy is below a particular 'transformation' temperature, it has a low yield strength and can be deformed easily into any desired shape which is the shape it retains. [1, Rao]

When the alloy is heated above its transformation temperature, this Shape Memory Effect (SME) occurs (due to temperature stimuli and stress-dependent shifts) which causes a change in the material's crystalline structure between two different phases - martensite (low-temperature phase) and austenite (high-temperature phase). [2, wiki]. The most common form of SMA is in the form of a wire (1-degree of freedom?). Due to its unique properties, it is used in many areas of aerospace, biomedical, automobile, vibration-dampers, etc. It offers a lot more advantages than piezoelectric actuators and motors.[?] They are able to generate large deformations and large restorative forces (if it encounters a resistance) at a much lower operating frequency.[?]

During this phase transformation, hysteresis and non-linear behavior can be observed which are major setbacks for the design and development of any SMA device. [? Maybe not required] Therefore, most SMA applications are based on only two-state transformation - maintaining a total austenite phase by continuous heating or maintaining a total martensite phase by continuous cooling. Due to the huge advantages provided by SMA, the control of SMA's phase transformation phenomenon is imperative.

In the pursuit of SMA wire application, it has been observed that a large number of factors come into play which has proved to be a hindrance. One of the main factors is the hysteresis which can't be modeled effectively even by 3rd-degree differential equations. Additionally, the apparatus used for measurement of strain recovery includes linear variable differential transducer (LVDT), which is very bulky and expensive. Many models were proposed to attempt to address this hysteresis problem including - Webb et al, Cruz-Hernandez, and Hayward, Song and Quinn, etc. One of the best approaches is to use Artificial Neural Networks (ANN) to solve this issue. Neural nets possess nonlinear function mapping and adaptation properties. This makes it a viable candidate to be used as it can be used as a feedforward controller in the actuation. Various researchers [??] have attempted to model the behavior of SMAs using neural networks but these models still had a high RMS error. This paper aims to implement an open-loop control of the SMA wire displacement device without the use of position sensors by using more complex neural networks for hysteresis compensation. This will not only make the entire setup less bulky, but it also reduces the high cost of the actuator.

For data collection, a SMA test rig is set. Two SMA wires are connected in antagonist way for bidirectional movement. We use current to produce the Joule heating effect to raise the temperature which triggers the phase transformation in the SMA wire. A heating module developed at CSIR-CEERI is used to provide constant current to both SMA in complementary form. We observe a hysteresis relationship between (i) current, input voltage, resistance v/s displacement and (ii) current, resistance v/s temperature. The first ANN architecture tries to model the hysteresis between the difference?? Resistance and the displacement of the SMA wire ~~by taking resistance difference (dR) as input and predicting the displacement of the SMA wire~~. The second ANN architecture to predict the temperature of the SMA wire with the current running through it and the resistance as the input. (POWER?)

It is also important to note that the hysteresis behavior isn't completely repeatable in nature. In every cycle, the voltage v/s displacement follows a slightly different curve. There can be as much as a 25% error in the value of the displacement (though the behavior becomes more stable after a few cycles and the voltage v/s displacement curves start to overlap). The curve isn't completely smooth either. This has been attributed to ambivalent conditions and environmental effects that can neither be controlled nor replicated. Hence, we should use a low pass filter and use the smoothed data for our model. We should also note that the current through the wire isn't sinusoidal like the applied voltage due to the change in resistance during phase transformation between Martensite and Austenite phases.

The models only use one representative hysteresis loop. To remove chattering present due to environmental noise, the displacement data was filtered using a third order low-pass filter.

The most commonly used SMA is an alloy of Nickel and Titanium called NiTi or Nitinol. It has very good electrical and mechanical properties, long fatigue life and high corrosion resistance. It can recover from up to 5% strain (500 MPa restoration stress). It has high resistance properties make it suitable to be actuated electrically with Joule heating (using current instead of directly using temperature for the actuation of the wire). Hence, it is a good approach to understand the relationship between the resistance difference and the displacement of the wire. Another experiment is on predicting the temperature of the SMA wire given the electrical power applied across it and finding the hysteresis relationship.

1. NEURAL NETWORK MODEL

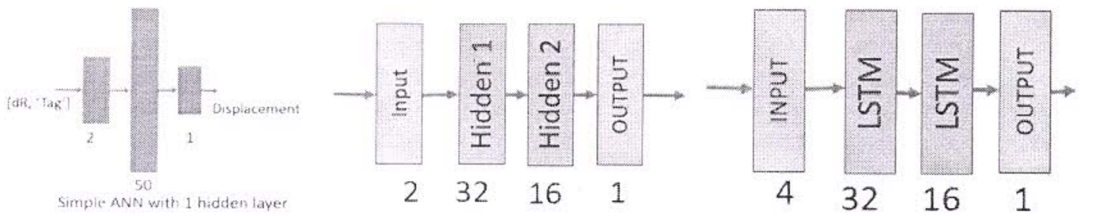
Displacement Prediction

The first architecture was a simple feedforward neural network. This took the resistance difference and a 'tag' value as input and tried to predict the displacement. The 'tag' is an input which takes a value of +2 for increasing voltage and -2 for decreasing voltage. It takes a value of +2 when there is no change in voltage to ensure continuity of signals. This input is to help the model know which part of the hysteresis is being followed. Essentially, the inputs are the applied voltage and information regarding the direction of the voltage (increasing or decreasing).

Our First model had an input layer, one hidden layer with 50 neurons and an output layer (the performance increased drastically as the number of neurons increased from 10 to around 50, after 50, the performance improves slightly). The Second model had an input layer, two hidden layers with 32 and 16 neurons and an output layer. Both these models take resistance difference and 'tag' value as input and outputs the predicted displacement.

The 'tag' value we used for the previous architecture was determined based on the previous value of 'voltage'. This value is essentially derived from the previous input or 'history' of the SMA. This motivated us to choose sequence models as our next architecture. This architecture takes the previous computed value (hidden state) and propagates this to the next time step. This means that we can omit the 'tag' value and give the previous input(s) along with the current input for the model to make a prediction. We chose LSTM for this purpose.

·Our LSTM model took the current resistance difference and the previous 3 resistance differences as input to make a prediction. This no longer uses the 'tag' value. The LSTM layer has 32 neurons. It has output is connected to a hidden layer with 16 neurons which is connected to the output layer.

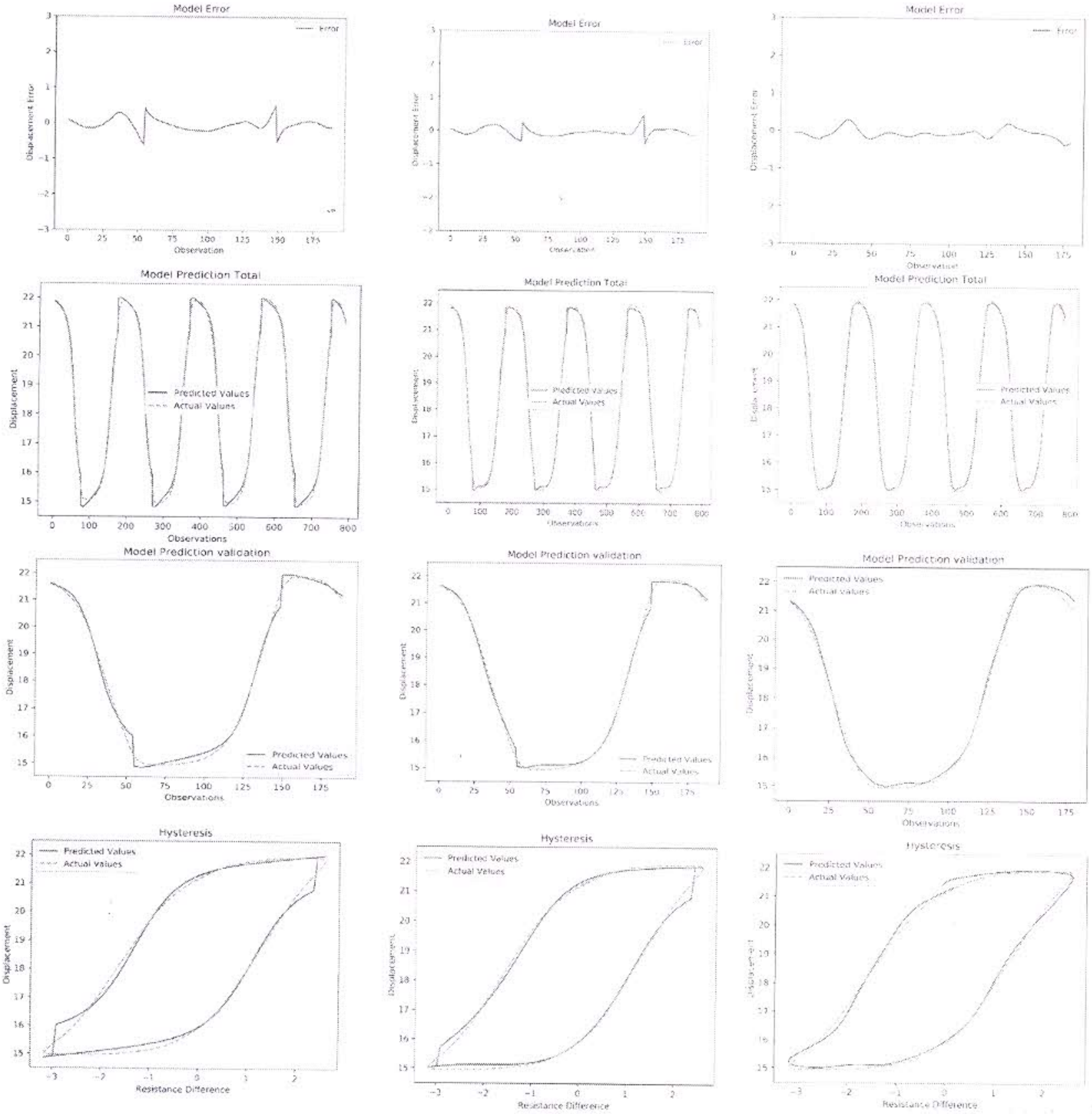


The results of both architectures are in the table 1. We can see that while the mean-squared loss and the maximum error (in mm) is comparable, the behavior that the architectures have learnt are very different. The simple feedforward model's graphs break the graph into multiple straight lines and learn that behavior. It models the extremes of the hysteresis curve very poorly (where the displacement changes for a given delta resistance change is very different for the upper and lower hysteresis curve). The cyclically large errors observed near the saturation ends of the hysteresis loops are caused due to truncation errors. In comparison to that, the LSTM has a very smooth curve and models the behavior of SMA well, even at the extreme ends.

TABLE 1. Results of both architecture for the displacement prediction

model	total loss	validation loss	max_error
LSTM Disp.hdf5	0.015508362158942	0.018811862501833	0.343526840209961
simple 2 layer	0.015808121623479	0.015939184003781	0.514333724975586
simple 1 layer disp.hdf5	0.029591650577529	0.029610275761495	0.604925155639648

Simple_1_layer Simple_2_layer LSTM



Temperature Prediction

Similar to the displacement prediction, we use similar architectures for temperature prediction.

The first architecture was a simple feedforward neural network. This took the resistance, current and a 'tag' value as input and tried to predict the temperature. Similar to displacement prediction, the 'tag' is an input which takes a value of +2 for increasing resistance and -2 for decreasing resistance. It takes a value of +2 when there is no change in resistance to ensure continuity of signals. This input is to help the model know which part of the hysteresis is being followed.

We had two models for this purpose. Both models have 1 input layer, 2 hidden layers with 32 and 16 neurons and an output layer. The difference between them is that one of them uses ambient temperature also as an input while the other one does not.

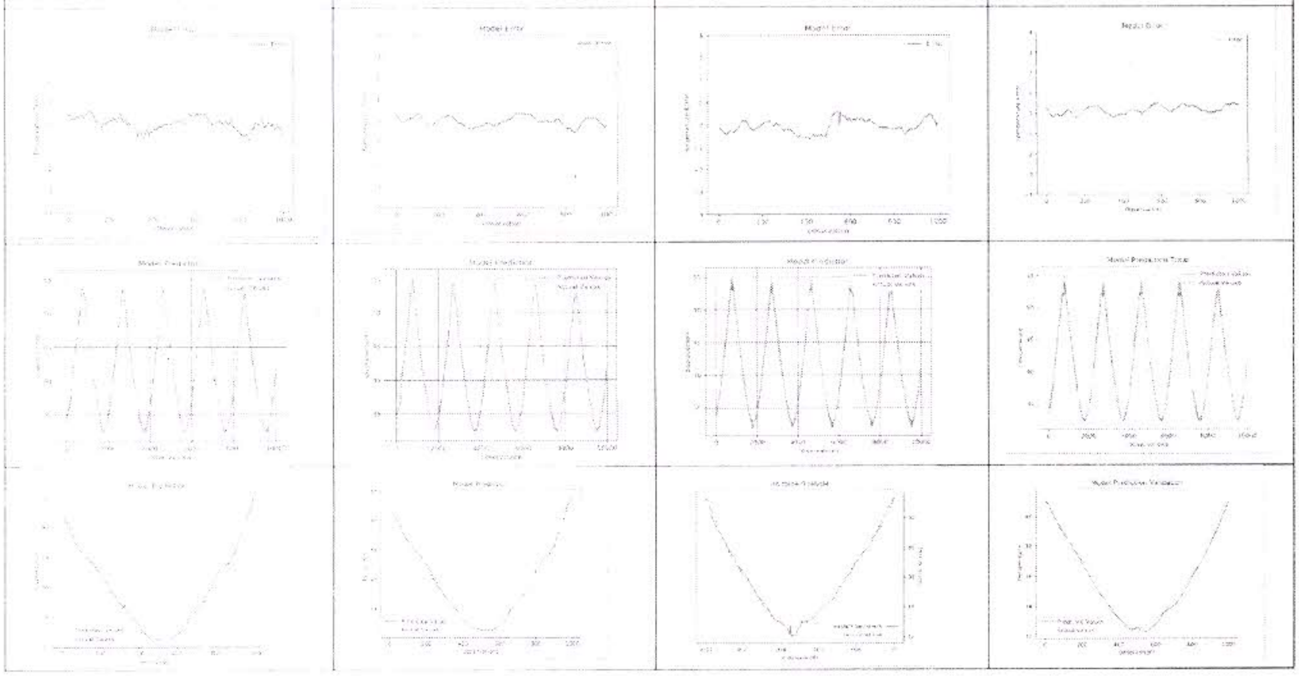
The second architecture is LSTM. This takes resistance and current as input and doesn't need the 'tag' value. Similar to the displacement prediction, it uses current input and previous 3 inputs as the input for making a prediction. The LSTM layer has 32 neurons. It's output is connected to a hidden layer with 16 neurons which is connected to the output layer.

The results of both architectures are in the table. We can see that while the mean-squared loss and the maximum error (in degree celsius) is comparable, the behavior that the architectures have learnt are very different. The simple feedforward model's graphs break the graph into multiple straight lines and learns that behavior. It models the extremes of the hysteresis curve very poorly (where the displacement change for a given delta resistance change is very different for the upper and lower hysteresis curve). The cyclically large errors observed near the saturation ends of the hysteresis loops are caused due to truncation errors. In comparison to that, the LSTM has a very smooth curve and models the behavior of SMA well, even at the extreme ends.

Table 2 **Results of both architecture for the temperature prediction**

model	*total_loss*	*validation_loss*	*max_error*
LSTM with ta	0.298160426028329	0.08660652843981	0.755020141601563
LSTM without ta	0.271552877681868	0.044232553626639	0.547409057617188
model with ta	0.327874886824191	0.085918510504067	0.633716583251953
model without ta	0.377871336997487	0.06007460474968	0.544509887695313

LSTM with ta	LSTM without ta	Simple with ta	Simple without ta



REFERENCES

1. Naresh, C., P. S. C. Bose, and C. S. P. Rao. "Shape memory alloys: a state of art review." *IOP Conference Series: Materials Science and Engineering*. Vol. 149. No. 1. IOP Publishing, 2016.
2. https://en.wikipedia.org/wiki/Shape-memory_alloy
3. G. SONG. "A Neural Network Inverse Model for a Shape Memory Alloy Wire Actuator", Journal of Intelligent Material Systems and Structures, 2003