POLITECNICO
MILANO 1863

SCUOLA DI INGEGNERIA INDUSTRIALE
E DELL'INFORMAZIONE

EXECUTIVE SUMMARY OF THE THESIS

Modeling and state estimation for the advanced control of a multimaterial shape memory based actuator

TESI MAGISTRALE IN AUTOMATION AND CONTROL ENGINEERING – INGEGNERIA DELL'AUTOMAZIONE

AUTHOR: DIEGO MAGRIN

ADVISOR: MARIA PRANDINI

CO-ADVISORS: FABIO LAZZARI, SIMONE PITTACCIO

ACADEMIC YEAR: 2020-2021

---

1 Introduction

In this thesis we model a multimaterial SMA (Shape Memory Alloy) based actuator made by a SMA wire coated by a silicone layer with a lumped parameters model.
We also design a nonlinear state observer which will be able to estimate the unmeasured variables of the system as well as some unknown parameters.
Finally, we show some preliminary results of a controller which exploits the nonlinear state observer for tuning its parameters based on the system operating zone.

2 Background on SMA

SMAs are a class of smart materials which are more and more adopted in a variety of domains, ranging from Biomedical to Aerospace Engineering, due to their peculiar properties, like the Shape Memory Effect (SME): SMAs can recover pseudo-plastic deformations obtained with the application of a load by being heated above a certain temperature, regaining their original shape, as schematized in Figure 1.

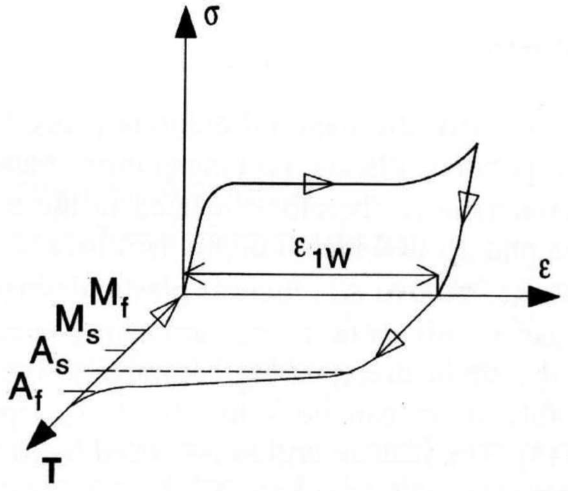
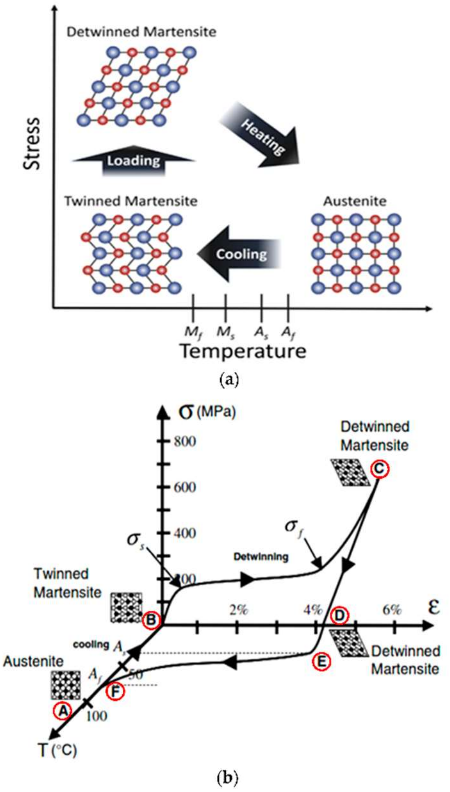


Figure 1: Shape Memory Effect

In NiTi based SMAs the SME is caused by a solid-to-solid phase change between Austenite, the parent phase with B12 (CsC1) lattice structure, which is stable at high temperatures, and Martensite, with monoclinic lattice structure B19’, stable at low temperatures, which can also be found in two configurations:

- Twinned Martensite, in which the orientation of the atoms is mirrored with respect to an atomic plane and is obtained by cooling below a certain critical temperature.
- Detwinned Martensite, formed under the effect of an externally applied stress and where the atomic layers tend to orient in the same direction of the stress.

The martensitic transformation is schematized in Figure 2.



(a)

(b)

Figure 2: Martensitic transformation in the σ-T plane σ-T- ε space

A suggested reference for the detained description of martensitic transformation in SMAs and their composition is found in (Otsuka & Ren, 2005).

3 Considered actuator

We consider a linear actuator composed by a NiTiCu wire 10 cm long and with a diameter of 0.65 mm surrounded by a silicon rubber coating 1mm thick in direct contact with the wire, operating while being surrounded by air.

The wire is actuated through a voltage source, with maximum voltage V_{SAT} = 1 V, and to maintain a good fatigue life is intended to be used for applications of 1000 cycles, with loads of approximately 200 MPa and maximum strain 3%.

4 SMA model

We develop a model of the SMA thermomechanical behavior based on the model from (Brinson, 1993), which was selected since it is already widely used, with plenty of examples from the literature, and appropriate for developing an observer for control application, which requires a relatively simple model and a good macroscopic description of the austenitic and martensitic phase changes.

4.1 Constitutive model equations

The constitutive equation 1 relates the state variables stress σ, strain ε and temperature T in the terms of the Martensite volume fraction ξ.

$$\sigma - \sigma_i = [D(\xi)(\varepsilon - \varepsilon_i) + \Omega(\xi)(\xi - \xi_i) + \Theta(T - T_i)] \qquad 1$$

$$D(\xi) = D_A + \xi(D_M - D_A) \qquad 2$$

$$\Omega(\xi) = -\varepsilon_L D(\xi) \qquad 3$$

where σ is the stress applied on the wire, ε is the wire strain, ξ is the total martensitic fraction, ξ_S is the stress induced Martensite fraction, D is the total elastic modulus, which is a linear combination of the elastic moduli of Austenite and Martensite, respectively D_A and D_M, proportionally to the total martensitic fraction ξ; Ω is the stress/transformation coefficient and ε_L is the maximum residual strain and finally Θ is the thermoelasticity modulus. The subscript *i* indicates the initial value of the variable at the start of the transformation.

4.2 Transformation kinetics

The SMA wire undergoes different phase transformations, depending on the position of its inner state in the σ-T plane shown in Figure 3.

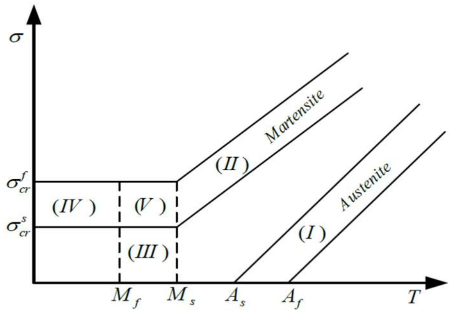


Figure 3: SMA phase plane with critical stresses

The laws governing the martensitic fraction change depending on the current zone of the σ-T plane in which the state finds itself.

There are a total of 5 zones, plus a neutral one where no transformation happens.

Zone 5 is also subdivided in 3 subzones, depending on the combination of the temperature and stress derivatives.

There are two equations for each zone, one for the Twinned Martensite ξ_T and one for the Detwinned Martensite ξ_S, these relations depend on the wire temperature, the applied stress and the martensitic fraction at the beginning of the transformation.

4.3 Thermal model

For our goal, we are not interested in the temperature distribution inside the silicon but only in the overall contribution of its presence in the heat transfer from the wire to the environment, thus we decided to use a lumped parameters approach. The two main state variables of the thermal model are T_1, the wire temperature (isothermal), and T_2, the average silicone temperature.

$$m_1 c_1 \dot{T}_1 = \frac{V^2}{R_{el}(\xi)} - \frac{T_1 - T_2}{R_{t1}} + m_1 L_h \dot{\xi} \qquad 4$$

$$m_2 c_2 \dot{T}_2 = \frac{T_1 - T_2}{R_{t1}} - \frac{T_2 - T_\infty}{R_{t2}} \qquad 5$$

where m_1 and m_2 are the masses of the wire and insulation, c_1 and c_2 are their respective specific heat capacities.

R_{el} is the electric resistance of the wire, which depends on the strain and martensitic fraction.

$$R_{el}(\xi) = \frac{L_0(1+\varepsilon(\xi))}{A_0(1-\nu_{NiTi}\varepsilon(\xi))^2}[\rho_A + \xi(\rho_M - \rho_A)] \qquad 6$$

where ρ_A and ρ_M are the Austenite and Martensite electrical resistivities, L_0 is the initial wire length and ν_{NiTi} is the SMA wire Poisson, which is equal to 0.33, and A_0 is the initial wire area.

r_{t1} and r_{t2} are the thermal resistivities per unit length of SMA wire and silicone respectively, R_{t1} and R_{t2} are their thermal resistances.

$$r_{t1} = \frac{ln\frac{a_2}{a_1}}{2\pi k_{Si}} \qquad 7$$

$$r_{t2} = \frac{1}{2\pi a_2 h} \qquad 8$$

$$R_{t1} = \frac{r_{t1}}{L_0} \qquad 9$$

$$R_{t2} = \frac{r_{t2}}{L_0} \qquad 10$$

5 Extended Kalman Filter

Because of the system nonlinearities like hysteresis and branching, the estimation is done through the use of an EKF (rather than a standard Kalman Filter), to which we feed the current control input of the system, the last available strain measurement and the last available current zone of our controller.

The state variables of the filter equations are:

- T_1 SMA wire average temperature;
- T_2 Silicon average temperature;
- ξ_S Stress induced Martensite fraction;
- ξ_T Thermal induced Martensite fraction;
- ξ Total Martensite fraction;
- ε Wire strain;
- R_{el} SMA wire electrical resistance.

The unknown parameters and disturbances to be estimated are:

- σ Wire stress;
- h Convection coefficient;

The control variable is the voltage V, the output of the system is the wire strain ε.

The system equations needed to be discretized before they can be used by the EKF.

6 Results

In this chapter we show the major results we obtained.

6.1 Model parameter optimization

We compared the stress recovery curves obtained from a real wire to the performance of our model. We executed many transformation cycles, tuning the model parameters until the simulations matched the measurements.

In Figure 4 we see the comparison between our simulated curves, the dotted lines, and the real wire behavior, the continuous lines.

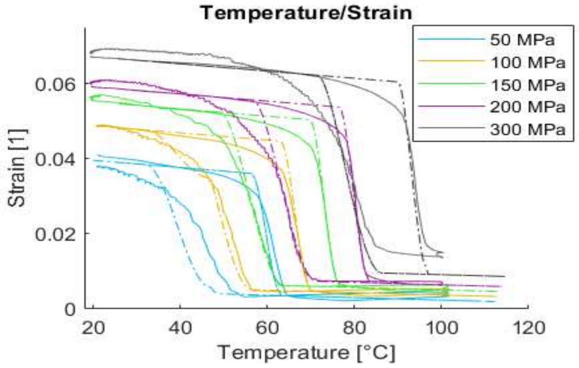


Figure 4: Comparison between real and simulated strain recovery curves

After the parameter optimization process the behavior of our system is very similar to the one of the real wire.

6.2 Silicone effect on temperature

We conducted a test to verify the effect of the presence of the silicone layer on the wire temperature, seen in Figure 5.

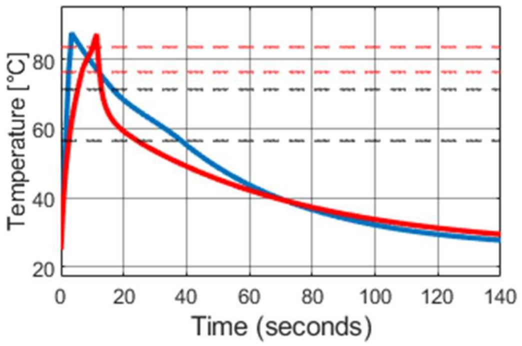


Figure 5: Comparison between wire temperature with (blue) and without (red) the silicone

We see that the presence of the silicone rubber slows down the wire heating but greatly helps the initial part of the cooling phase, this is because the cooling is helped by the lower silicone temperature T_2.

6.3 EKF performance

Here we present the performance of our EKF.

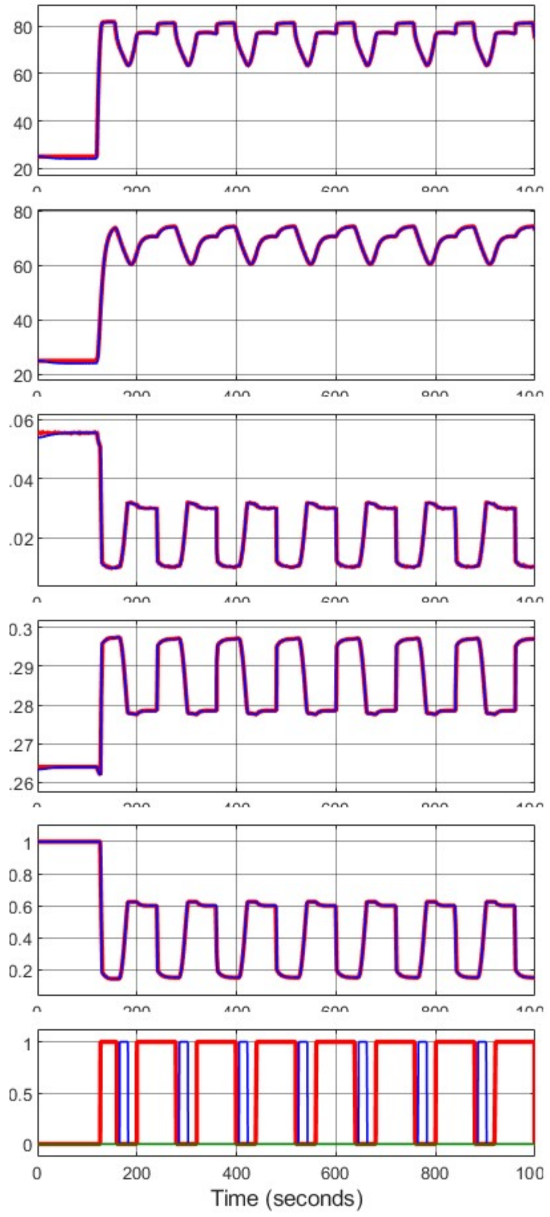


Figure 6: State variables estimated (blue) and simulated (red) with $h_{0f} = h_0 - 10$ and $\sigma_{0f} = \sigma_0 - 20$

In Figure 6 we see the estimated variables, respectively T_1, T_2, ξ_S, ε, R_{el} and transformation zones (zone 1 in red, zone 2 in blue) of our model controlled by a PID following a square wave reference starting with wrong initial conditions on the states.

We can see that the EKF is able to correctly estimate the state variables of the system.

In Figure 7 we see the unknown parameters estimation in the same conditions.

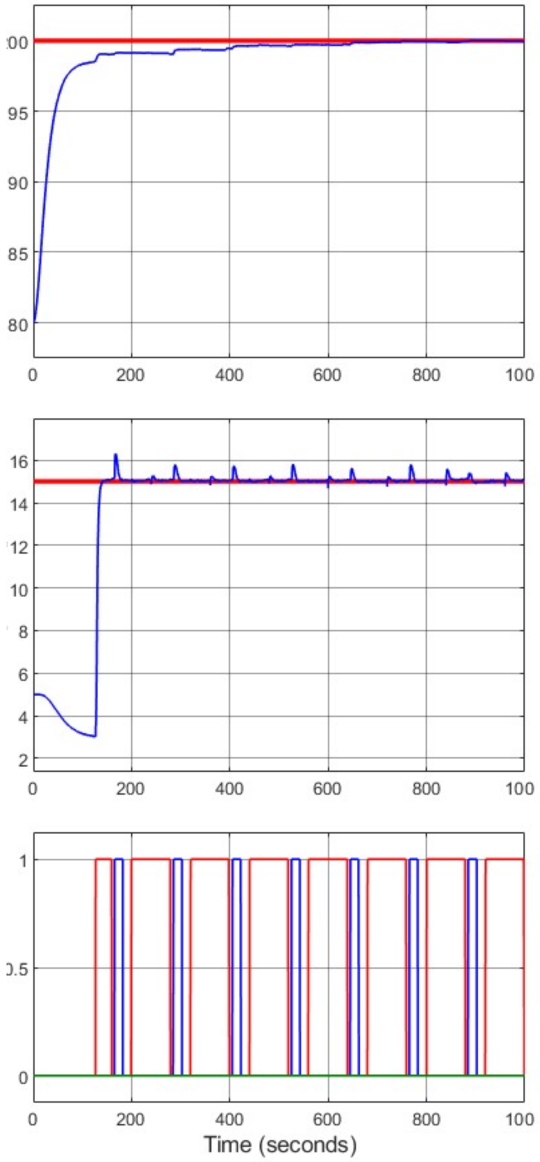


Figure 7: Unknown parameter estimated (blue) and simulated (red) with $h_{0f} = h_0 - 10$ and $\sigma_{0f} = \sigma_0 - 20$

We can see that after the initial parameter conditioning time of 120 s the EKF corrects the estimate of the unknown parameters and after the initial transitory the estimated σ has mean value 199. 63 MPa, standard deviation 0.323 MPa and its estimation error has maximum absolute value 0.05 MPa; while the estimated h has mean value 15.09 $W/K/m^2$, standard deviation 0.157 $W/K/m^2$ and maximum error absolute value 1.30 $W/K/m^2$.

6.4 Preliminary controller results

In this chapter we provide some initial results and implementation of a state variable dependent controller.

The preliminary nature of the current implementation depends on the fact that the rules for PID gain modulation were established empirically based on our observations of the controlled system behavior in the different zones.

In Figure 8 we see the controller performance in blue compared to a reference PID in red on the variables ε, T_1, T_2, ξ_S and electrical power P_{el}.

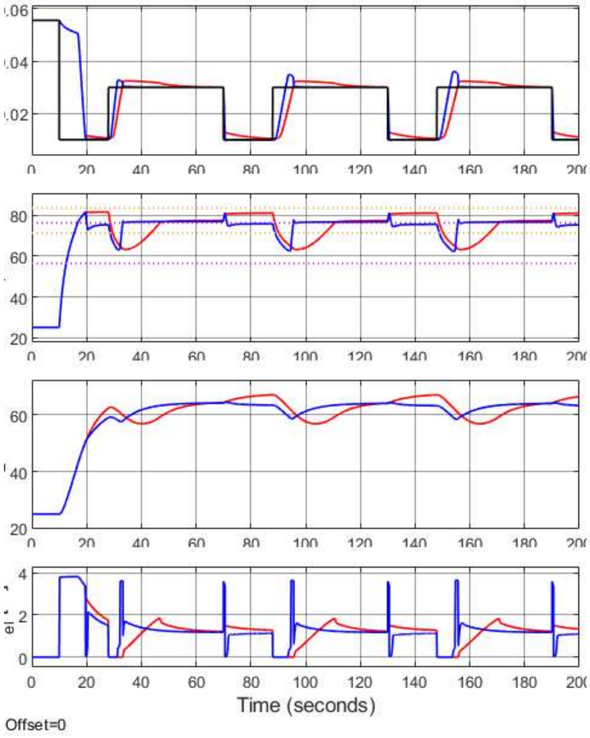


Figure 8: Performance of the proposed controller with square reference

During the first heating phase, the two controllers have a similar performance, while on the following heating zones our proposed controller is slightly faster than the reference PID and manages to converge more quickly to the reference and most importantly manages to keep the wire and silicone cooler; this is crucial during the cooling phase, where our variable PID controller can start cooling faster than the reference one.

7 Conclusions

In this thesis we implemented the Brinson model for the thermomechanical behavior of a SMA and a lumped parameter dynamic model of the heat exchange with the silicone, obtaining a Simulink model which allowed us to simulate the behavior of the real wire actuator, allowing us to conduct meaningful experiments without a physical setup.

To model the SMA hysteretic and conditional behavior in Simulink, we introduced a zone selection block which changes the model equations based on the current wire temperature, stress and their derivatives.

We also produced a state observer for the model in the form of an EKF, which is able to estimate all the unmeasured state variables and at the same time estimate the unknown parameters wire stress σ and convective heat coefficient h, despite the hysteresis and the conditional behavior of the model.

In parallel with the EKF works another zone selection block, based on the estimated values of T_1 and σ.

Finally, we showed how this added knowledge of the system could be used for improving a model-based controller, especially reducing the requested electrical power for the actuation, which is one of the biggest limiting factors for SMA based actuators.

Future research is needed to investigate model-based methods exploiting the proposed EKF.

8 Bibliography

Auricchio, F., Marfia, S., & Sacco, E. (2003). Modelling of SMA materials: Training and two way memory effects. *Computers and Structures, 81*, 2301-2317.

Brinson, L. (1993). One-Dimensional Constitutive Behavior of Shape Memory Alloys: Thermomechanical Derivation with Non-Constant Material Functions and Redefined Martensite Internal Variable. *JOURNAL OF INTELLIGENT MATERIAL SYSTEMS AND STRUCTURES, 4*, 229-242.

Otsuka, K., & Ren, X. (2005). Physical metallurgy of Ti–Ni-based shape memory alloys. *Progress in Materials Science, 50*, 511-678.

Scarpa, F., & De Rosa, M. (2016, August). Transient heat conduction in wires with heat sources; lumped and distributed solution

9 References

10 Acknowledgements

I want to thank Fabio and Simone for their great help, patience and for the opportunity to work at CNR ICMATE.

I want to thank prof. Prandini for her cordiality and her availability despite her many other commitments.

I want to thank my parents, my girlfriend, my family and my friends for all the support and love throughout this period.

POLITECNICO
MILANO 1863

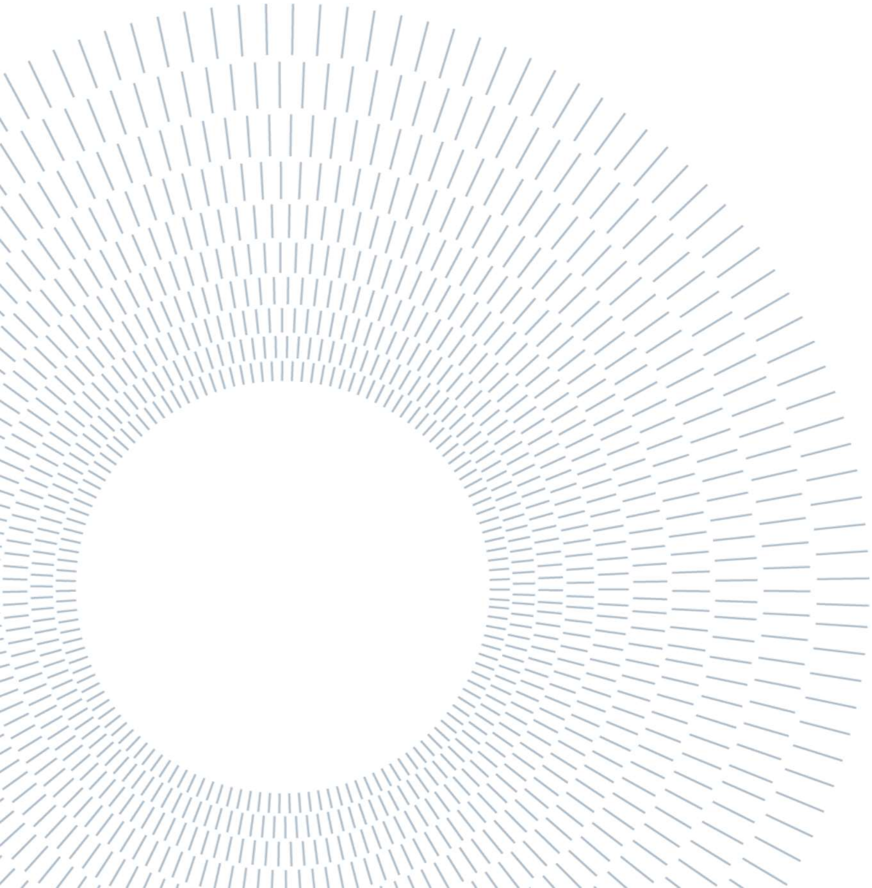
SCUOLA DI INGEGNERIA INDUSTRIALE
E DELL'INFORMAZIONE

Modeling and state estimation for the advanced control of a multimaterial shape memory based actuator

TESI DI LAUREA MAGISTRALE IN
AUTOMATION AND CONTROL ENGINEERING
INGEGNERIA DELL'AUTOMAZIONE

Author: **Diego Magrin**

Student ID:	920553
Advisor:	Maria Prandini
Co-advisors:	Fabio Lazzari, Simone Pittaccio
Academic Year:	2020-21



Abstract

Shape Memory Alloys (SMAs) are a class of smart materials with interesting thermomechanical properties, such as the ability to recover very large deformations when heated. A peculiar application of SMA actuators is soft robotics, where a SMA wire can be embedded inside a soft material, such as a polymeric matrix.

In this thesis, we expand the SMA model proposed in the literature by Brinson so as to develop a new thermomechanical multimaterial model of a SMA-based actuator made by a SMA wire coated by a silicone rubber layer. The system is highly nonlinear, time-varying and hysteretic, adding to the complexity of controlling actuation.

A new switched controller concept is proposed in the form of a Variable PID, whose parameters are tuned based on current SMA operating zone, which is determined via a state estimator and a logical block.

Key to the switched controller implementation is the development of an Extended Kalman Filter (EKF) for state estimation The EKF was tested and showed robustness in predicting the state.

Preliminary simulations of the Variable PID show that the innovative strategy allows for faster heating and cooling in the reaching some reference position compared to a conventional PID. Future work should complete the controller implementation, test its ability to improve energy efficiency, and provide better evidence of its functioning.

Abstract italiano

Le leghe a memoria di forma (SMA) sono una classe di materiali intelligenti con interessanti proprietà termomeccaniche, come la capacità di recuperare deformazioni molto grandi quando riscaldate. Un'applicazione peculiare degli attuatori SMA è la robotica morbida, in cui un filo SMA può essere incorporato all'interno di un materiale morbido, come una matrice polimerica.

In questo lavoro di tesi, abbiamo ampliato il modello SMA di Brinson proposto in letteratura per sviluppare un nuovo modello multimateriale termomeccanico di un attuatore basato su SMA realizzato da un filo SMA rivestito da uno strato di gomma siliconica. Il sistema è altamente non lineare, tempo variante e isteretico, aggiungendo complessità al controllo dell'attuazione.

Un nuovo concetto di controllore a commutazione viene proposto sotto forma di un controllore PID variabile, i cui parametri vengono modificati in relazione alla zona di lavoro dello SMA, che viene determinata utilizzando uno stimatore dello stato ed un blocco logico.

Determinante è quindi l’introduzione di un filtro di Kalman esteso per la stima dello stato. Il filtro introdotto è stato testato e ha mostrato di essere robusto nel prevedere lo stato del sistema e nel rigettare gli errori.

Simulazioni preliminari delle prestazioni del PID variabile hanno mostrato che la strategia innovativa proposta consente un riscaldamento e un raffreddamento più rapidi nel raggiungere la posizione di riferimento rispetto ad un PID convenzionale.

Sviluppi futuri riguardano la messa a punto del controllore proposto, testandone la capacità di migliorare l'efficienza energetica e fornendo una più evidente dimostrazione delle sue prestazioni.

Index

0 Introduction and thesis objective

0.1 Context

Shape Memory Alloys (SMAs) are a class of smart materials which have interesting thermomechanical properties, such as the ability to recover very large deformations when heated.

These properties can be employed through SMA based actuators, which are gaining importance in the industrial, aerospace and biomedical fields due to their small size and high force-to-mass ratio.

A peculiar application of SMA actuators is soft robotics, the branch of robotic concerned with highly compliant robotic actuators that mimic the behavior of living organisms, where a SMA wire can be embedded inside a soft material, such as a polymeric matrix, this opens many interesting applications but introduces challenges when modeling the heat exchange between the SMA and the matrix around it.

0.2 Thesis objective

The goals of this thesis are to model a multimaterial actuator made by a Shape Memory Alloy (SMA) wire coated by a silicone layer and to design an observer to estimate its state for possible application to advance controller design. The plant will include the wire/silicone system, a known source of loading (e.g., a hanging mass, an applied stress, a loading spring) and the surrounding atmosphere.

It is envisaged that, under appropriate conditions, this setup may present some opportunities regarding the actuation speed, because the silicone layer could help the wire cooling, and the applicability of the actuator for biomedical applications, as it could be in direct contact with the skin without letting the high temperatures reached by the wire burn it.

A controller will be proposed, able to exploit information obtained through the designed state observer, with the aim to control efficiently and precisely the position and movement stroke (0-3% strains) dynamically.

It is expected that an active estimation of the unmeasured state variables of the system will be able to improve the control action.

The factors expected to influence the actuation are the silicone temperature, the wire temperature, the convective heat transfer coefficient, and the load.

Out of these, the wire and silicone temperatures will be treated as dynamic variables and their evolution in time will be modeled, while the convective heat transfer coefficient and the load will be treated as unknown parameters which can have big fluctuations, the first because of environmental factors such as sudden gusts of air, the latter because it could vary accidentally or be measured poorly.

1 Background on Shape memory alloys

Shape Memory Alloys (SMAs) are a class of smart materials which are finding more and more uses, from Biomedical to Aerospace Engineering, due to their peculiar properties, like:

- Pseudo-Elasticity (PE): SMAs have the ability to isothermally recover big deformations, up to 6-8%, when the applied load is removed.
- Shape Memory Effect (SME): SMAs can recover pseudo-plastic deformations obtained with the application of a load by being heated above a certain temperature and regaining their original shape.
- Two-Way Shape Memory Effect (TWSME): SMAs have in addition the ability to recover a pre-determinate shape in response to cooling; to achieve this effect the material must have received previous thermomechanical training treatment.

The best-known SMAs belong to the quasi-stoichiometric NiTi intermetallic system, with possible ternary additions made to optimize specific properties. The alloys with this range of compositions proved excellent for the design of actuators and functional elements.

A suggested reference for the detained description of martensitic transformation in NiTi-based alloys and their composition is found in (Otsuka & Ren, 2005).

1.1 Martensitic transformation

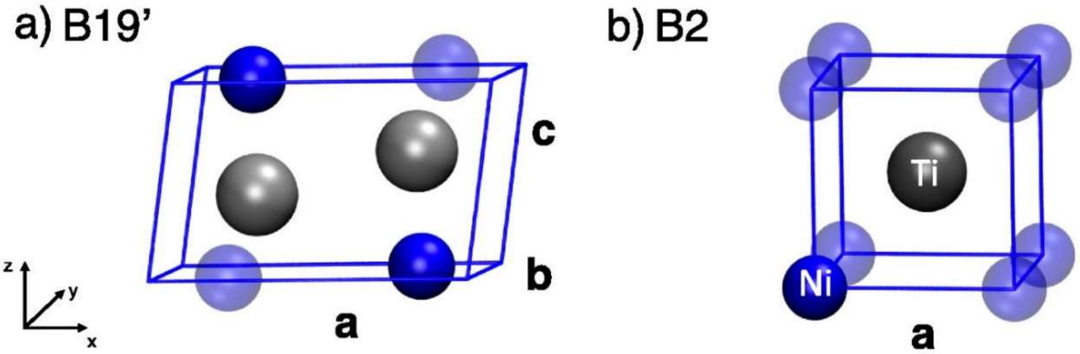


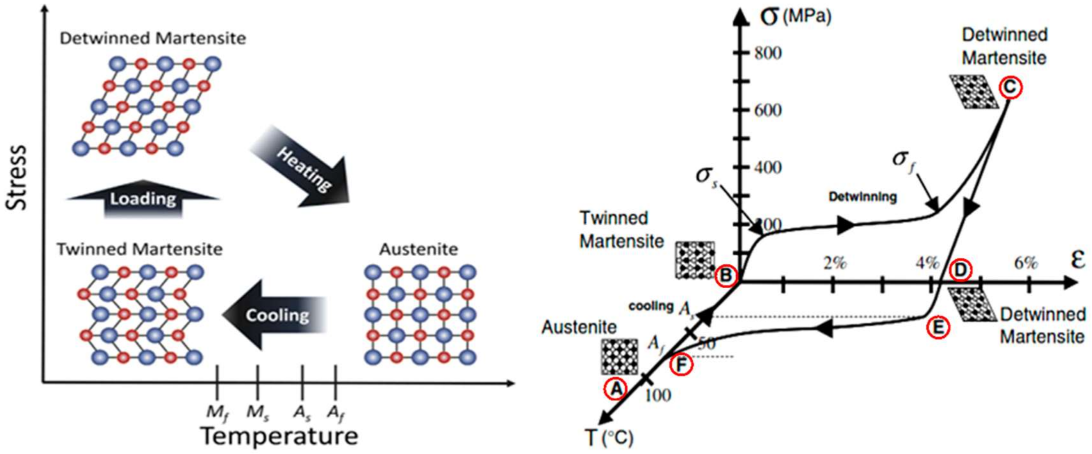
Figure 1 - Martensite crystal cell B19' (a) and Austenite cell B12 (b)

SMAs have two stable solid phases, which can be seen in Figure 1. In particular, for the SMAs based on the NiTi system, the phases are:

- Austenite, the parent phase, which is stable at high temperatures, has a B12 (CsC1) lattice structure and has a higher Young's modulus than Martensite.
- Martensite, stable at low temperatures, which has a monoclinic lattice structure B19', lower stiffness and can be found in 24 different variants. More precisely, Martensite can be found in two configurations:
    - Twinned Martensite, in which the orientation of the atoms is mirrored with respect to an atomic plane. It is obtained by cooling the SMA below a certain critical

temperature under quasi-null stress. Multiple twin variants accommodate elastically with each other most of the times.

- Detwinned or deformed Martensite, that is formed by one or few well-oriented variants grown to the expense of others under the effect of an externally applied stress. The atomic layers tend to orient towards the same direction of loading, producing a net macroscopic shape change. This can only be reversed if the temperature is raised above the critical temperature to stabilize Austenite.



(a) (b)

Figure 2 - SMA phase transitions

In NiTi-based alloys, the transition between Austenite and Martensite is a self-accommodating solid to solid transformation of the first order, because heat (transformation enthalpy) is released during the direct transformation (towards Martensite) and absorbed through the inverse one (towards Austenite). There are two simultaneous phenomena in the martensitic transformation:

- Lattice deformation (Bain distortion), in which the atoms rearrange into the new martensitic crystal structure, one atomic layer at the time, through small atomic movements, as seen in Figure 3, this new configuration has a different shape and volume with respect to the starting one.

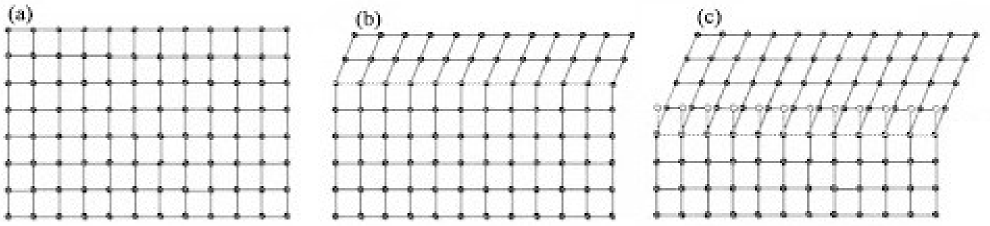


Figure 3 - Lattice deformation from Austenite to Martensite

- Plane-invariant shear, in which the new structure accommodates to the parent lattice configuration through two mechanisms that guarantee the invariance of the habit (interface) plane: a lattice rotation and an inhomogeneous lattice-invariant deformation, which may be slip or twinning, as displayed in Figure 4:

- Crystal plane slip, in which the atomic planes slip on each other, breaking some existing bonds. This process is plastic and irreversible, so in order to have the SME it must happen rarely or not at all.
- Twinning, in which two differently oriented domain accommodate without breaking bonds. In SMAs this process is reversible because it generates elastic stresses at the border between monovariant domains, which allow the recovery of the original shape when the material is heated.

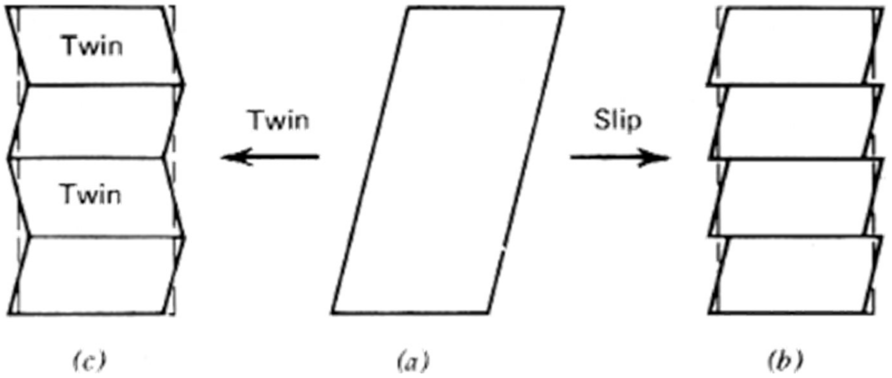


Figure 4 - Twinning or crystal plane slip can enforce invariance of the habit plane (dashed)

The main method of accommodation in NiTi is by twinning, and the transformation is highly reversible. Because of the friction generated and the energy dissipated in the direct and inverse transformations, it is observed that the martensitic transformation is a with hysteresis. The material will transition to austenite between the temperatures A_S and A_F, which are known respectively as Austenite Start and Austenite Finish temperatures, when heating and similarly it will transition to martensite between M_S and M_F, which are similarly known as Martensite Start and Finish temperature, when cooling. These temperatures are different from one another, as shown in Figure 5 and according to (Tanaka, 1986), (Liang & Rogers, 1997) and (Brinson, 1993) are linearly dependent on stress through the Clausius-Clapeyron coefficients C_m and C_a.

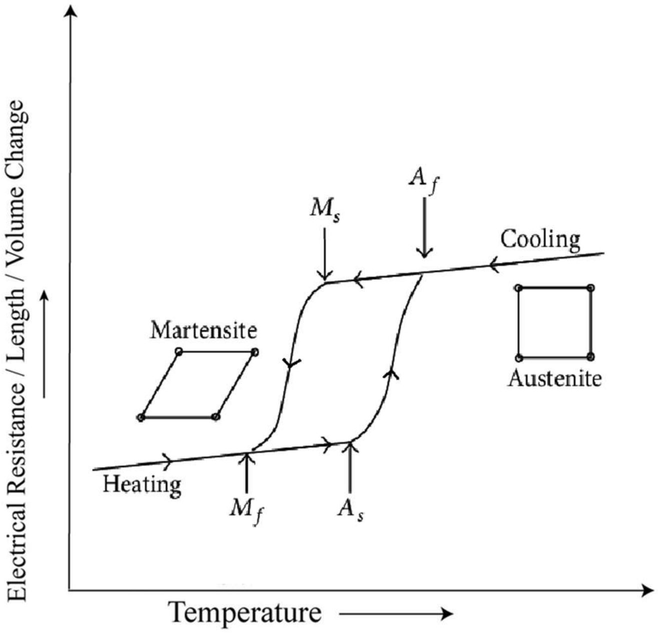


Figure 5 - SMA hysteretic behavior

As explained by Otsuka et al. in (Otsuka & Ren, 2005), many binary and ternary SMAs have a third phase between Austenite and Martensite, which is rhombohedral for Ni-Ti and Ni-Ti-Fe or Orthorhombic for Ni-Ti-Cu. The presence of a two-step transformation including this additional intermediate phase change can be observed through a DSC measurement as the one in Figure 6. Depending on the thermomechanical history of the material the martensitic transformation can be direct or two-step.

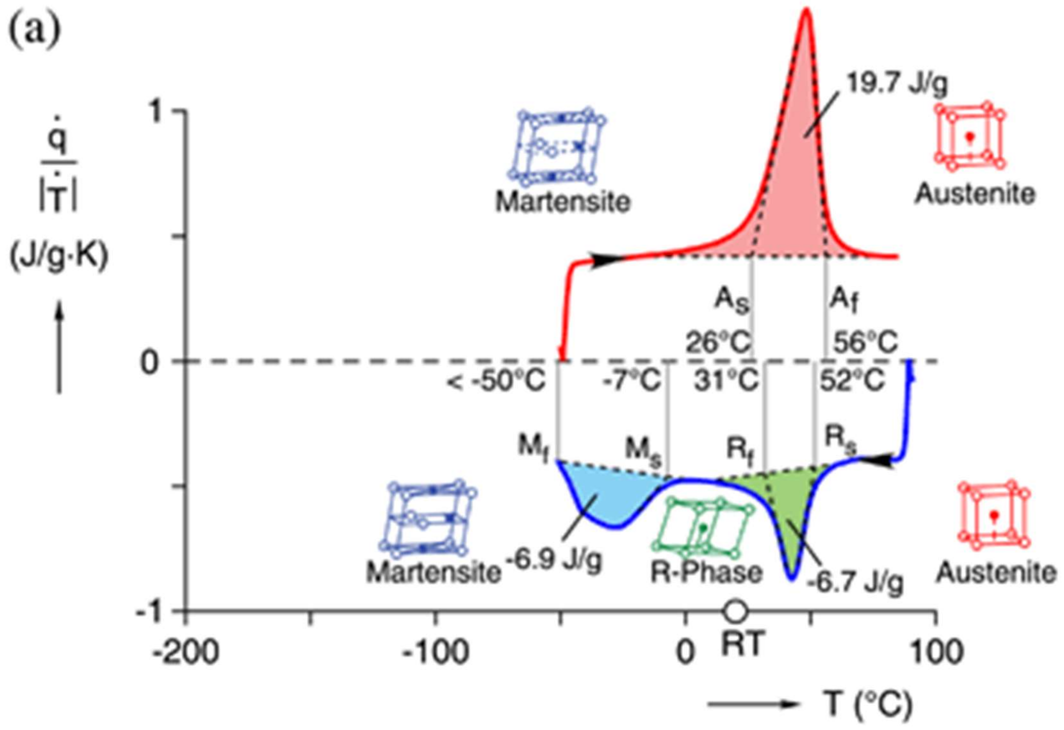


Figure 6 - DSC of a Ni-Ti SMA wire showing the two-step transition and including the rhomboedric phase

1.2 Pseudo-Elasticity

When the material is in the Austenitic phase, above A_f, if a sufficient load is applied to induce a transformation into detwinned Martensite, the observed change in shape is directly recovered upon load removal. This recovery, also known as pseudo-elasticity, is due to the Martensite instability without the applied load, which causes the material to revert back to Austenite and to its original shape, without any further increase in temperature.

This behavior is observed only below a temperature called M_d, above which the energy for slip is lower than the energy to initiate and propagate the martensitic transformation: in these conditions the SMA behaves most like a conventional elastoplastic metal.

1.3 (One-way) Shape Memory Effect

When the temperature is below M_F the SMA is in Martensite phase and cannot recover large deformations in a pseudo-elastic way.

When a load is applied to the SMA in Martensite phase, it causes all the differently oriented martensitic domains to reorient along the direction of the applied stress, and when the stress is removed the selected variant remains in most of the deformed material volume.

Even though the new crystal lattice rearrangement and the deformation persist after the load removal, this phenomenon is not plastic, instead is due to detwinning, and inside the material are present only elastic accommodation stresses, which help the return to the original shape when enough energy is given through heating the crystal above A_S. Raising the temperature, the transformation into Austenite progresses and is completed as A_F is reached. This is called Shape Memory Effect (SME), a scheme of the process is shown in Figure 7.

This effect is observed even though the stress is not removed prior to heating. In this case the shape recovery is incomplete, depending on the level of stress. This working mode is typical of actuation applications.

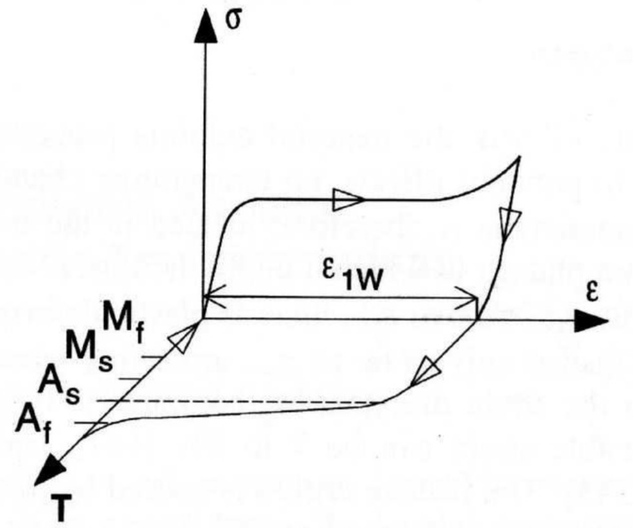


Figure 7 - Shape Memory Effect

1.4 Two-Way Shape Memory Effect

According to (Auricchio, Marfia, & Sacco, 2003) when a SMA sample is subjected to specific cyclic thermomechanical loading, known as training, the material response changes gradually, until it reaches a stable path, where we can observe, with respect to the untrained material:

- a decrement of the initial and final stress threshold for the phase transformations
- an increment of a residual permanent deformation

While the reaching of a stable deformation path improves the repeatability of actuation, the generation of (limited) permanent deformation by the training process allows to obtain reversible spontaneous shape changes during cooling and heating processes without application of any external stress, this is known as the Two-Way Shape Memory Effect (TWSME).

There are two main micromechanical explanations for this phenomenon, the first states that the change in response is due to oriented residual stresses occurring during the dislocation arrangement, which favors the growth of those martensite variants who are relaxed by the shape change. Moreover, since dislocation are linked to plasticity, it follows that the training effect produces an increment in the residual plastic strain.

The second interpretation points to the increment in residual permanent oriented martensite variants, that occur because of the dislocation arrangement and tend to grow during subsequent cycles, which do not allow the complete martensite-austenite transformation during unloading and heating.

To model the TWSME, we consider that a fraction of the austenite that should transform into twinned martensite when cooling is instead transformed into the detwinned variant, which is normally generated by applying an external stress.

1.5 Most important SMAs

Shape memory alloys are made of compositions of different metals such as:

- NiTi-based alloys: NiTi, NiTiCu, NiTiPd, NiTiFe, NiTiNb, NiFeGa, NiTiCo…

- Cu-based alloys: CuZn, CuZnAl, CuAlNi, CuAlNiMn, CuSn ...
- Fe-based alloys: FePt, FeMnSi, FeNiC ...

Currently, the Nickel-Titanium alloys are the most commonly used shape memory alloys, possessing transition temperatures ranging from − 50 °C up to 110 °C. There are several reasons for the preference of the Nickel-Titanium alloys in comparison to other alloys:

- Can be fabricated with common metalworking techniques.
- The transition temperature of the SME can be tailored to specific needs by simply changing the ratio between Nickel and Titanium.
- Are biocompatible.
- Possess a high shape memory strain of up to 8% and high transformation stresses.

1.6 SMA wire actuator systems

In this paragraph we present some of the configurations of SMA wire actuators that emerged during our literature review, which bear some resemblance to our target plant. SMA have many advantages over traditional actuators but do suffer from a series of limitations that may impede practical application. SMA actuators are typically actuated electrically, by an electric current resulting in Joule heating. Deactivation typically occurs by free convective heat transfer to the surrounding environment. Consequently, SMA actuation is typically asymmetric, with a relatively fast actuation time and a slow deactivation time.

1.6.1 Simple SMA actuators

The simplest actuator is a SMA wire that pulls a counterweight, shown in Figure 8.

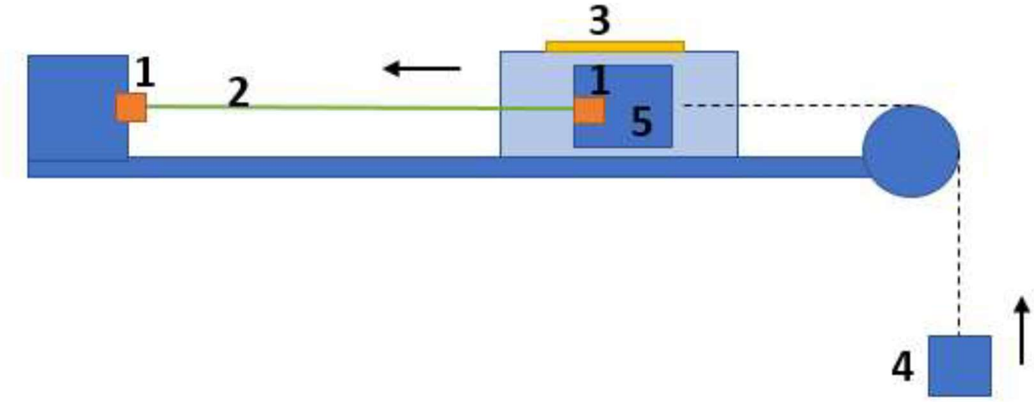


Figure 8 - Simple linear SMA actuator

This simple configuration uses the weight of the load to induce via stress the inverse transformation, from austenite to martensite, which binds the speed of the transformation to the load, this makes it viable only for testing the wire properties and for applications with sufficiently high loads, as seen in (Lee, Jin, & Ahn, 2013), (Lee & Kim, 2019) and (Daohui, Xingang, Jianda, Xiaoguang, & Bi, 2019). To enable the reverse transformation, a biasing force is needed, this is usually achieved, other than by a dead weight, through the use of a bias spring, as seen in (Shameli, Alasty, & Salaarieh, 2005) and (Lee, Jin, & Ahn, 2013). An example is shown in Figure 9.

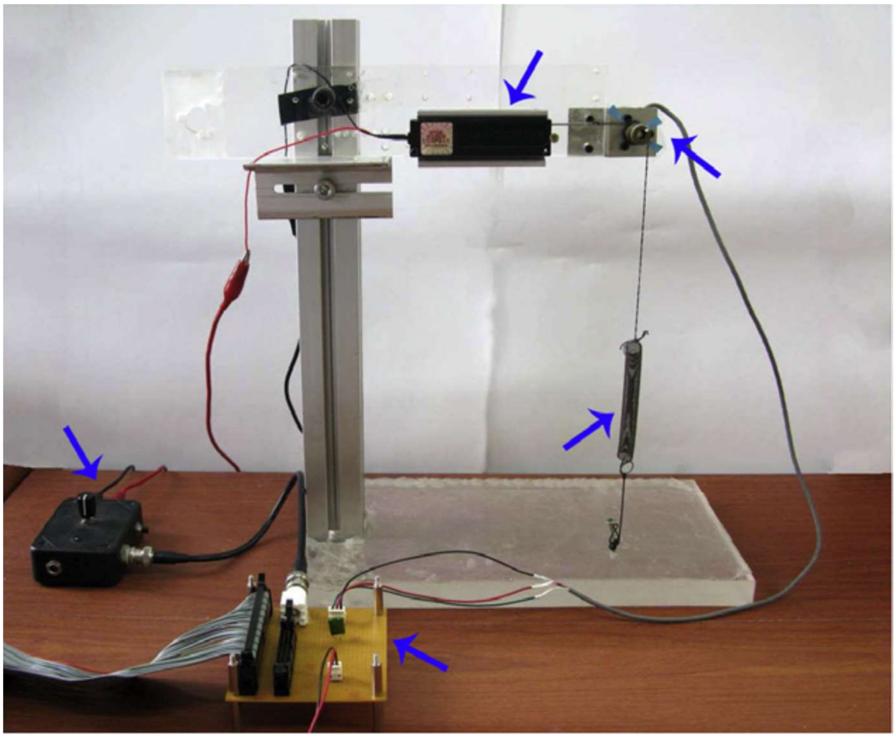


Figure 9 - SMA actuator with bias spring

Apart from examples of bare wires, papers in literature also describe actuators in which the wire is covered or lined with other materials. For instance, (Villoslada, Flores, Copaci, Blanco, & Moreno, 2015) and (Copaci, Blanco, & Moreno, 2019) devised a control for a flexible SMA-based actuator that consists of a wire wrapped a number of times around two pulleys inside a Bowden cable, shown in Figure 10.

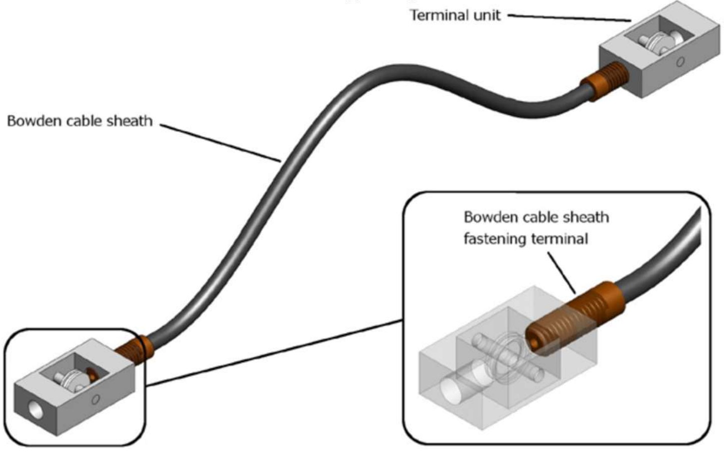


Figure 10 - SMA actuator with bowden cable and terminal pulleys

This configuration can exert the same strength as an actuator as long as the SMA wire while remaining shorter and allowing flexibility, however the presence of the Bowden cable limits the heat exchange with the environment. Counterworking SMA wires can work as the tendons for small robotic arms, as shown by (Ashrafiuon & Elahinia, 2015) and (Elahinia & Ashrafiuon, 2002).
Figure 11 shows an example of a robotic arm from (Elahinia & Ashrafiuon, 2002).

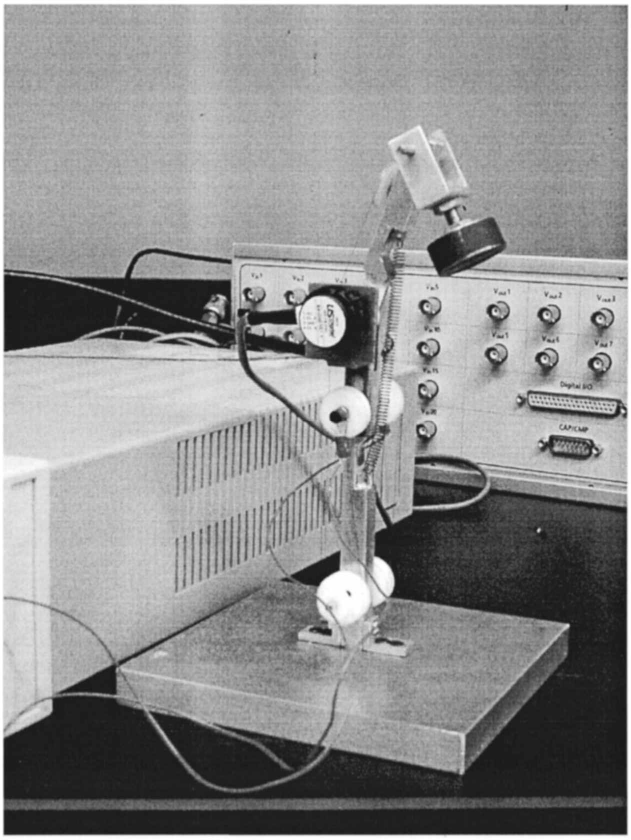


Figure 11 - 1 d.o.f. robot arm actuated with SMA wire and bias spring

1.6.2 SMA soft actuators

In recent years SMA wires have found use as a substitute for classical drives in soft robots, which are highly flexible and compliant robots, made by soft materials such as polymers or silicone rubber, as the name implies.

A review of some interesting SMA-actuated robots was published by M.M. Kheirikhah et al. (Kheirikhah, Samaneh, & Edalat), including crawlers and walkers, as in Figure 12, and even flowers like the one in Figure 13.

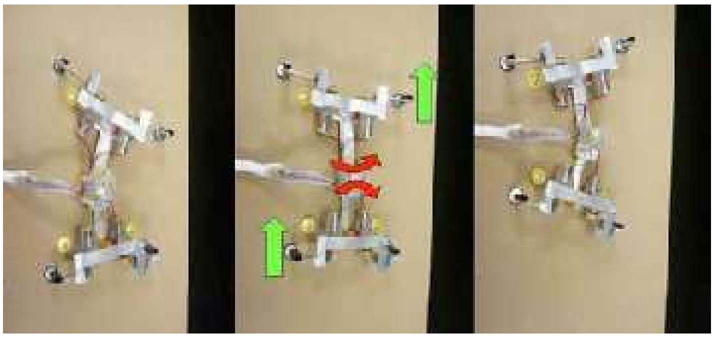


Figure 12 - SMA actuated walker robot

Figure 13 - SMA actuated flower robot

There are also some biomedical uses, such as for the prosthetic hand in Figure 14 developed by (Kim, Han, Song, & Ahn, 2016).

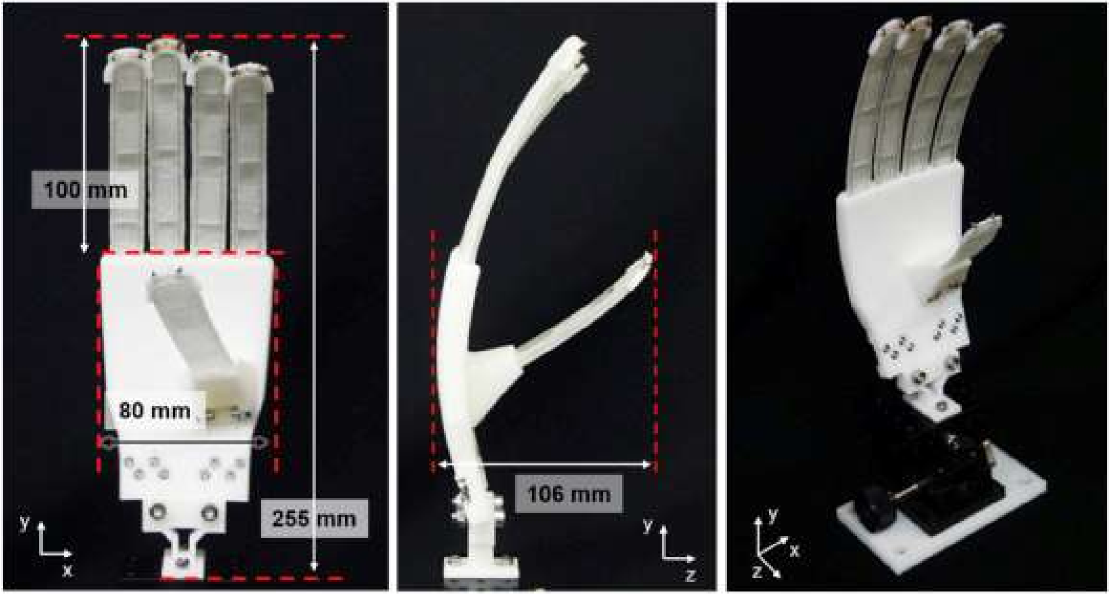


Figure 14 - SMA actuated prosthetic hand

Regarding the use of SMA wire in soft gripper actuators, there is a very rich literature on bending and twisting actuators.

An example of bending actuator is given in Figure 15, taken from (Rodrigue, Wang, Kim , & Ahn, 2017), where a SMA wire is embedded inside a polymeric matrix out of the neutral plane so that, when the wire contracts, the whole matrix bends.

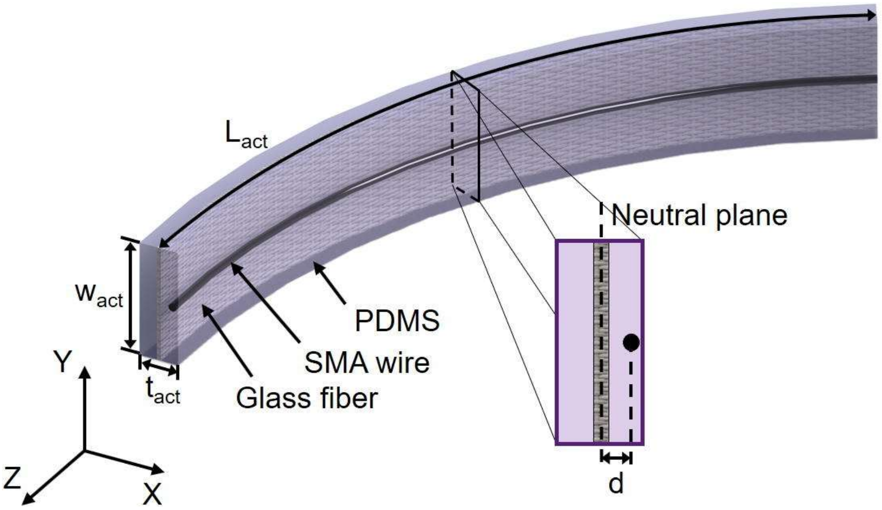


Figure 15 - Soft bending SMA actuator

Other similar grippers are found in (Kim, Han, Song, & Ahn, 2016) and (Lee, Chung, & Rodrigue, 2019). In all these applications, where the SMA wire is embedded in a heterogenous matrix, the heat transport properties and hence the transformation characteristics especially of timing, bandwidth and energy are different from bare wires. The control strategies will have therefore to be developed accordingly.

1.6.3 Our target actuator

The actuator we based on for all our thesis work is as a linear actuator composed by a SMA wire 10 cm long and with a diameter of 0.65 mm surrounded by a silicon rubber coating 1mm thick in direct contact with the wire, operating while being surrounded by air.

The wire diameter was chosen so that the actuator would be fast to cool and easy to heat up without too much power consumption.

The silicone diameter needed to be large enough for the coating to be fabricated easily and to ensure a large enough external surface area to facilitate the convective heat exchange, while being small enough to neglect its mechanical properties and not possessing too big of a heat capacity for the system to not heat up above the austenitic transformation temperatures.

The wire is actuated through a voltage source, with maximum voltage V_{SAT} = 1 V, and to maintain a good fatigue life is intended to be used for applications of 1000 cycles, with loads of approximately 200 MPa and maximum strain 3%.

The presence of the silicone complicates the modeling of the system because it must be included in the heat exchange path. Its mechanical contribution can instead be neglected with respect to the external applied force due to its small mass and the long actuation times of the actuator.

Considering the actuator performance, the presence of the silicone could help during the wire cooling, the main limiting factor on the bandwidth of this actuator, because of its large specific heat capacity which can act as a heat well for the wire and its larger surface area which helps speed up the convective heat transfer. The gradual heating of the silicone however may be a problem, and it is envisaged that its temperature will strongly depend on the actuation history.

2 Review of SMA constitutive models

For the purpose of this thesis, it is required to have a model of the actuator system, which for the most part coincides with a model of a SMA wire. The modelled SMA will be used to mimic the behavior of the physical system and simulate its interaction with the innovative controller to be devised.
In the literature we find countless models for SMAs, based on many different approaches.
The aim of modeling SMAs are:

- Understanding the underlying physics and mechanisms that cause the observed effects.
- Identify the material properties or processing parameters that can yield the desired effects in terms of SME, SE, TWSME etc.
- Predict the material response.

A broad classification of some salient approaches to SMA modeling is:

- Microscopic approach
    - Free energy
    - Micromechanics
- Macroscopic approach
    - Hysteresis based
    - Free energy based
    - Phenomenological

2.1 Microscopic approach

Microscopic models start from microstructural features at the lattice or grain-crystal level, such as phase nucleation, interface motion and martensite twin growth, to describe the behavior of the SMA. They are mostly based on the Landau-Devonshire free energy or molecular dynamics.
Most of these models are based on a notion of unit cell or Representative Volume Element (RVE) and use suitable volume averaging schemes to capture the overall effect in the SMA. Even though these models can describe the SMA behavior from the highest level, they usually are very large and complex and thus require large computational powers to yield accurate results. As a reference, some examples of this type of models are found in (Cisse, Zaki, & Ben Zineb, 2016).

2.2 Macroscopic approach

Macroscopic models, as the name suggests, attempt to capture the SMA response at a macroscopic scale (typically > 100 μm). Suitable assumptions and approximations are made to account for the macroscopic effects. The most important aspects of this modeling approach are phenomenology and the use of thermodynamic principles: while some models possess a thermodynamic framework, others rely significantly on phenomenology, assuming that the thermodynamic laws are satisfied a priori. Another class of models attempts to capture the hysteretic input-output response without explicitly accounting for underlying physics. These models use different types of hysteresis operators to predict specific stress-strain-temperature response.

2.2.1 Hysteresis models

This approach models the complex hysteresis response in SMAs by constructing a suitable input-output model that captures the observed hysteresis between a forcing function and the response variable. Such models are commonly used in modeling magnetic and ferroelectric materials and are frequent in control literature. In (Smith, 2005) R. Smith gives a detailed discussion on this approach for modeling SMAs defining two main types of hysteresis models, which differ in terms of the nature of the hysteresis operator that is used to track the response:

1. Preisach type models
2. Duhem-Madelung models.

Preisach models follow an integration of the response of individual relays that switch between two fixed states, each representing in this case, the type of transformation. Critical driving forces are prescribed for switching in both the directions. Papers such as (Ortin, Preisach modeling of hysteresis for a pseudoelastic Cu-Zn-Al single crystal, 1992) and (Ortin & Delaey, Hysteresis in shape-memory alloys, 2002) used this approach to capture the SE and SME hysteresis in SMAs. One of the advantages of this approach is the ease to model inner loops and return point memory.

Duhem-Madelung models capture the hysteresis in an incremental manner using two differential operators, one for the loading segment and the other for the unloading segment of the load path are used to capture the response. In (Ivshin & Pence, 1994) Y. Ivshin and T.J. Pence used this approach to capture the hysteresis in SMAs.

Due to its simplicity, this approach is employed in real time control of SMA devices and to predict inner loop response. However, the approach becomes tedious when effects like TWSME, changes in hysteresis and multiple variants are considered. Furthermore, it is not possible to use it in the design of the material or a of system since it does not provide sufficient insight into the phase transformation process in the material.

2.2.2 Free energy-based models

In this approach, a more consistent attempt is made to identify a suitable form of free energy or potential that represents the state of the system and introduces the phase-transformation-related effects either as changes in free energy itself or through a generalized constitutive function like dissipation potential. The choice of field or state variables determines the type of free energy, e.g., Gibbs or phenomenological Landau-Devonshire forms.

The free energy can be mechanistically decomposed like proposed in (Fried & Gurtin, 1994) and (Srinivasa & Rajagopal, 1999) as the sum of:

$$Free\ energy = Mechanical\ energy + Thermal\ energy + Exchange\ energy + Latent\ heat$$

The exchange energy is a function of the phase volume fraction and it usually determines the nature of the transformation kinetics through the dissipation potential.

Usually, the interaction between individual variants of Martensite is ignored and only the interaction between Martensite and Austenite is considered. The exchange energy also determines the nature of hysteresis and memory in the material response.

2.2.3 Phenomenological models

Phenomenological models separate out the constitutive stress-strain-temperature relationship from the driving forces and the evolution (kinetics) of the phase transformation. In general, most of the salient differences in approaches stem from the choice of internal variables and the description of their evolution in terms of the thermomechanical driving forces.

Their simplicity and their good ability to capture the macroscopic SMA response, at least in a 1-D setting, has resulted in large number of phenomenological models.

Several models for evolution kinetics in SMAs use a σ-T phase diagram to identify appropriate transformation (active) and dead zones. In each active zone appropriate evolution functions are defined to compute the extent of phase transformation. While most of these models are 1-D in terms of uniaxiality of stress and strain, some 3-D models exist that account for multi-axial stress. This approach leads to simplified models facilitating their use as design tools. K. Tanaka's model (Tanaka, 1986) was among the first to use this approach for SMA to study superelasticity with only the stress-induced martensitic fraction (ξ_S) as an internal variable, then an evolution law was deduced in terms of stress and temperature. The basic assumption he made was that the thermomechanical process of the SMA material is fully expressed by three major state variables: strain, temperature, and Martensite volume fraction.

The constitutive equation *2.1* relates the state variables stress σ, strain ε and temperature T in the terms of the Martensite volume fraction ξ.

$$\sigma - \sigma_i = [D(\xi)(\varepsilon - \varepsilon_i) + \Omega(\xi)(\xi - \xi_i) + \Theta(T - T_i)] \tag{2.1}$$

$$D(\xi) = D_A + \xi(D_M - D_A) \tag{2.2}$$

$$\Omega(\xi) = -\varepsilon_L D(\xi) \tag{2.3}$$

Where D is the total elastic modulus, which is a linear combination of the elastic moduli of Austenite and Martensite, respectively D_A and D_M, proportionally to the total martensitic fraction ξ; Ω is the transformation tensor and ε_L is the maximum residual strain, Θ is the thermoelasticity modulus; the subscript *i* indicates the initial value of the variable at the start of the transformation. The transformation kinetics laws differ depending on the current temperature and stress:

$if\ \left[(T > M_F) \wedge \left(C_M(T - M_S) < \sigma < C_M(T - M_F)\right)\right]\ then$

$$\xi = 1 - e^{a_M(M_S - T) + b_M\sigma} \tag{2.4}$$

$if\ \left[(T > A_S) \wedge \left(C_A(T - A_F) < \sigma < C_A(T - A_S)\right)\right]\ then$

$$\xi = 1 - e^{a_A(A_S - T) + b_A\sigma} \tag{2.5}$$

where a_A, a_M, b_A and b_M are material constants, dependent on the transition temperatures A_S, A_F, M_S and M_F, C_A and C_M are two material constants called stress-influence or Clausius-Clapeyron coefficients, which indicate the influence of stress on the transition transformation and are obtained from experimental tests. To obtain a better fit to the experimental data, in (Liang & Rogers, 1997) C. Liang and C.A. Rogers modified the phase kinetics from an exponential to a cosine-based function.

$if\ \left[(T > M_F) \wedge \left(C_M(T - M_S) < \sigma < C_M(T - M_F)\right)\right]\ then$

$$\xi = \frac{1 + \xi_{S_i}}{2} + \frac{1 - \xi_{S_i}}{2}\left\{cos\left[\frac{\pi}{M_S - M_F}\left(T - M_F - \frac{\sigma}{C_M}\right)\right]\right\} \tag{2.6}$$

$if\ \left[(T > A_S) \wedge \left(C_A(T - A_F) < \sigma < C_A(T - A_S)\right)\right]\ then$

$$\xi = \frac{\xi_i}{2}\left\{cos\left[\frac{\pi}{A_F - A_S}\left(T - A_S - \frac{\sigma}{C_A}\right)\right] + 1\right\} \tag{2.7}$$

Both Tanaka (1986) and Liang and Rogers (1990) were successful in capturing the superelastic behavior. Based on the broad framework of Tanaka (1986) and Liang and Rogers (1990), Brinson (1993) proposed a modified model to also account for the shape memory effect by splitting the martensitic phase fraction into two parts, temperature induced twinned fraction (ξ_T) and stress induced detwinned fraction (ξ_S), this differentiation of the phase fractions is used to capture stress/strain recovery. One of the important simplifying assumptions is that the twinned Martensite does not contribute to the recovery stress. A phase diagram was constructed to describe phase transformation involving Austenite and the two

Martensite variants. L.C. Brinson's model (Brinson, 1993) has since been used widely for modeling SMAs.

2.3 Brinson (1993) model

We chose Brinson's model from (Brinson, 1993) since it is the most complete among those phenomenological models we considered and it is widely used, with many examples from literature to refer to.

It is also possible, with some modification, to account for partial cycles, that is when the transformation is interrupted before completion and then reversed.

2.3.1 Constitutive law

The constitutive law of the Brinson model is a modified version of (Tanaka, 1986), accounting for the two Martensite variants.

$$\sigma - \sigma_i = \left[D(\xi)\varepsilon - D(\xi_i)\varepsilon_i + \Omega(\xi)\xi_S - \Omega(\xi_i)\xi_{S_i} + \Theta\left(T_1 - T_{1_i}\right)\right] \tag{2.8}$$

$$\varepsilon = \frac{1}{D(\xi)}\left[\sigma - \sigma_i + D(\xi_i)\varepsilon_i - \Omega(\xi)\xi_S + \Omega(\xi_i)\xi_{S_i} - \Theta\left(T_1 - T_{1_i}\right)\right] \tag{2.9}$$

$$D(\xi) = D_A + \xi(D_M - D_A) \tag{2.10}$$

$$\Omega(\xi) = -\varepsilon_L D(\xi) \tag{2.11}$$

Where σ is the stress applied on the wire, ε is the wire strain, D is the total elastic modulus, which is a linear combination of the elastic moduli of Austenite and Martensite, respectively D_A and D_M, proportionally to the total martensitic fraction ξ; Ω is the stress/transformation coefficient and ε_L is the maximum residual strain and finally Θ is the thermoelasticity modulus. The subscript *i* indicates the initial value of the variable at the start of the transformation.

2.3.2 Transformation kinetics

The SMA wire undergoes different phase transformations, depending on the position of its inner state in the σ-T plane shown in Figure 16.

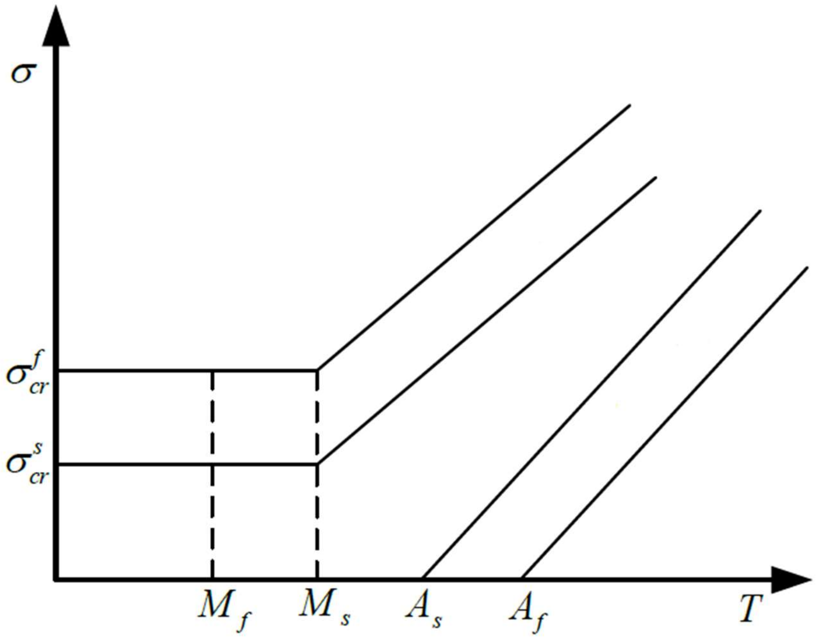


Figure 16 - SMA phase plane with critical stresses

The transformation kinetic laws according to Brinson are again a modified version of those from (Liang & Rogers, 1997).

Conversion to detwinned Martensite

$if\ [(T > M_S) \wedge (\sigma_S{}^{cr} - C_M(T - M_S) < \sigma < \sigma_F{}^{cr} - C_M(T - M_S))]\ then$

$$\xi_S = \frac{1 + \xi_{S_i}}{2} + \frac{1 - \xi_{S_i}}{2} cos\left\{\frac{\pi}{\sigma_S{}^{cr} - \sigma_F{}^{cr}}[\sigma - \sigma_F{}^{cr} - C_M(T - M_S)]\right\} \qquad 2.12$$

$$\xi_T = \frac{\xi_{T_i}}{1 - \xi_{S_i}}(1 - \xi_S) \qquad 2.13$$

$if\ [(T < M_S) \wedge (\sigma_S{}^{cr} < \sigma < \sigma_F{}^{cr})]\ then$

$$\xi_S = \frac{1 + \xi_{S_i}}{2} + \frac{1 - \xi_{S_i}}{2} cos\left[\frac{\pi}{\sigma_S{}^{cr} - \sigma_F{}^{cr}}(\sigma - \sigma_F{}^{cr})\right] \qquad 2.14$$

$$\xi_T = \frac{\xi_{T_i}}{1 - \xi_{S_i}}(1 - \xi_S) + \Delta_{T\xi} \qquad 2.15$$

$if\ [(M_T < T < M_S) \wedge (T < T_0)]\ then$

$$\Delta_{T\xi} = \frac{1 + \xi_{T_i}}{2}\left\{cos\left[\frac{\pi}{M_S - M_F}(T - M_F)\right] + 1\right\} \qquad 2.16$$

$else$

$$\Delta_{T\xi} = 0 \qquad 2.17$$

Conversion to Austenite

$if\ [(T > A_S) \wedge (C_A(T - A_S) < \sigma < C_A(T - A_S))]\ then$

$$\xi = \frac{\xi_i}{2}\left\{cos\left[\frac{\pi}{A_F - A_S}\left(T - A_S - \frac{\sigma}{C_A}\right)\right] + 1\right\} \qquad 2.18$$

$$\xi_S = \frac{\xi_{S_i}}{\xi_i}\xi \qquad 2.19$$

$$\xi_T = \frac{\xi_{T_i}}{\xi_i}\xi \quad 2.20$$

3 SMA control techniques: state of the art

Over the years, many control methods have been proposed for the many different SMA actuators, although we can see that some methodologies are more used than others.
A comprehensive review of control schemes for SMA is presented by M.H. Elahinia et al. in (Elahinia, Esfahani, & Wang, 2010), providing five key characteristics of all SMA actuators and effective controllers:

1. Bandwidth is limited by passive cooling.
2. Rapid heating to near transformation temperature should always be included in the control scheme.
3. Adaptive control laws help maintain stability to disturbances and un-modeled parameters.
4. A biasing force is required for controlled two-way motion.
5. Combining position, force, and temperature control can help reduce sensitivity to disturbances and un-modeled parameters.

3.1 PID-based control

Many methods found in literature are based upon variations in the linear PID control scheme. These models are easy to implement since they do not need to model the system. Many people adopted a PID whose parameters are tuned via fuzzy logic, such as S.H. Lee. And S.W. Kim in (Lee & Kim, 2019). This closed loop scheme uses a combination of the error and its derivative to modify the proportional, integral and derivative gains via the Mamdani rule-based process, which is outlined in Figure 17.

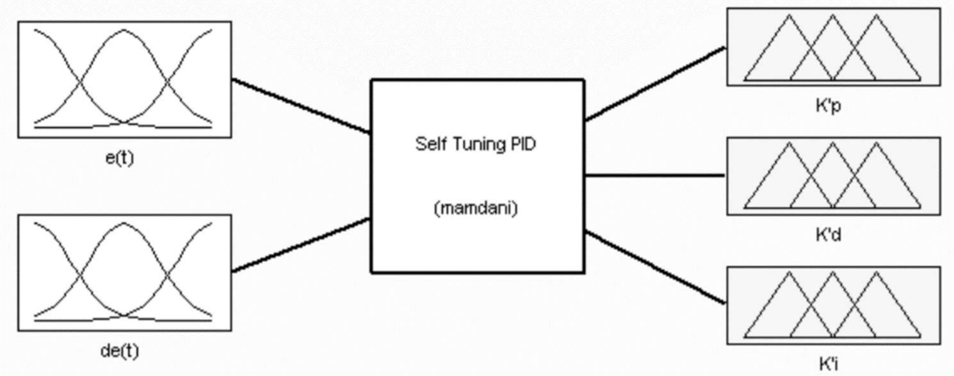
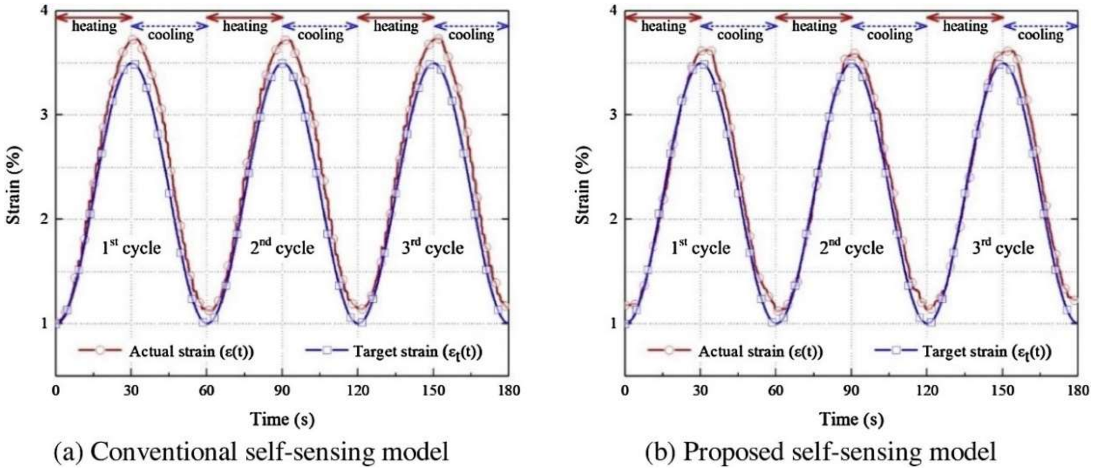


Figure 17 - Fuzzy PID scheme with membership functions

In this scheme, some labels such as “large”, “medium” or “small” are defined for the inputs of the controller (position error and derivative) and outputs (the controlled variable, mostly strain but (Lee & Kim, 2019) uses as feedback variable the wire resistance).
Through a membership function, to each input is assigned a membership value from 0 to 1 in each label, this is called fuzzification.
After this, all the inference rules are applied to obtain the fuzzified outputs from the inputs, which are then defuzzified to obtain crisp output values.
In (Lee & Kim, 2019) the proposed a fuzzy PID controller is applied in conjunction with a self-sensing model, which closes the control loop around the wire resistance and then obtains the wire strain in open loop, based on a curve fitting of both the heating and cooling strain-resistance curves.
In Figure 18 we see the performance of their fuzzy PID and their self-sensing model compared to another one they considered, which they found in the literature.



(a) Conventional self-sensing model (b) Proposed self-sensing model

Figure 18 - Performance of the fuzzy PID scheme by S.H. Lee and S.W. Kim with self sensing models

As another example, E. Shameli et al. (Shameli, Alasty, & Salaarieh, 2005) proposed a PID-P3 controller to perform a precise position tracking. Their PID-P3 controller has a cubic term of positioning error. This term has considerable effects in reducing the settling time and overshoot of the positioning system.

3.2 Variable structure control

Another type of control scheme found often in literature is Variable Structure Control (VSC) and its subcategory Sliding Mode Control (SMC).

The SMC algorithm utilizes a switching control law to drive the nonlinear state trajectory of the plant onto a user-specified surface in the state space, and to maintain it on this surface for all subsequent times. When the state trajectory of the plant is above the surface the control path has a given structure and when the trajectory drops below the surface a different control is applied.

(Lee, Jin, & Ahn, 2013) proposed a SMC with a hyperbolic tangential sliding surface in the $e, \dot{e}$ plane, seen with equation

$$s = \dot{e} + \lambda_1 \, tanh(ke) \tag{3.1}$$

where k is a positive constant value which is related to the slope of the hyperbolic tangent and λ_1 denotes the velocity limit of the SMA actuator.

The above sliding surface is designed in order to attenuate the unfavorable effects of large velocity error in the SMA actuator system.

The hyperbolic tangential sliding surface *3.1* is continuous and differentiable, with first derivative

$$\dot{s} = \ddot{e} + \lambda_1 \frac{k}{coth^2(ke)} \dot{e} \tag{3.2}$$

A dynamic attractor of the reaching phase is

$$\dot{s} + \lambda_2 s = 0 \tag{3.3}$$

The closed loop error dynamics thus becomes

$$\ddot{e} + \left(\lambda_1 \frac{k}{coth^2(ke)} + \lambda_2\right) \dot{e} + \lambda_1 \lambda_2 \, tanh(ke) = 0 \tag{3.4}$$

In Figure 19 we can see the closed loop error dynamics in the $e, \dot{e}$ plane, after the application of the proper control input which eliminates the nonlinearities in the system.

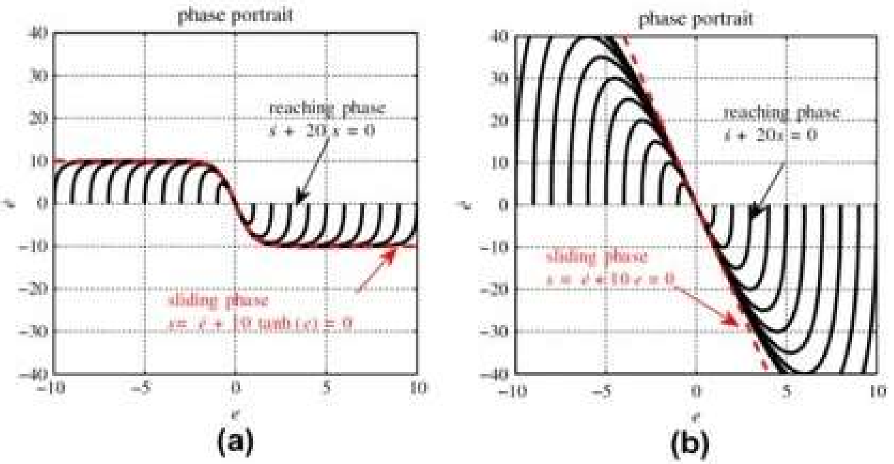


Figure 19 - Phase portrait of closed-loop error dynamics: (a) hyperbolic tangential sliding surface; (b) linear sliding surface.

In Figure 20 we see the performance of this controller compared to a normal PID controller.

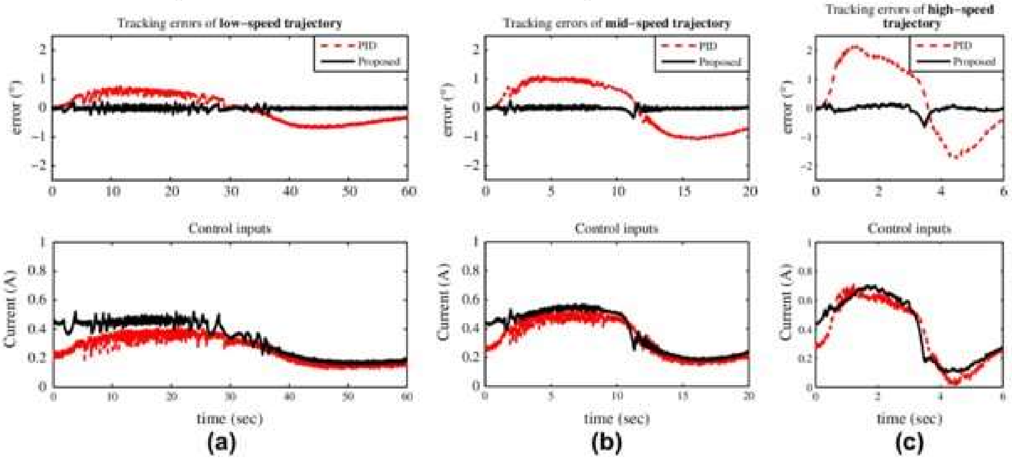


Figure 20 - Performance of Lee's SMC compared to a regular PID

3.3 Other methods

As an example of alternative control methods, it is interesting to mention schemes that incorporate estimators for unmeasured variables.

In (Daohui, Xingang, Jianda, Xiaoguang, & Bi, 2019) Z. Daohui et al. proposed a modified feedback linearization control consisting of three parts, feedback linearization, active modeling and outer-loop controller, schematize in Figure 21.

The Kalman filter estimates the enlarged states, consisting of the regular states $Z(t) = (T - T_{amb}, d, \dot{d})^T$ and the modeling errors $f = (f_1, f_2, f_3)^T$, considered as constant and noisy unknown parameters.

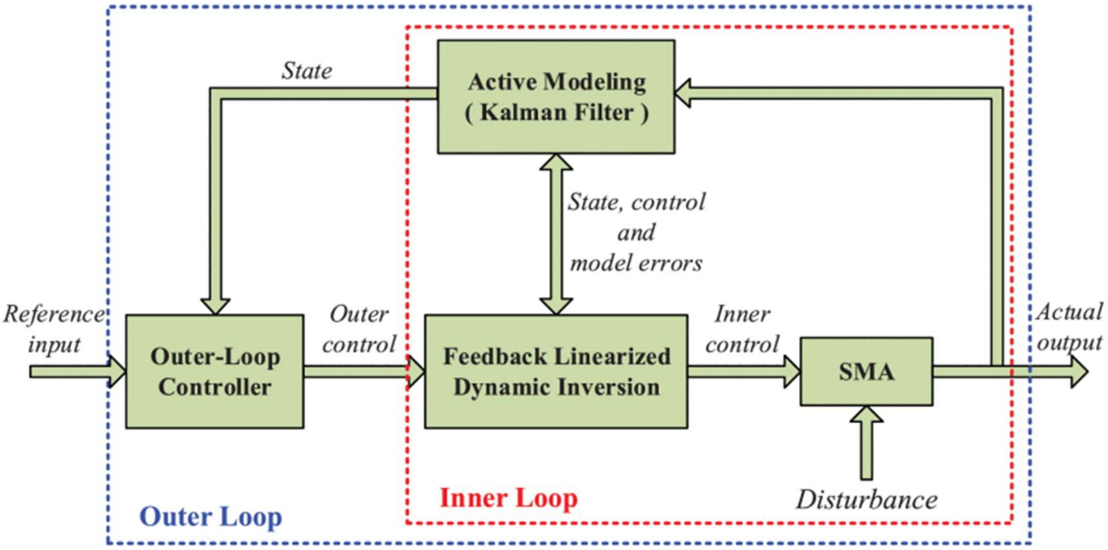


Figure 21 - Daohui's control scheme

In Figure 22 we can see the performance of this Active Modeling State Feedback Controller (AM-SFC) confronted with the state feedback controller without active estimation of modeling errors they improved on.

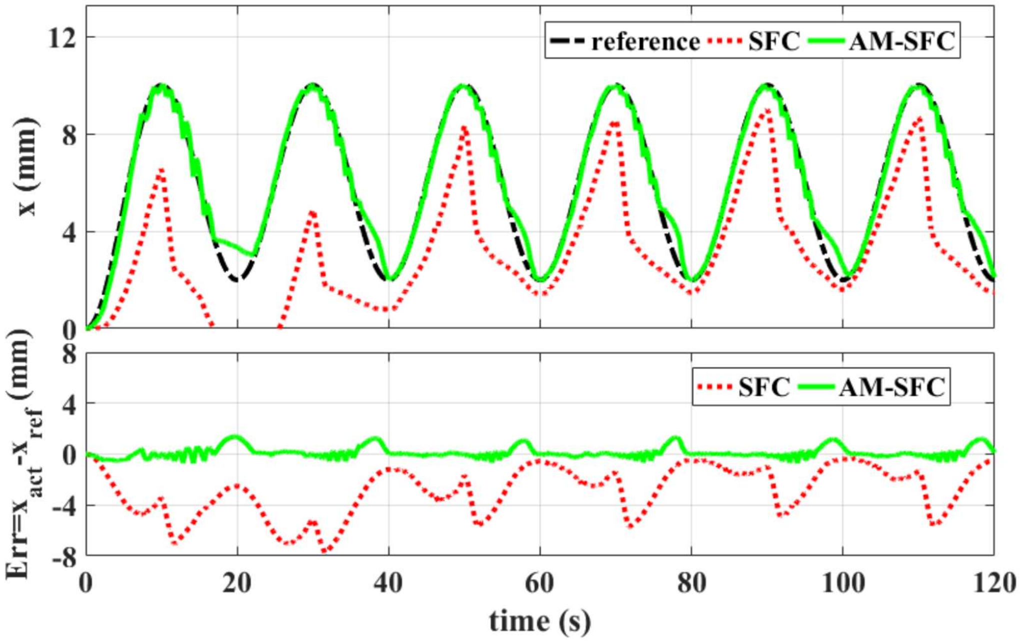


Figure 22 - Comparison between state feedback control with and without active state modeling

3.4 Discussion and choice of a control strategy

After reviewing the literature, we can draw some conclusions.

PID based controllers are quite simple and versatile for this type of application but they require additional actions to compensate for the system nonlinearities.

SMC shows high accuracy in reference tracking both stationary and variable input signals and a good handling of uncertainties and nonlinearities in SMA.

Feedback linearization coupled with active modeling is a very interesting concept, however it is very heavily dependent on accurate parameter estimation, which can be problematic for a branched nonlinear system like a SMA.

In order to make the most of the different approaches and of the state observer presence, we propose a PID based controller which parameters are variable and dependent on the state variables, which will be estimated through active system modeling by our Extended Kalman Filter (EKF), with the dual function of state estimator and error corrector.

We expect that this added knowledge on the thermomechanical characteristics of the system will allow us to obtain some which could be used to improve the performance of a PID type controller.

4 Materials

Here we discuss the procedures and methods used to prepare and characterize a SMA wire for our application and those used to obtain its physical parameters.

4.1 Choice of the SMA wire

The SMA for our application had to satisfy some specifications, mainly:
Good cycling stability, which translates into having no residual stress after a complete heating and cooling cycle.
Transformation temperatures compatible with actuation at room temperature.
Limited hysteresis in order to have fast transformation and low energy losses.
Recoverable transformation strain greater than 2% at 250 MPa.
Considering different binary and ternary NiTi-based alloys seen in (Otsuka & Ren, 2005), NiTiCu appeared to be a suitable composition, which best satisfied these specifications. The addition of Cu to NiTi considerably decreases thermal hysteresis, increases the transformation temperatures and lowers the transformation stresses. Different amounts of Cu produce these effects to variable levels. In addition, mechanical (strain hardening) and heat (ageing) treatments, play a fundamental role in determining the final properties.
NiTiCu has often displays a 1-step transformation from austenite to martensite for Cu contents lower than approximately 7%, unlike for example NiTi, which normally has a rhomboidal phase between austenite and martensite: this helps keep the hysteresis low and thus limits energy losses.
Ultimately, $NiTiCu_6$ was selected for the purpose.

4.2 Wire preparation

The starting material was $NiTiCu_6$ wire (Kyokalloy), with a diameter of 1mm supplied from the producer in cold-drawn state.
Considering the specifications for the actuator set out in section 1.6.3, it was decided that a diameter of 0.65mm would be appropriate: with a smaller diameter, the energy expenditure and the cooling times decrease, and the required force to deform martensite is smaller.
The procedure to reduce the wire cross-section is cold drawing, so that the grain size is maintained and a suitable work hardening (increase in dislocation density) is produced. These dislocations introduce elastic stresses in the material which help the material recover its original shape by facilitating the phase change from martensite to austenite and strengthen the material.
The wire section reduction, especially from a cold-drawn supply state, cannot generally happen with only one wire drawing step, so the material needs to be annealed in order to be drawn again. The annealing process consists in heating the wire at 700 °C, allowing the lattice to relax and eliminate most dislocations.
The final drawing process achieves the desired wire section by reducing it by 40%.
The last step in the wire fabrication is shape setting, through which the wire is heated to around 400-500 °C while kept in tension, this straightens the wire and partially reduces hardening, achieving a compromise between maximum recoverable strain and material strength, depending on the application. This last process, combined with the cold working, has great influences on the material transformation parameters, such as the transformation temperatures.
In our case the selected ageing (shape setting) treatment is 460°C 11min followed by quenching in water. This was optimized using the published method (Pittaccio & Garavaglia, 2014).

4.3 SMA wire characterization methods

Here we discuss the measurements done on the SMA wire sample to identify the model parameters useful for our purposes.

4.3.1 DSC

The first measurement done on our SMA wire was DSC (Differential Scanning Calorimetry).
With this technique, the difference in the amount of power required to increase the temperature of a sample vs. a reference is measured as a function of temperature. Both the sample and reference are maintained at nearly the same temperature throughout the experiment. The result of a DSC experiment is a curve of heat flow versus temperature or versus time, in Figure 23 we can see a typical DSC curve for a SMA.

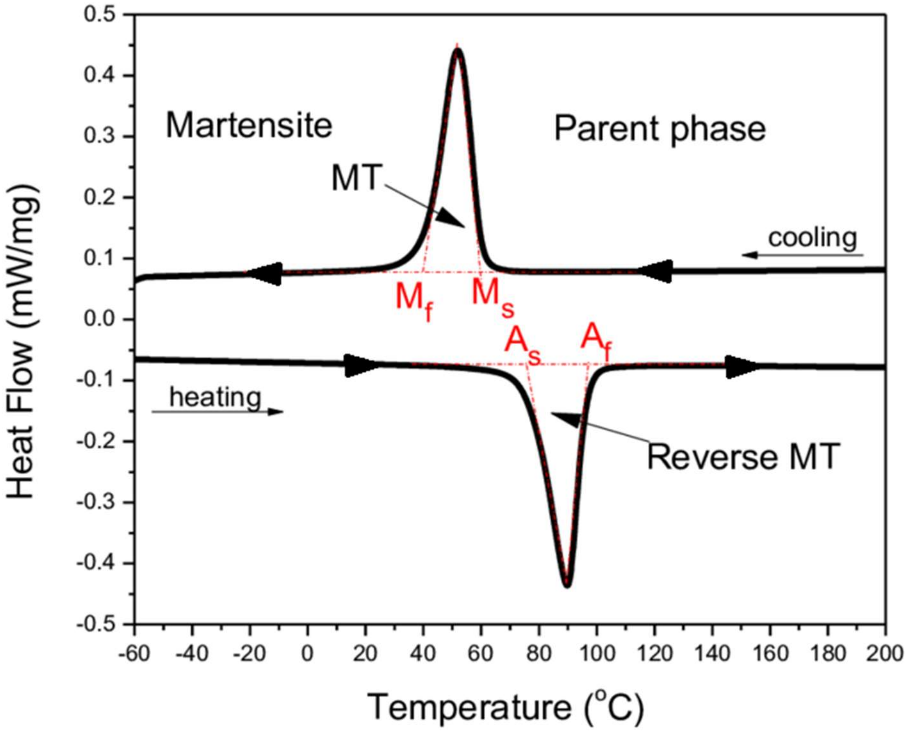


Figure 23 - Typical DSC curve for a SMA with 1-step transformation

When a peak or valley appears on the curve a phase change is occurring in the material. This curve can be used to calculate enthalpies of transitions by evaluating the area under the peak corresponding to a given transition.
The width of the peak gives us an information about the spread in local transformation characteristics, so that a broader peak corresponds to microstructures with large local variations in crystal energy (e.g., internal stresses or strains, complex dislocation networks, etc.).
The beginning and the end of the peak lets us estimate the starting and finishing temperatures of each transformation. The narrower the intervals among the transition temperatures, the lower the hysteresis, the faster the phase transitions for a set heating/cooling rate.

4.3.2 SMA stabilization

After the DSC measurement, the SMA wire is subjected to a stabilization process: the wire is repeatedly pulled until the desired strain, in our case 3%, and then released, while the temperature is maintained 10 °C above A_F; the loading cycle is repeated 50 times in 5 minutes, this process is illustrated in Figure 24.

This training process allows the wire to reach a steady state with repeatable loading cycles up to the desired strain without residual stress when the load returns to 0.

This training process allows the wire to sustain loading cycles up to the desired strain without residual stress when the load returns to 0.

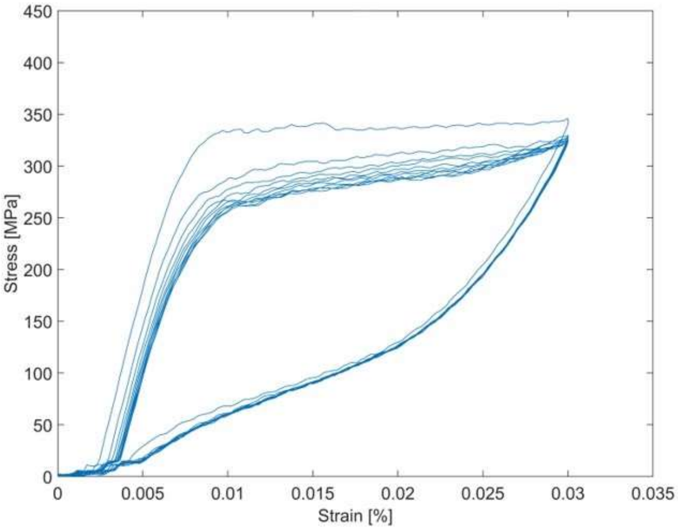


Figure 24 - SMA training cycles

4.3.3 Isothermal stress/strain measurements

After stabilization, the SMA wire can be tested mechanically in isothermal stress/strain cycles at various temperatures. Tensile tests are conducted in a universal test machine (Electropulse E3000, Instron), by which the sample is stretched to increasing strains from 1% to 5% at two temperatures, one 10°C below M_F and one 10°C above A_F, to test the material mechanical properties when starting from pure martensite and pure austenite phase respectively. In Figure 25 is shown a typical set of curves at low and high temperatures.

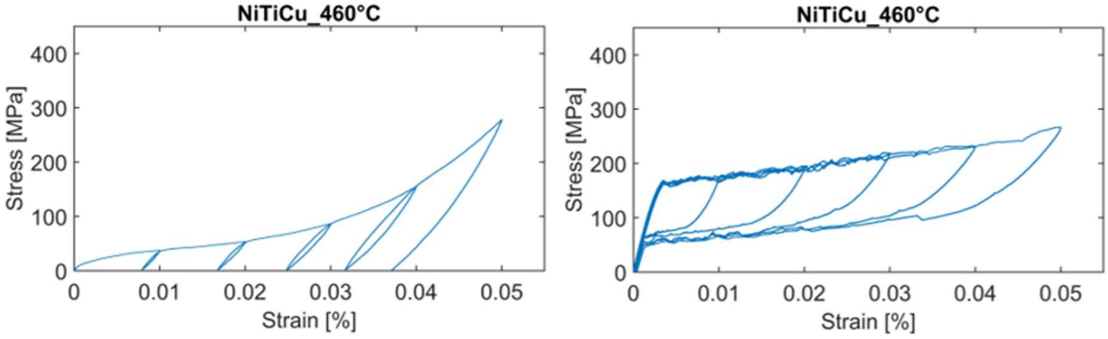


Figure 25 – Comparison between stress/strain cycles for stabilized SMA starting from Martensite (left) and Austenite (right)

In the curves obtained at temperatures higher than A_f we can clearly distinguish the initial elastic behavior, where the slope of the graph is the austenite elastic modulus, then around 170 MPa the change into martensite starts and we find the transformation plateau. The transformation into martensite is gradual and we cannot observe the martensite elastic behavior since the material does not transform completely within the range of these graphs. After reaching the desired strain, the material elastically contracts until it starts the reverse transformation back to austenite. We can notice that the residual strain at 0 load is basically null, which is due to the previous stabilization process. In the low-temperature curves, the initial elastic part is not present because of the detwinning of the Martensite, we cannot however observe a full transformation into detwinned Martensite since it would occur at a stress higher than the SMA yielding stress.

4.3.4 Strain recovery measurements

To test the functional behavior of the wire, we then executed a series of strain-recovery measurements, in which we heated the SMA wire from room temperature to above A_F and then let it cool back at the same controlled rate of 10°C/min to the initial temperature, completing the cycle; we executed various cycles under different stresses, which can be seen in Figure 27.

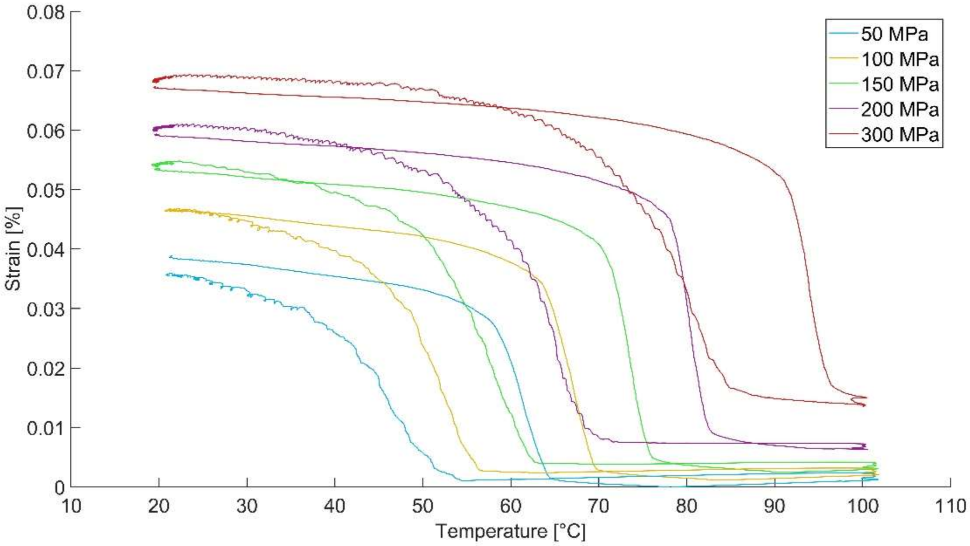


Figure 26 - Strain recovery curves at different stresses

From these curves we can clearly observe the flag shape of the cycle, in which the starting temperature of one transformation differs from the finishing one of the other one, the area comprised between the two branched paths of the cycle represents the energy lost because of the hysteresis.

The first path followed by the SMA is the heating path, the upper one, once the temperature of 100 °C is reached, the cooling path is then followed.

In both these paths we can recognize the phase change by looking at where the path stops being linear and instead becomes sigmoid-shaped.

We can also see that the starting and final temperatures of the transformations increase as the stress at which the cycle is executed increases; this effect is modeled through the Clausius-Clapeyron coefficients and is the same showed in Figure 16.

4.4 SMA parameters identification

Here we present the results of the tests described in the previous subchapter, the wire which we treated and on which we carried out all the measurements is made of $NiTiCu_6$ and has the same dimensions as the one specified in subchapter 1.6.3.

4.4.1 DSC results

Figure 27 displays the measured DSC curve for our wire.

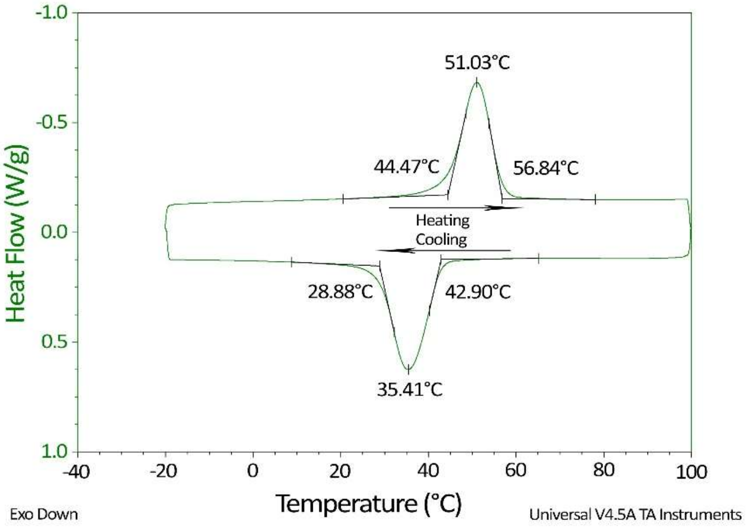


Figure 27 - DSC results for our SMA wire

We can recognize the austenitic transformation peak in the heating curve between 44.47°C and 56.84°C and in the cooling curve we can see the martensitic transformation between 42.90°C and 28.88°C. These temperatures are taken at the intersection between the tangents of the curves inside and outside the peak, conducted from the curve inflection point; they give us a first estimation of the transformation temperatures, the actual ones we will be using in the model will be optimized by simulating the system behavior and fitting the measured curve with the simulated one.

4.4.2 Stress/strain measurements results

The results of the isothermal tensile measurements were used to identify the mechanical parameters of the SMA wire through graphical interpolation, as seen schematically in Figure 28.

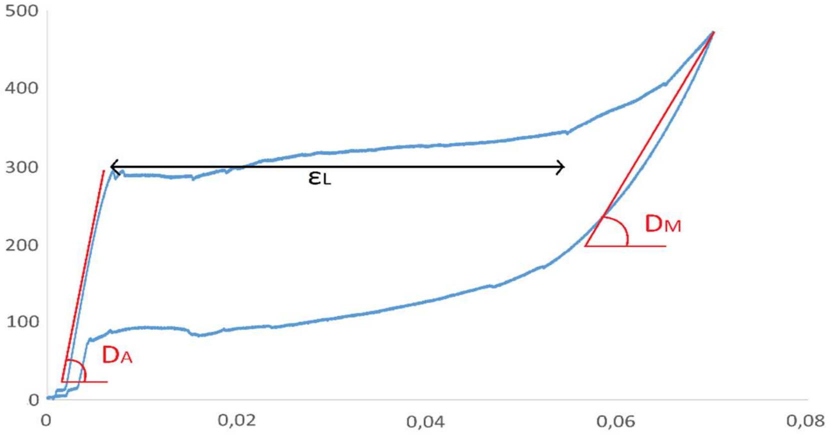


Figure 28 - Parameter estimation from stress/strain graph

Firstly, we estimated the maximum residual strain ε_L as the length of the central plateau in terms of strain, this graphical interpolation can be tricky, but the point of this measures is giving a first estimate of the selected parameter, which in this case is estimated to be around 0.055.

Then, we estimated the Young's moduli of Austenite and Martensite by looking at the slope of the linear parts at low and high stresses respectively, we estimated E_A = 56 MPa and E_M = 10 MPa.

4.4.3 Strain recovery measurements results

From these curves we graphically extrapolated the starting and finishing temperatures of each phase changes, MF in blue, MS in purple, AS in orange and AF in red, they can be seen in Figure 29.
Based on these temperatures we estimated the Clausius-Clapeyron coefficients, which were CA = 7 MPa/°C and CM = 6.5 MPa/°C.

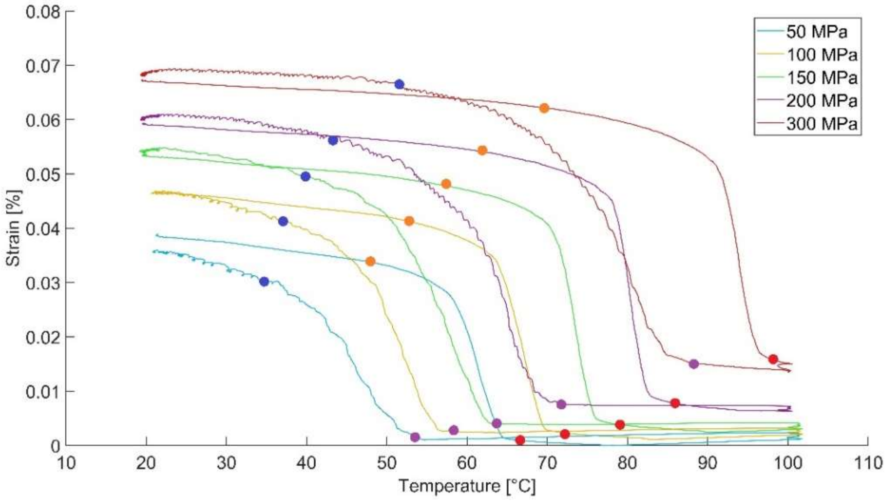


Figure 29 - Strain recovery curves with highlighted transformation temperatures

4.5 SMA wire tabulated parameters

The identified parameters to be employed for the SMA model are summarized here. In Table 1 we can see the results of the SMA wire characterization. In are the SMA wire parameters that can be found in the literature, in particular from (Brinson, 1993) and (Mattey, s.d.).

Table 1 - SMA wire estimated parameters

Parameter	**Value**
MF	28.88 °C
MS	42.90 °C
AS	44.47 °C
AF	56.84 °C
εL	0.055
DA	59.78 GPa
DM	17.92 GPa
σS^{cr}	30 MPa
σF^{cr}	180 MPa
CA	7 MPa /°C
CM	6.5 MPa /°C

Table 2 - SMA tabulated parameters

Parameter	**Value**
d_{NiTi}	6540 kg/m^3
ϱ_A	1 µΩ*m
ϱ_M	0.8 µΩ*m
c_1	836.8 J*Kg/K
Θ	0.55 MPa/°C
L_h	24.184 kJ/kg

4.6 Silicone characteristics

Regarding the silicone rubber, we decided to use parameters found in the literature without loss of validity for our system since these parameters are from a renowned source and they will be used only during simulations, any future work can use silicone parameters suitable to their own application. In Table 3 we see the parameters we used, taken from (materials, s.d.).

Table 3 - Silicone tabulated parameters

Parameter	**Value**
d_{Si}	1100 kg/m^3
c_2	1300 J*Kg/K
k_{Si}	0.2 W/m/K

5 Simulated actuator model

In this chapter we present the complete system equations used for the SMA wire model. The same basic equations, with appropriate modifications will also be relevant for the implementation of the Extended Kalman Filter, described in section 7.4 and in appendix 0.

5.1 Modified Brinson model for the SMA

During preliminary simulation tests, we found some differences between the simulated and real wire behaviors, thus some changes were made to the Brinson model to account for unmodelled effects.

5.1.1 TWSME modeling

One modification was the explicit introduction of the TWSME in the equations. Due to repeated thermomechanical actuation, NiTiCu often develops this effect. Hence, when the martensitic transformation occurs at low stresses, we have the formation of a fraction of stress-induced martensite ξ_S with respect to an untrained SMA.
To model this effect, in the affected zones we first calculate auxiliary martensitic fractions ξ'_S and ξ'_T without the TWSME, then we account for the stress-induced martensitic fraction created by the TSWME by multiplying ξ'_T by 1-χ, obtaining ξ_T; χ is called the TWSME retransformation fraction and is equal to 0.8,
Finally, we add to ξ'_S what was taken away from ξ'_T because of the TWSME, obtaining the actual fractions ξ_S.

5.1.2 Accounting for partial cycles

Some discontinuities problems were present with partial cycles: when the transformation is stopped in the middle and then resumed the resulting martensitic fraction may be discontinuous.
This is the case since the argument of the cosine function is supposed to be 1 (and thus the cosine is 0) at the entering border of the zone only, which means that when the transformations starts in the middle of a zone the cosine function is not null at the starting instant of the transformation. To solve this, we introduced the provision that the system updates the starting temperature T_i and stress σ_i at the beginning of the transformation and included those variables in the equations in place of the fixed starting points.

5.1.3 Creation of new zones

As equations 2.15, 2.16 and 2.17 accounts for a region which is difficult to model by a single, albeit conditional, equation, we divided the σ-T plane into 5 main zones where the phase change happens, plus a neutral zone where the change does not occur. Of the main zones, zone 5, remained quite difficult to handle as a single domain because there the laws governing the transformation depend on two independent conditions on the differentials of T and σ. The proposed solution that we implemented was to subdivide zone 5 into 3 subzones according to different combinations of conditions on the signs of $\dot{T}$ and $\dot{\sigma}$, corresponding to different zone-traversing paths.
The most interesting equations for the martensitic fraction transformation are the ones for zone 1 and 2. Zones 3, 4 and 5 treat cases bear a lesser interest for our application purposes.

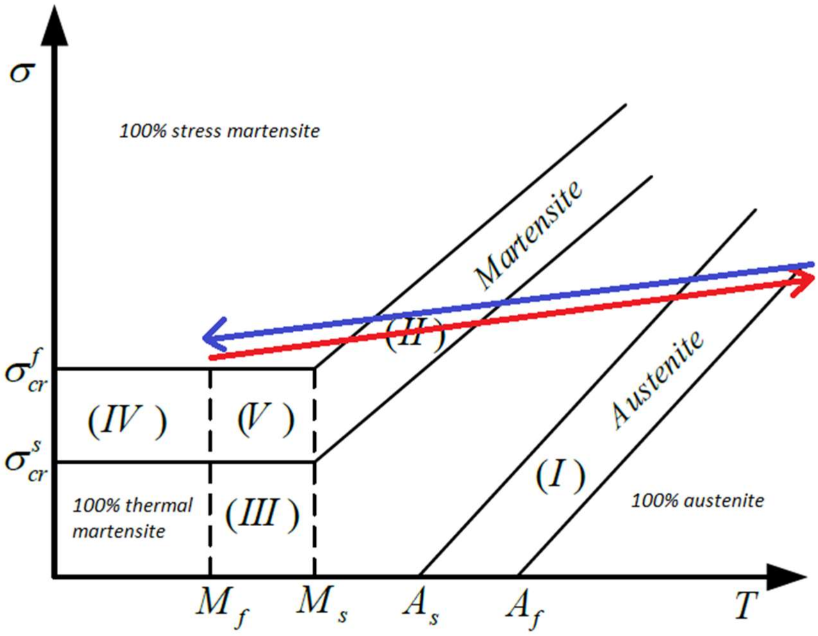


Figure 30 - Expected SMA working cycle in the phase plane

The arrows in Figure 30 highlight the typical transformation paths of a SMA wire in actuators:

- the red one represents the path of heating and increasing of stress, with the formation of austenite and the shortening of the wire.
- the blue one represents the path of cooling and decreasing of stress, with the formation of martensite and the elongation of the wire.

The laws governing the martensitic fraction in each zone are as follows:

- Zone 0: No transformation;

$$\xi = \xi_i \tag{5.1}$$

- Zone 1: Austenite zone;

$$if\ \left[\left(A_S < T - \frac{\sigma}{C_A} < A_F\right) \wedge \left(\dot{T} - \frac{\dot{\sigma}}{C_A} > 0\right)\right] then$$

$$\xi = \frac{\xi_i}{2}\left\{cos\left[a_A\left(T - T_i - \frac{\sigma}{C_A}\right)\right] + 1\right\} \tag{5.2}$$

$$\xi_S = \frac{\xi_{S_i}}{\xi_i}\xi \tag{5.3}$$

$$\xi_T = \frac{\xi_{T_i}}{\xi_i}\xi \tag{5.4}$$

- Zone 2: Linear stress Martensite zone;

$$if\ \left[(T > M_S) \wedge ({\sigma_S}^{cr} < \sigma - C_M(T - M_S) < {\sigma_F}^{cr}) \wedge \left(\dot{\sigma} - C_M\dot{T} < 0\right)\right] then$$

$$\xi_S = \frac{1 + \xi_{S_i}}{2} + \frac{1 - \xi_{S_i}}{2} cos\{a_M[\sigma - {\sigma_F}^{cr} - C_M(T - M_S)]\} \tag{5.5}$$

$$\xi_T = \frac{\xi_{T_i}}{1-\xi_{S_i}}(1-\xi_S) \tag{5.6}$$

- Zone 3: Thermal Martensite zone;

$if\ \left[\left(\dot{T} \leq 0\right) \wedge (M_S < T < M_F) \wedge (\sigma < \sigma_S{}^{cr})\right]\ then$

$$\xi'_S = \xi_{S_i} \tag{5.7}$$

$$\xi'_T = \frac{\left(1-\xi_{S_i}\right)+\xi_{T_i}}{2} + \frac{\left(1-\xi_{S_i}\right)-\xi_{T_i}}{2} cos[a_{TM}(T-M_F)] \tag{5.8}$$

$$\xi_S = \xi'_S + \chi\xi'_T \tag{5.9}$$

$$\xi_T = \xi'_T(1-\chi) \tag{5.10}$$

- Zone 4: Constant stress Martensite zone;

$if\ [(T_1 < M_F) \wedge (\sigma_S{}^{cr} < \sigma < \sigma_F{}^{cr}) \wedge (\dot{\sigma} \geq 0)]\ then$

$$\xi_S = \frac{1+\xi_{S_i}}{2} + \frac{1-\xi_{S_i}}{2} cos[a_\Sigma(\sigma-\sigma_F{}^{cr})] \tag{5.11}$$

$$\xi_T = \frac{\xi_{T_i}}{1-\xi_{S_i}}(1-\xi_S) \tag{5.12}$$

- Zone 5: Overlap Martensite zone. This zone is peculiar since the equations depend also on the combination of the temperature and stress derivatives

$if\ [(M_S < T < M_F) \wedge (\sigma_S{}^{cr} < \sigma < \sigma_F{}^{cr})]\ then$

  - Zone 5′: cooling and loading

    $if\ \left[\left(\dot{T}_1 < 0\right) \wedge (\dot{\sigma} > 0)\right]\ then$

$$\xi'_S = \frac{1+\xi_{S_i}}{2} + \frac{1-\xi_{S_i}}{2} cos[a_\Sigma(\sigma-\sigma_F{}^{cr})] \tag{5.13}$$

$$\xi'_T = \frac{(1-\xi_S)+\xi_{T_i}}{2} + \frac{(1-\xi_S)-\xi_{T_i}}{2} cos[a_{TM}(T-M_F)] \tag{5.14}$$

$$\xi_S = \xi'_S + \chi\xi'_T \tag{5.15}$$

$$\xi_T = \xi'_T(1-\chi) \tag{5.16}$$

  - Zone 5′′: heating and loading

    $if\ \left[\left(\dot{T} \geq 0\right) \wedge (\dot{\sigma} > 0)\right]\ then$

$$\xi'_S = \frac{1+\xi_{S_i}}{2} + \frac{1-\xi_{S_i}}{2} cos[a_\Sigma(\sigma-\sigma_F{}^{cr})] \tag{5.17}$$

$$\xi'_T = \frac{\xi_{T_i}}{1-\xi_{S_i}}(1-\xi_S) \tag{5.18}$$

$$\xi_S = \xi'_S + \chi\xi'_T \tag{5.19}$$

$$\xi_T = \xi'_T(1-\chi) \tag{5.20}$$

- Zone 5′′′: cooling and unloading

  $if\ \left[\left(\dot{T}<0\right)\wedge\left(\dot{\sigma}\le 0\right)\right]\ then$

$$\xi'_S = \xi_{S_i} \tag{5.21}$$

$$\xi'_T = \frac{\left(1-\xi_{S_i}\right)+\xi_{T_i}}{2} + \frac{\left(1-\xi_{S_i}\right)-\xi_{T_i}}{2} cos\left[a_{TM}(T-M_F)\right] \tag{5.22}$$

$$\xi_S = \xi'_S + \chi\xi'_T \tag{5.23}$$

$$\xi_T = \xi'_T(1-\chi) \tag{5.24}$$

Zone 5′′′ is functionally equal to zone 3, but without the two-way memory effect.
The parameters appearing in the equations are:

$$a_A = \frac{\pi}{A_F - T_i + \frac{\sigma_i}{C_A}} \tag{5.25}$$

$$a_M = \frac{\pi}{\sigma_i - \sigma_{CR}^F - C_M(T_i - M_S)} \tag{5.26}$$

$$a_{TM} = \frac{\pi}{T_i - M_F} \tag{5.27}$$

$$a_\Sigma = \frac{\pi}{\sigma_i - \sigma_{CR}^F} \tag{5.28}$$

5.2 Silicone heat transfer model

The heat transfer in the silicon coating can be fully described by solving the Fourier heat equation; this equation is a partial differential equation including spatial and time derivatives.
However, for the scope of our modeling we are not interested in the temperature distribution inside the silicon but only in the overall contribution of the presence of the silicon in the heat transfer from the wire to the environment, thus we decided to use a lumped parameters approach.
We based our model on the work by F. Scarpa and M. De Rosa (Scarpa & De Rosa, 2016) on the heat conduction in insulated electric wires, since our situation is very similar having a metal wire coated by an electrically and thermically isolating material, the one big difference being the inclusion of a latent heat term L_h which depends on the derivative of the martensitic fraction and accounts for the energy absorbed by the phase change in the SMA. This necessary modification to the equations we did implement.
The two main state variables of the thermal model are T_1, the wire temperature, which is assumed to be isothermal, and T_2, the silicone average temperature. The use of the average temperature instead of the temperature distribution is a valid simplification under the condition that the Biot number $Bi = \frac{hL}{k}$ is smaller than 0.1, which is met for our cable dimensions (Bi = 0.0662). The lumped parameter heat transfer equations are:

$$m_1 c_1 \dot{T}_1 = \frac{V^2}{R_{el}(\xi)} - \frac{T_1 - T_2}{R_{t1}} + m_1 L_h \dot{\xi} \tag{5.29}$$

$$m_2c_2\dot{T}_2 = \frac{T_1 - T_2}{R_{t1}} - \frac{T_2 - T_\infty}{R_{t2}} \quad 5.30$$

where m_1 and m_2 are the masses of the wire and insulation, c_1 and c_2 are their respective specific heat capacities.

R_{el} is the electric resistance of the wire, which depends on the strain and martensitic fraction.

$$R_{el}(\xi) = \frac{L_0\big(1+\varepsilon(\xi)\big)}{A_0\big(1-\nu_{NiTi}\varepsilon(\xi)\big)^2}[\rho_A + \xi(\rho_M - \rho_A)] \quad 5.31$$

Where ρ_A and ρ_M are the Austenite and Martensite electrical resistivities, L_0 is the initial wire length and ν_{NiTi} is the SMA wire Poisson, which is equal to 0.33, and A_0 is the initial wire area.

r_{t1} and r_{t2} are the thermal resistivities per unit length of SMA wire and silicone respectively, R_{t1} and R_{t2} are their thermal resistances.

$$r_{t1} = \frac{ln\frac{a_2}{a_1}}{2\pi k_{Si}} \quad 5.32$$

$$r_{t2} = \frac{1}{2\pi a_2 h} \quad 5.33$$

$$R_{t1} = \frac{r_{t1}}{L_0} \quad 5.34$$

$$R_{t2} = \frac{r_{t2}}{L_0} \quad 5.35$$

k_{Si} is the thermal conductivity of the silicone, ε is the strain of the wire and a_1 and a_2 are the wire and cable radii.

5.3 Model implementation

In the present section we discuss the implementation and coding of the model into MATLAB and Simulink.
Regarding Simulink files, we preferred to use Simulink blocks only to code equations instead of MATLAB scripts, in order to speed up the software execution.

5.3.1 SMA wire model states and inputs

Our SMA wire model is described by both dynamic and algebraic equations.
The state variables of the dynamic equations are:

- T_1 SMA wire average temperature;
- T_2 Silicon average temperature;

While the algebraic ones are:

- ξ_S Stress induced Martensite fraction;
- ξ_T Thermal induced Martensite fraction;
- ξ Total Martensite fraction;
- ε Wire strain;
- R_{el} SMA wire electrical resistance.

The control variable is the voltage V.
The unknown parameters and disturbances to be estimated are:

- σ Wire stress;
- h Convection coefficient;

The measured variable of the system is the wire strain ε.

5.3.2 SMA model blocks

Here we present in detail all of the blocks present in our Simulink model, whose general layout is schematized in Figure 31.
The simulations are executed with a time step t_s of 0.005s.

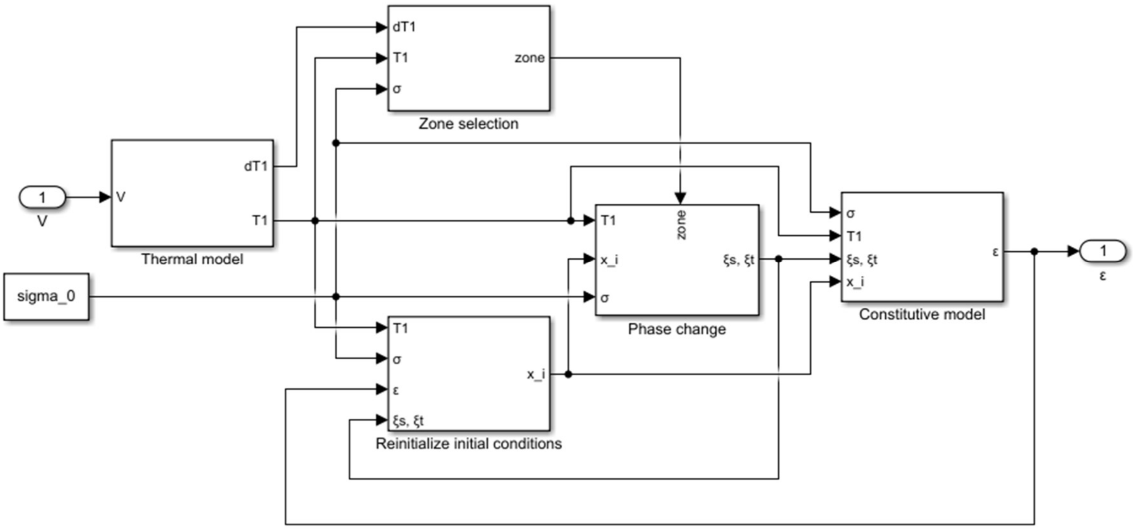


Figure 31 - SMA wire Simulink model layout

5.3.3 Zone selection block

This block determines the current zone of the σ-T plane based on the current information on the stress, the wire temperature T_1 and their derivatives.
It outputs one Boolean signal for each zone, 1 if we are inside that zone, 0 otherwise, this information is then fed to the constitutive model equations block, which change accordingly.
The conditions and the equations for each zone are specified in subchapter 5.1.3.

5.3.4 Reset initial value block

This block resets the initial values of the system variables at the beginning of the new zone (the variables with subscript *i*) whenever a zone signal changes from 1 to 0 or vice versa.
When a commutation happens, the block updates the old initial value with the value of that variable at the previous time instant, the block holds this value and outputs it constantly until the next zone commutation.
The variables which need their value to be stored at the end of the zone are T_1, σ, ξ_S, ξ_T and ε.

5.3.5 Thermal model block

Inside this block we coded the dynamical thermal model of the system using equations 5.22 and 5.23 for the wire and silicone temperatures T_1 and T_2 from subchapter ◎, , plus all the equations from 5.24 to 5.28, which defined the parameters of the thermal model.
This block also updates some state-variable-dependent model parameters with the current value of the governing variables, such as the electric resistance R_{el}, which is updated to the current ξ and ε.
Here is also present the derivation of the current ξ for the computation of the latent heat of transformation term. This derivative is low-pass filtered before use to avoid unwanted noise or chattering problems.

5.3.6 Phase transformation block

This is one of the most important and complex blocks in the model, as it contains all the equations and zones from subchapter 5.1.3 to compute the current ξ_S and ξ_T, each contained in a specific subsystem, which is enabled by the respective zone Boolean signal; all the outputs of these blocks are then connected to the same Merge block, that selects only the output of the subsystem active at the moment (i.e. the output of the equation currently governing the transformation).

5.3.7 Constitutive model block

Compared to the other blocks in the model this one is quite straightforward; it computes the strain ε according to equation 2.9, which has not been modified by us and is mostly a linear combination of the other state variables except for the division by the total wire Young's modulus D updated to the current ξ.
Another small block outside of the physical system model computes the current wire length using the strain.

5.4 Model parameter optimization

In this section we present the preliminary simulations we ran to optimize the model parameters so that the model would mimic as much as possible the behavior of the real SMA wire, especially in our operating zone of interest, and the results of said simulations.

5.4.1 Test protocols for the model

In order to test the model accuracy, we need to test its behavior and compare it to that of a real wire with the same parameters.

The first test aims to verify the complete operating cycle by applying a voltage pulse to the SMA at room temperature, letting it cycle between 100% Austenite and back to 100% Martensite by heating it above A_F and then letting it cool below M_F; we also compare the model behavior to the measurements obtained from a real wire.

The duration of the step is 100s and the voltage is 0.8V.

5.4.2 Stress recovery curve fitting

We used the stress recovery curves discussed in 4.3.4 to compare the performance of our model to the behavior of the real wire, we executed many transformation cycles, toggling the model parameters one at the time to better approximate the real wire behavior, in Figure 31 we can see the comparison between our simulated curves, the dotted lines, and the real wire behavior, the continuous lines.

These simulations are referred to complete heating cycles.

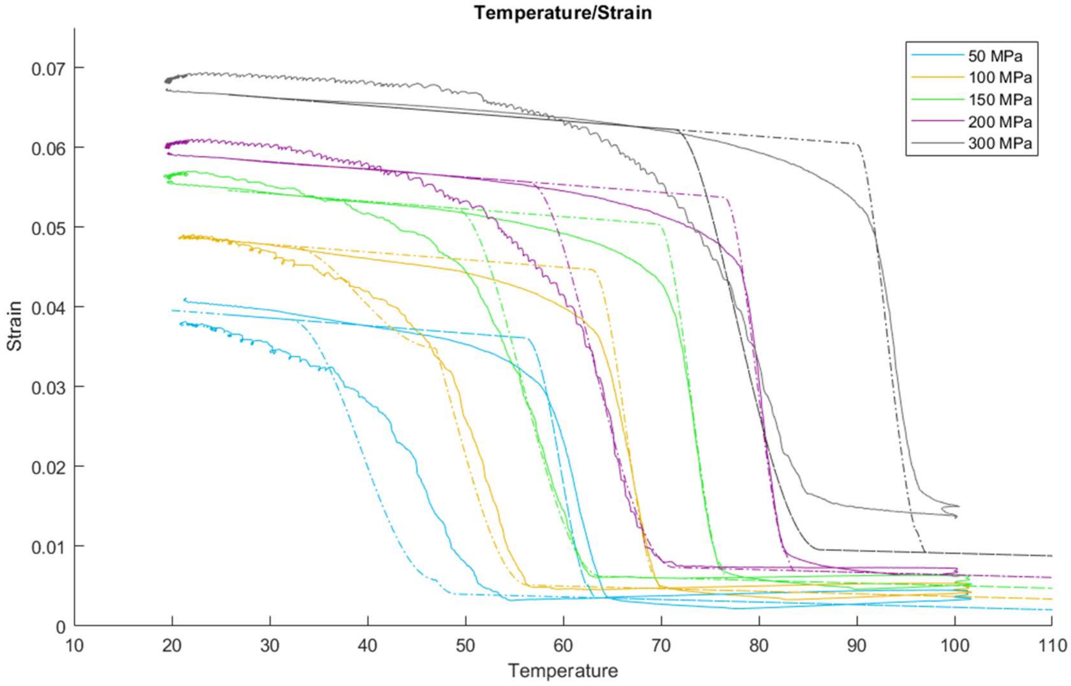


Figure 32 - Comparison between real and simulated strain recovery curves

It can be seen that for medium stresses (from 100 to 200 MPa) and below the maximum training stress of 3%, we have a good accordance between the simulated and real behavior, this will be the operating range for our actuator and we are most interested in optimizing the model behavior here through the proper choice of model parameters.

In the 50 MPa curve, at low temperatures we have some divergence between the simulated and real behavior, this is most likely due to an incomplete modeling of the TWSME.

Another divergence is present in the 300 MPa curve at high temperatures, where our model simulates a recovered strain up to 1%, when in reality this stress is too high and the real wire can recover only up to 1.5% of strain, thus our actuator must be used at lower stresses.

Furthermore, in all curves there are some discrepancies at the beginning of the austenitic transformation and at the end of the martensitic one, this is because of the possible presence of some orthorhombic phase, however these discrepancies are present at strains higher than our maximum working strain of 3%, this means that they will have a limited effect on our model in the normal operating conditions.

5.4.3 Conclusion on model parameters optimization

After the parameter optimization process the behavior of our system is very similar to the one of the real wire, especially in our operating zone of interest. The final values of the model parameters are shown in Table 4.

Parameter	Value
M_F	32.08 °C
M_S	47 °C
A_S	49.4 °C
A_F	56.6 °C
ε_L	0.03953
D_A	37 GPa
D_M	12.5 GPa
σ_S^{cr}	36.7 MPa
σ_F^{cr}	136.7 MPa
C_A	7.4 MPa/°C
C_M	6.7 MPa/°C

Table 4 - Optimized values of the estimated model parameters

From this point on, we will only use our simulated model as a reference for performance tests, instead of the real wire.

5.4.4 Silicone effect on the temperature

We conducted some tests to verify the effect of the presence of the silicone layer on the wire temperature by comparing the simulated heating of a naked wire and another one coated with 1 mm of silicone rubber until the end of the Austenite transformation.

The stress was set at 200 MPa, the wires were heated to the same target temperature of A_F + 4 °C, adjusted by the stress according to the Clausius-Clapeyron's relationship. The tests were conducted at two different convective dissipation coefficients, 20 W/m^{2}K and 5 W/m^{2}K.

The two simulated temperatures of the first test are shown in Figure 33.

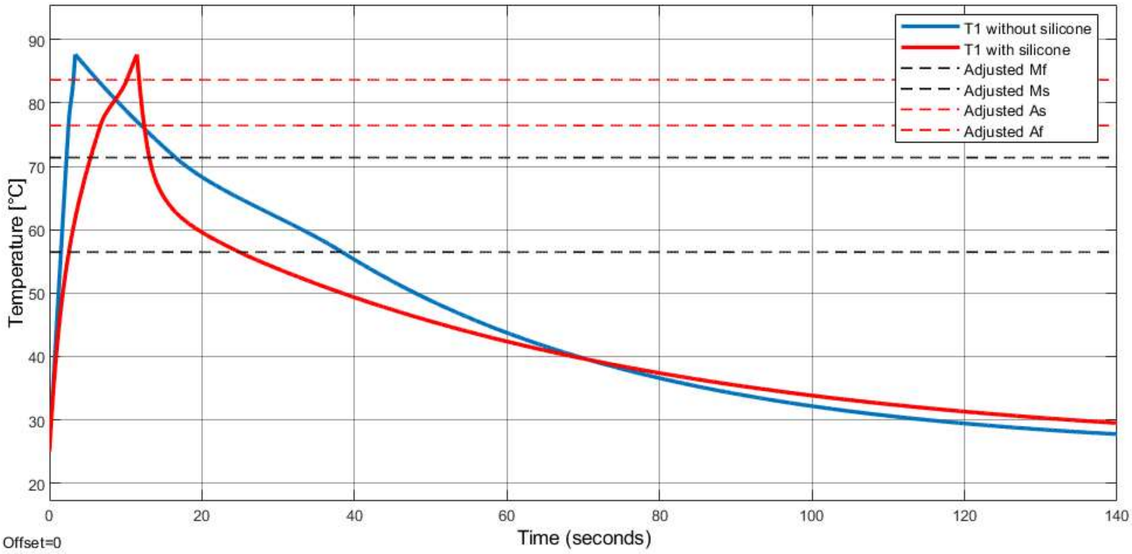


Figure 33 – Comparison between naked wire temperature (blue) and coated wire temperature (red) at h = 20 W/m^{2}K

We see that the presence of the silicone rubber slows down the heating but greatly helps the initial part of the cooling phase, this is because the cooling is helped by the lower silicone temperature T_2, as shown in Figure 34.

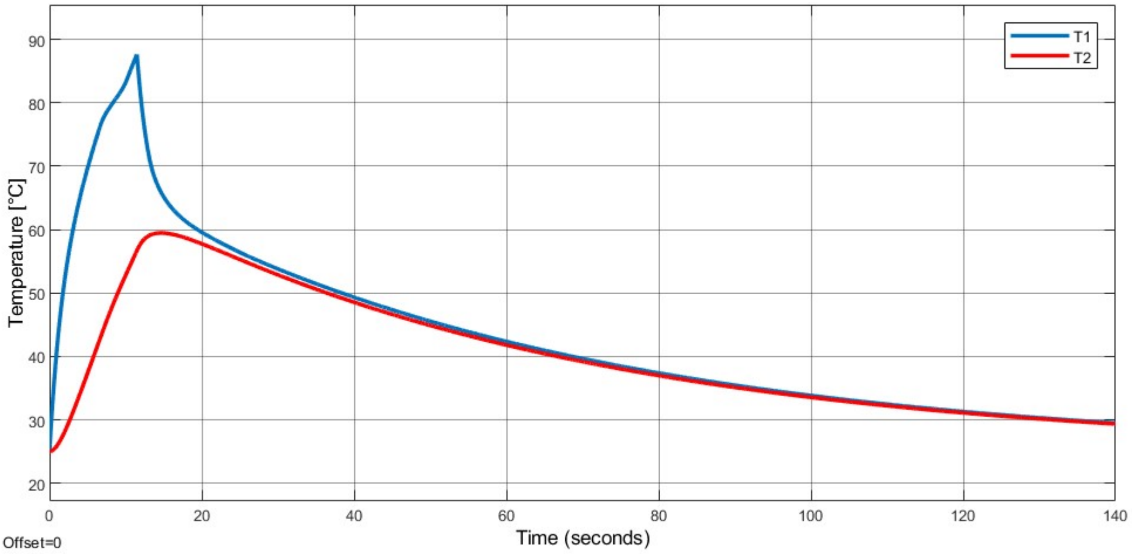


Figure 34 – Wire temperature (blue) and silicone temperature (red) at h = 20 W/m²K

Here we clearly see the initial cooling phase, in which the temperature difference between wire and silicone is quickly recovered. It can be anticipated that, an appropriate control strategy that does not allow the silicone to heat excessively in repeated cycles, could safeguard its cooling action, which, in turn, should be regarded as a means of both shortening the cycle duration and improving the controllability, particularly during cooling.

In Figure 35 we see the electric power fed to the system.

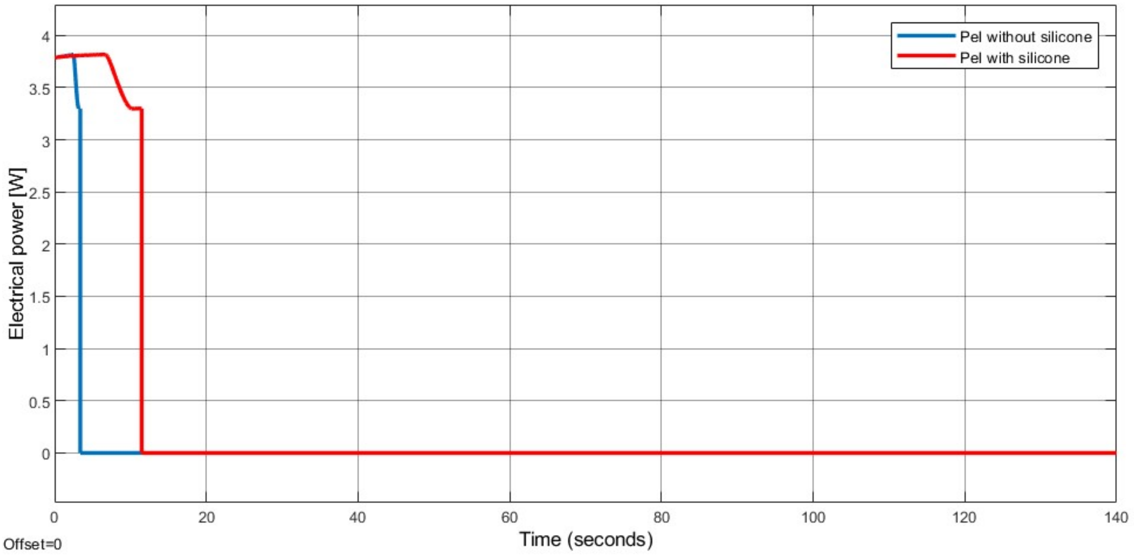


Figure 35 – Electrical power to reach $A_F(\sigma) + 4$ for naked wire (blue) and coated wire (red) at h = 20 W/m²K

Predictably, the electric power to heat up the wire in the presence of the silicone rubber thermal capacitance in direct contact with it is larger.

In Figure 36, Figure 37 and Figure 38 we see the results of the test conducted at h = 5 W/m²K.

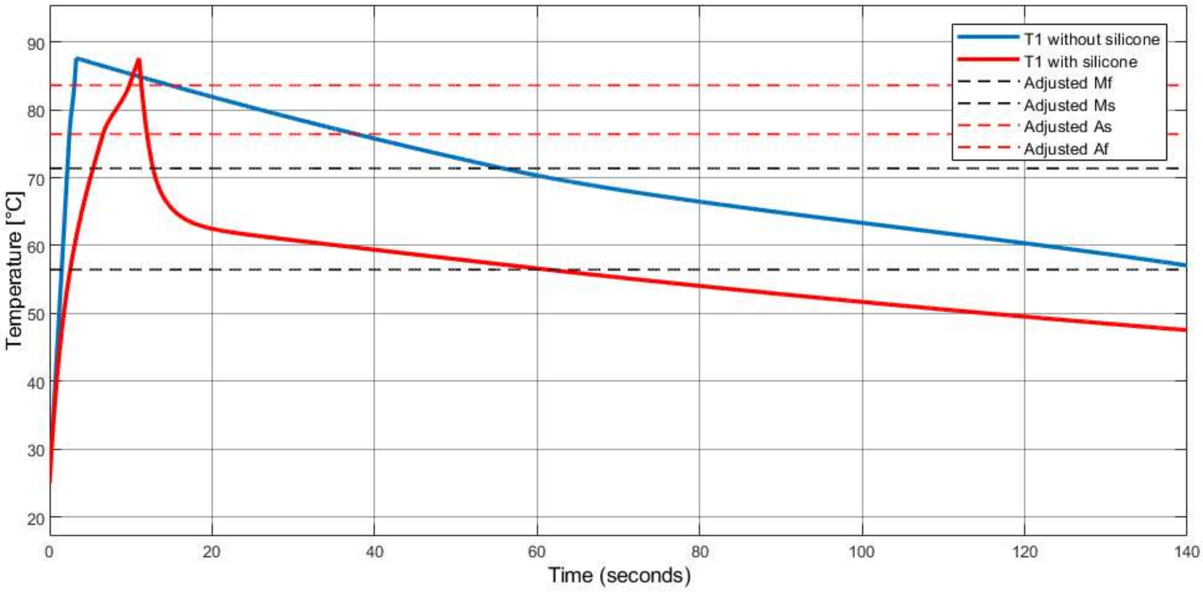


Figure 36 – Comparison between naked wire temperature (blue) and coated wire temperature (red) at h = 5 W/m²K

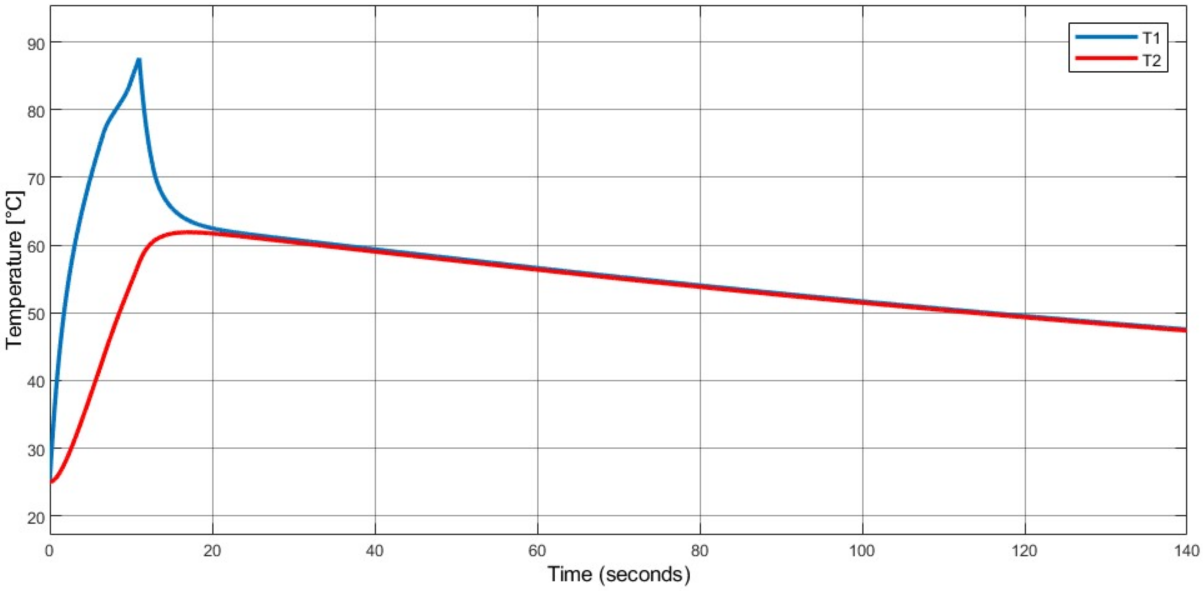


Figure 37 – Wire temperature (blue) and silicone temperature (red) at h = 5 W/m²K

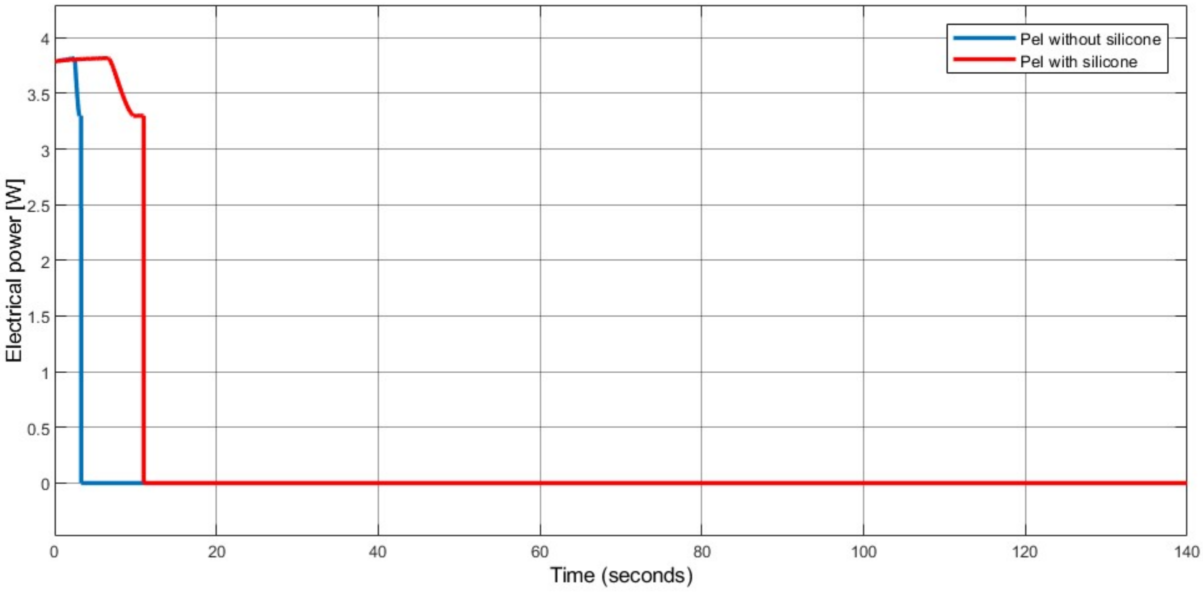


Figure 38 – Electrical power to reach $A_F(\sigma) + 4$ for naked wire (blue) and coated wire (red) at h = 5 W/m^2K

We can observe that with a lower dissipation the system cools off slowly but we can still see the initial rapid cooling of the wire temperature in Figure 39.

6 Concept of a variable PID controller

In this chapter we explain the general concept of the proposed controller, introducing the use of an EKF and of a zone selection scheme to modulate the PID parameters.

6.1 Proposed controller

A proposed controller which uses the estimated state variables from the EKF could consists of a variable-structure PID, which structure and gains change based on the current control zone; these could reflect the conditional and hysteretic behavior of the SMA and would allow us to adjust the control action to the changing SMA response. The current control zone is determined like in the simulated SMA model, where a logic block detects when the current value of a combination of state variables passes a certain threshold, indeed the control zones are very similar to the ones introduced in subchapter5.1.3.
The nominal operating range of this controller is with stresses between 200 MPa and 300 MPa and for strains between 1% and 3%.
The state variables T_1 and T_2 and parameter σ used to determine the current zone, as well as convective coefficient h, are estimated by an Extended Kalman Filter (EKF).
We also want to take into account the energy present in the silicon, which can help the controller reduce the electric power when heating the wire.

6.1.1 General controller layout

In Figure 40 we can see the most important parts of the layout of our improved controller.
The first block we focus on is the EKF. This block takes as input the same input fed to the SMA wire model (V), its output (ε) and the current zone. Using this information, it produces an estimate of the unmeasured state variables of the model and then feeds them to the zone selection and the variable PID blocks.
The zone selection block uses the EKF state estimation to determine the current control zone, as its name implies.
Finally, the Variable PID block acts on the SMA wire model with a control action dependent on the current strain error by some coefficients which change depending on the estimated control zone, it also adjusts the control action with respect to the temperature T_2 and thus the energy stored in the silicone rubber.

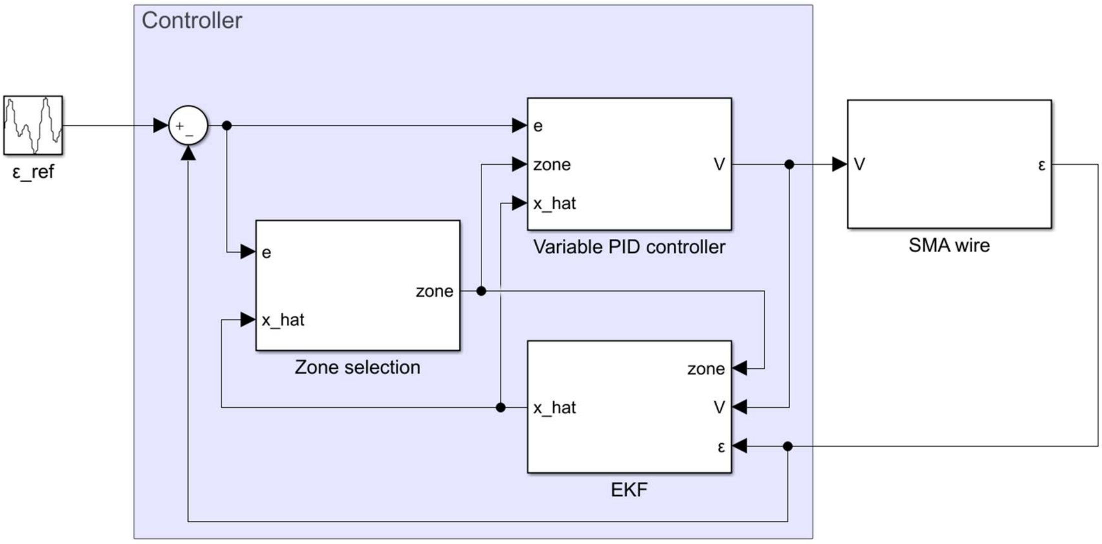


Figure 40 - Controller layout

In this thesis the main focus is on the EKF and the zone selection blocks, while only preliminary results will be provided for the implementation of the variable PID controller.

6.2 Reference controller

In order to assess the performance of the proposed variable PID controller, we used a PID controller with constant gains as a reference, the comparison will be seen chapter 0.
The same reference PID controller will also be used to control the wire model in the evaluation of the EKF performance, seen in chapter 8.
Its proportional gain was tuned on the heating phase by incrementing it until we reached an overshoot, at that point we halved it and we started incrementing the integral part, the integral of the error was also limited between -0.08 and 0 during the tuning phase, finding the best compromise between overshoot on heating and low steady state error, in the end its gains were $K_P = -200$, $K_I = -20$ and $K_D = 0$.
Its input is the strain error e and its output is the voltage V; the output is saturated with lower limit 0 since our system is heated via Joule effect, which means that a negative voltage will have the same effect as a positive one, and upper limit V_{SAT}, which is set at 1V. We will be comparing its response with the one of our proposed controller to check on our improvements.

7 Extended Kalman Filter (EKF)

In this chapter we focus on the EKF structure, parameters, and implementation.

7.1 Parameter estimation

Because of the system nonlinearities like hysteresis and branching, the estimation is done through the use of an EKF (rather than a standard Kalman Filter), to which we feed the current control input of the system, the last available strain measurement and the last available current zone of our controller.
The interesting variables estimated are the wire temperature T_1, the average silicon temperature T_2, the stress σ and the convective heat transfer coefficient h.
While T_1 and T_2 are dynamic variables which follow a known, albeit nonlinear, dynamic equation, σ and h are unknown parameters whose evolution is unmodelled and in case of h also stochastic; to estimate them we suppose them constant and noisy, leaving the EKF able to tune them.
Due to our operating conditions, which will be discussed in subchapter 4.3, we will only be working with stresses above σ_F^{cr} and thus we simplified the low temperature zones below this stress to simplify the EKF equations while still accounting for the zones outside our scope of interest.

7.2 EKF formulation

In the EKF, the state transition function f and the measurement function h are not required to be linear functions of the state but may instead be generic nonlinear differentiable functions.
The discrete model of the system takes the form

$$x_k = f(x_{k-1}, u_{k-1}) + w_{k-1} \tag{7.1}$$

$$y_k = h(x_k) + v_k \tag{7.2}$$

where x_k and x_{k-1} are the state vector at the current and previous time instant; u_k is the current input of the system; y_k is the system output; w_k are v_k are the process and measurement noise, assumed white gaussian noise with covariances Q and R.
The function f is used to compute the predicted state $\hat{x}_k$ from the previous state estimate. Similarly, the function h is used to compute the predicted measurement $\hat{y}_k$ from the predicted state.
However, f and h cannot be applied to the covariance directly, instead a matrix of partial derivatives, the Jacobian, is computed.

$$F_K = \left.\frac{\partial f}{\partial x}\right|_{\hat{x}_{k-1|k-1}, u_{k-1}} \tag{7.3}$$

$$H_K = \left.\frac{\partial h}{\partial x}\right|_{\hat{x}_{k|k-1}} \tag{7.4}$$

The Jacobian is evaluated with the current predicted states, this matrix can then be used in the Kalman filter equations.
This process is essentially a linearization of the nonlinear system functions around the current estimate.
In order to estimate the current state, at each time step, the EKF alternates two processes:

1) Prediction step
    a. Predict the current state estimate

$$\hat{x}_{k|k-1} = f\left(\hat{x}_{k-1|k-1}, u_{k-1}\right) \tag{7.5}$$

b. Predict the state prediction error covariance estimate P

$$P_{k|k-1} = F_k P_{k-1|k-1} {F_k}^T + Q \quad 7.6$$

2) Update step

a. Compute the output innovation z

$$\hat{z}_k = y_k - h(\hat{x}_{k|k-1}) \quad 7.7$$

b. Compute the innovation covariance S

$$S_k = H_k P_{k|k-1} + R \quad 7.8$$

c. Compute the near-optimal Kalman gain K

$$K_K = P_{k|k-1} {H_k}^T {S_k}^{-1} \quad 7.9$$

d. Update the state estimate

$$\hat{x}_{k|k} = \hat{x}_{k|k-1} + K_K \hat{z}_k \quad 7.10$$

e. Update the state prediction error covariance estimate

$$P_{k|k} = (\mathbb{I} - H_k K_K) P_{k|k-1} \quad 7.11$$

7.3 EKF zones

In this subchapter we define the zones used by the Extended Kalman filter, which will be identified with a letter, to distinguish them from those defined in subchapter 5.1.3, also because some zones are taken directly from that section, like zones a and b which are identical to zones 1 and 2.
Some zones fall outside of the working range for our controller, defined in subchapter 6.1, but are modeled anyway, in order to make our parameter estimation more robust with respect to estimation errors that could cause the variables to be outside the working range and in turn may cause severe computational errors.
The zones and their conditions in detail are:

- $if\left(A_S < T - \frac{\sigma}{C_A} < A_F\right) \wedge \left(\dot{T} - \frac{\dot{\sigma}}{C_A} > 0\right) then$

  Transformation to Austenite zone (zone a)

  In this zone the Austenite transformation occurs and so we have a contraction of the wire very fast and sensitive to the controller action.

  In Figure 41 we see in red the position of zone a in the σ-T plane and the direction of the shortest transformation paths, represented by the dark red arrows.

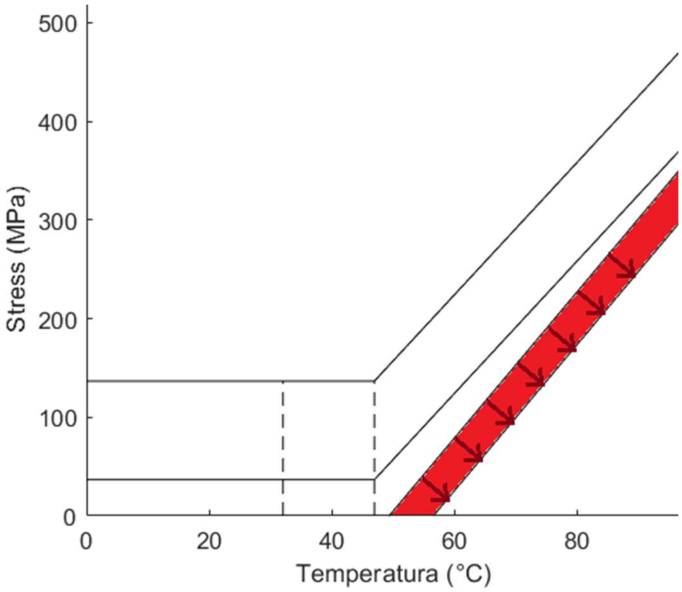


Figure 41 - Zone a in the σ-T plane

- $if\ \left[(T > M_S) \wedge ({\sigma_S}^{cr} < \sigma - C_M(T - M_S) < {\sigma_F}^{cr}) \wedge \left(\dot{\sigma} - C_M\dot{T} < 0\right)\right]\ then$

  Transformation to Stress-Induced Martensite zone (zone b)

  In this zone where we have a wire elongation due to the martensitic transformation that happens while cooling, shown in blue in Figure 42 , alongside the direction of the shortest transformation paths (arrows).

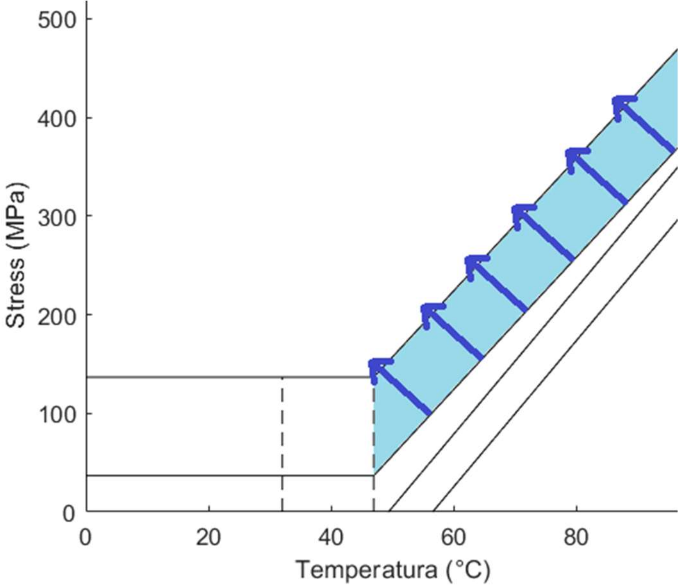


Figure 42 - Zone b in the σ-T plane

- $if\ \left[\left(\dot{T} \leq 0\right) \wedge (M_S < T < M_F) \wedge (\sigma < {\sigma_F}^{cr})\right]\ then$

  Transformation to Thermal-Induced Martensite zone (zone c)

  In this zone we have the creation of thermal Martensite when the temperature decreases, this zone is the result of merging zones 3 and 5′′′ from subchapter 5.1.3 and can be seen in green in Figure 43.

  We decided to merge the two zone for two main reasons, dependent on the fact that we expect to be working with constant stresses most of the time:

  - There is no difference between zones 3 and 5 if the stress is constant.
  - The presence of errors in the stress estimation in zone 5 would lead to a change in operating zone between all the subzones, which we can automatically exclude since we know the stress is not varying.

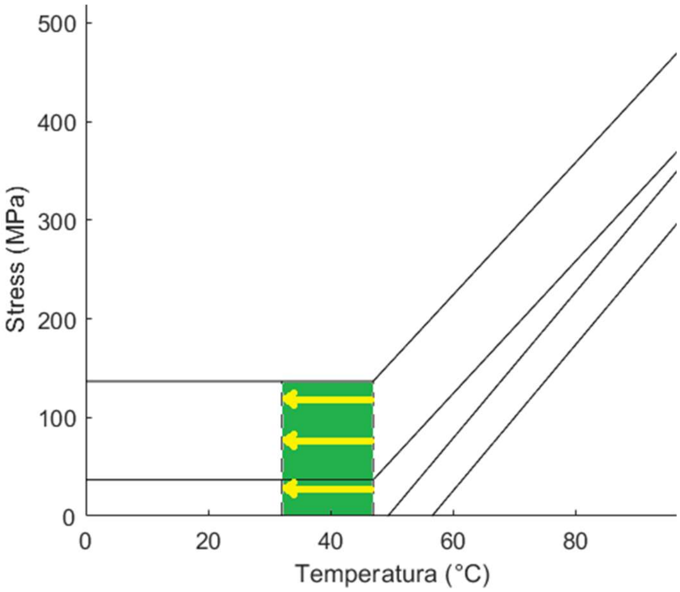


Figure 43 - Zone c in the σ-T plane

When the transformation conditions are not satisfied, we find ourselves in the neutral zones. In these zones there is no transformation, however the controller still needs different operating zones, according to the strain error (e) sign and the value of the martensitic fraction (ξ). The condition for the neutral zones is:

- $if\ (zonea = 0) \wedge (zoneb = 0) \wedge (zonec = 0) \wedge (e < 0) \wedge (\xi \neq 0)\ then$

  Free heating zone (zone d)

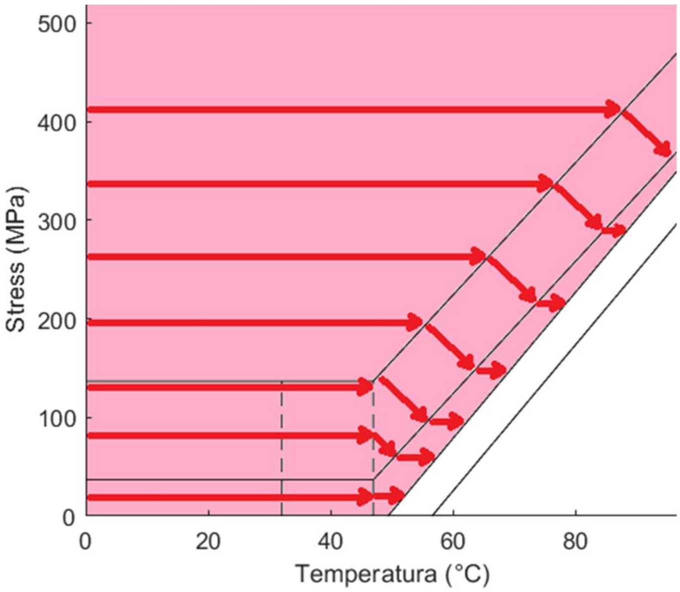


Figure 44 – Zone d in the σ-T plane

- $if\ (zonea = 0) \wedge (zoneb = 0) \wedge (zonec = 0) \wedge (e > 0) \wedge (\xi \neq 1)\ then$

  Free cooling zone (zone e)

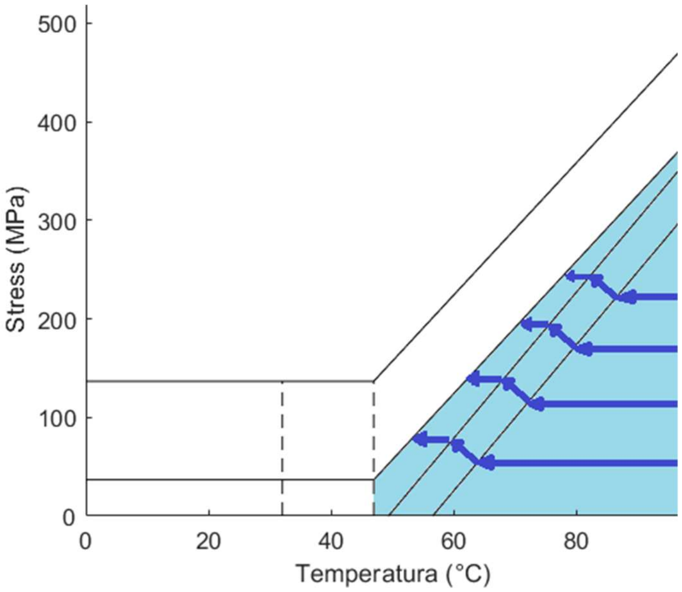


Figure 45 – Zone e in the σ-T plane

The last set of zones included are the zones after the complete transformations, both for Austenite and Martensite, when the reference strain has not been reached yet.

If not treated separately from the free zones, these could become problematic if the reference strain is unreachable due to the actuator limitations, they are:

- $if\ (\xi = 0) \wedge (e > 0)\ then$

  Austenite 100% zone (zone f)

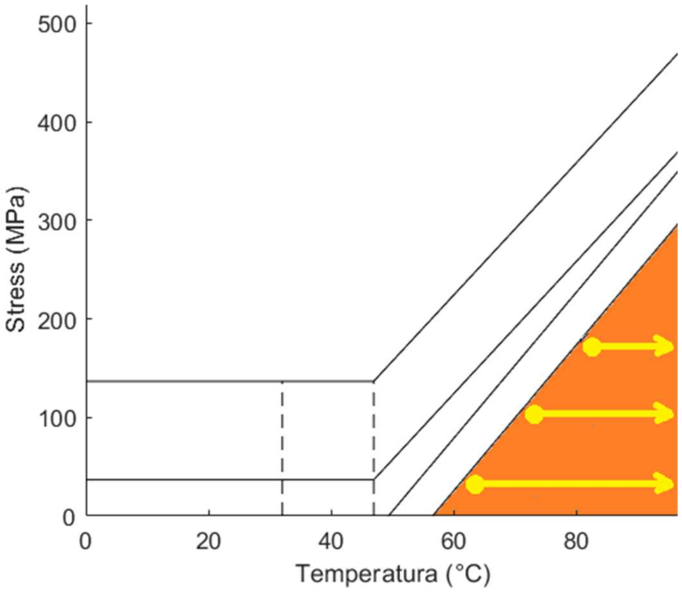


Figure 46 - Austenite 100% zone in the σ-T plane

- $if\ (\xi = 1) \wedge (e < 0)\ then$

  Martensite 100% zone (zone g)

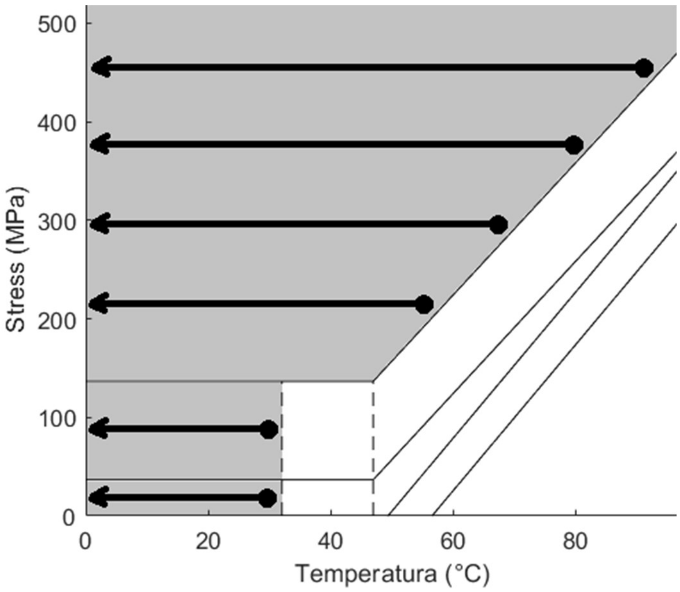


Figure 47 - Martensite 100% zone in the σ-T plane

7.4 EKF implementation

For the implementation of our state observer, we used the EKF block present in Simulink library. This block needs many parameters and inputs, in particular two MATLAB functions which represent the state transition function *f* and output transformation *h* present in the discrete-time formulation of the EKF, to avoid over encumbering this section all the equations of the model are presented in Appendix 0.

The Jacobian of *f* is computed numerically at each time step by Simulink.

These functions depend on the state vector, which in our case is formed by 6 proper state variables and 2 unknown parameters to be estimated, the state variables are:

- T_1 SMA wire average temperature;
- T_2 Silicon average temperature;
- ξ_S Stress induced Martensite fraction;
- ξ_T Thermal induced Martensite fraction;
- ξ Total Martensite fraction;
- ε Wire strain;
- R_{el} SMA wire electrical resistance.

The unknown parameters and disturbances to be estimated are:

- σ Wire stress;
- h Convection coefficient
- We also need to define the values of the process and measurement covariances matrices Q and R; whether the noise is addictive, like in the equations above, or *f* depends on the noise in some other way; and finally, the initial values for the state x_0 and the process covariance Q_0.

The control variable is the voltage V and the output is the wire strain ε, this means that while the state transition function is very complicated and nonlinear, the measurement function is linear and trivial.

7.5 Zone selection block

This logic block is the one responsible for determining the current control zone, as its name suggests, it outputs one Boolean signal for each zone which is true if we are inside it, false otherwise.

The current zone is then fed to both the variable PID block and the EKF.

It checks the conditions defined in subchapter 5.1.3 using two combinations of the state variables wire temperature T_1 and stress σ, plus the two variables themselves.

These linear combinations of variables are the Austenitic temperature T_{AUST} and the Martensitic stress σ_{MART}, and are defined as

$$T_{AUST} = T_1 - \frac{\sigma}{C_A} \tag{7.12}$$

$$\sigma_{MART} = \sigma - C_A(T_1 - M_S) \tag{7.13}$$

Their gradient represents the direction of fastest transformation in the σ-T plane in their respective zones; these gradients, as well as the wire temperature derivative, are computed inside the block and filtered through a low-pass filter with cutoff frequency 50Hz to reject the noise, which would cause the computed derivative to change its sign with every random fluctuation of the noisy signal, the presence of the filter however introduces a slight delay in the zone detection.

To make the zone detection more robust and to mitigate the delay in the zone selection, we introduced another logical condition on the sign of the measured strain derivative, whenever a change in sign happens, a logic block

holds the current zone for up to 0.1s after the commutation since we know that the system dynamic is much slower than the Simulink sample time and thus it is impossible to have multiple zone commutations in such a small time interval.

7.6 Test protocols for the EKF

In this section we will describe the waveforms and the conditions we selected to test the EKF performance. The purpose of the tests is to compare the simulated wire state variables (representing the ground truth) with the ones estimated by the EKF (to be used for the setting of the Variable PID coefficients).

All the tests will start from ambient temperature T_{amb} = 25 °C.

In order to mimic a realistic operating condition, a measurement white noise is added to the simulated output strain with mean 0 and variance $5x10^{-8}$, low pass filtered with cutoff frequency 10 Hz.

First, we want to test the EKF performance when estimating the model response when it is given a voltage pulse train in open loop, this will allow us to test multiple complete heating cycles with different convective heat exchange coefficients, to simulate different environments.

The stress on the wire at which these tests will be carried out is 200 MPa, the input voltage is 1V, the pulse train will have an initial delay of 60s while the pulse lengths will vary depending on the current convection coefficient h, we will test h = [5, 15, 35].

After this, we wanted to verify the performance of the EKF on a closed loop system, consisting of a PID controller and the simulated system, to investigate its performance on a real albeit simplified control case.

The controller will be tracking a random steps strain reference varying between 1% and 3%, each step is 120s long and there will be an initial delay of 120s.

We will carry out tests in the two extreme conditions for the convection with h = [5, 35].

Finally, we will test the performance of the EKF on the same closed loop system but giving it wrong initial conditions on stress and the convective coefficient.

The closed loop system will be subjected to cyclic loading, following a square wave strain reference, with initial delay 120s, period 120s, on time 40s, minimum strain value 1% and maximum strain value 3%.

8 EKF performance

Here we present the performance of our EKF, verified through the tests defined in section 7.6.

8.1 Estimations of open loop system with correct initial conditions

In this part we show the results of the estimations of variables with an exact knowledge of the initial conditions of the system.

8.1.1 Tests with h = 5

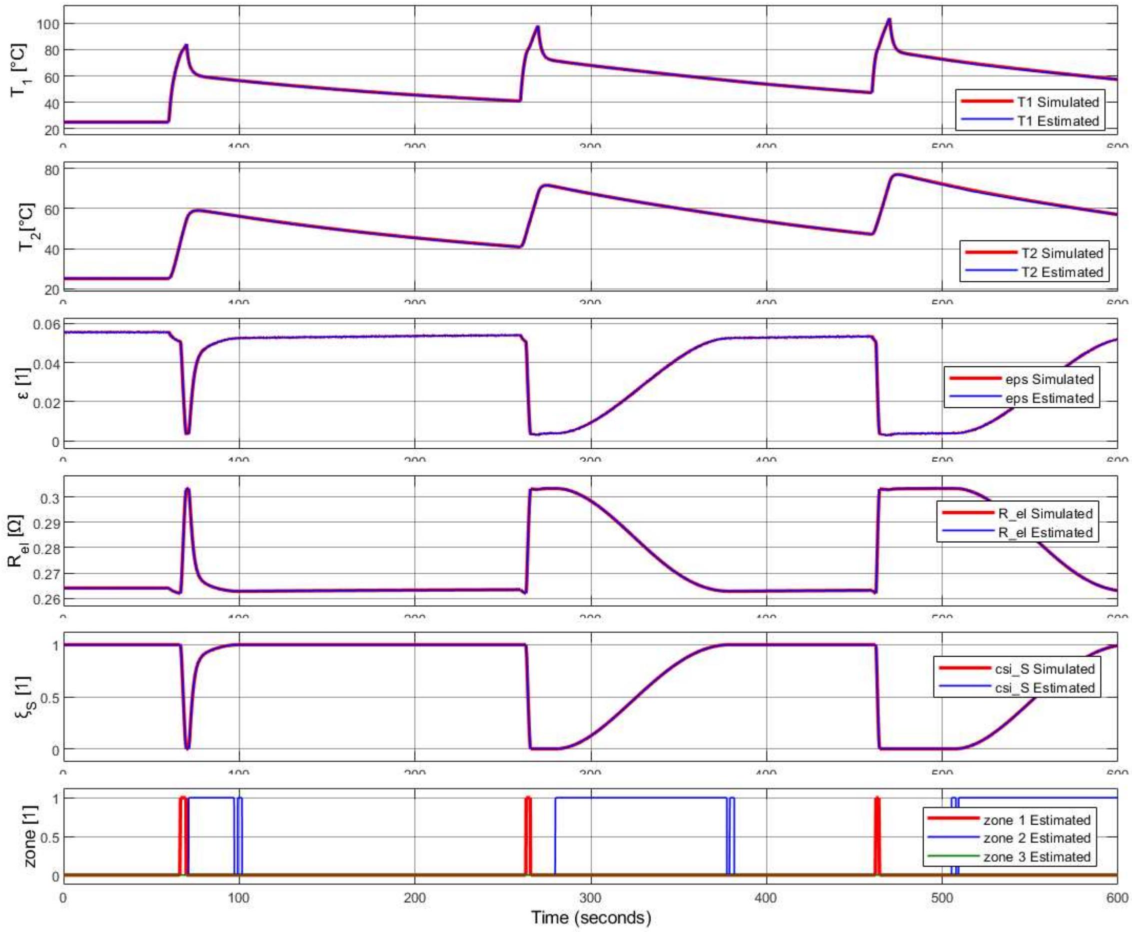


Figure 48 - State variables estimation on complete cycles with h = 5 W/m²K

In Figure 48 we see the comparison between simulated and estimated state variables, we can see that the estimated variables follow quite well the simulated ones.

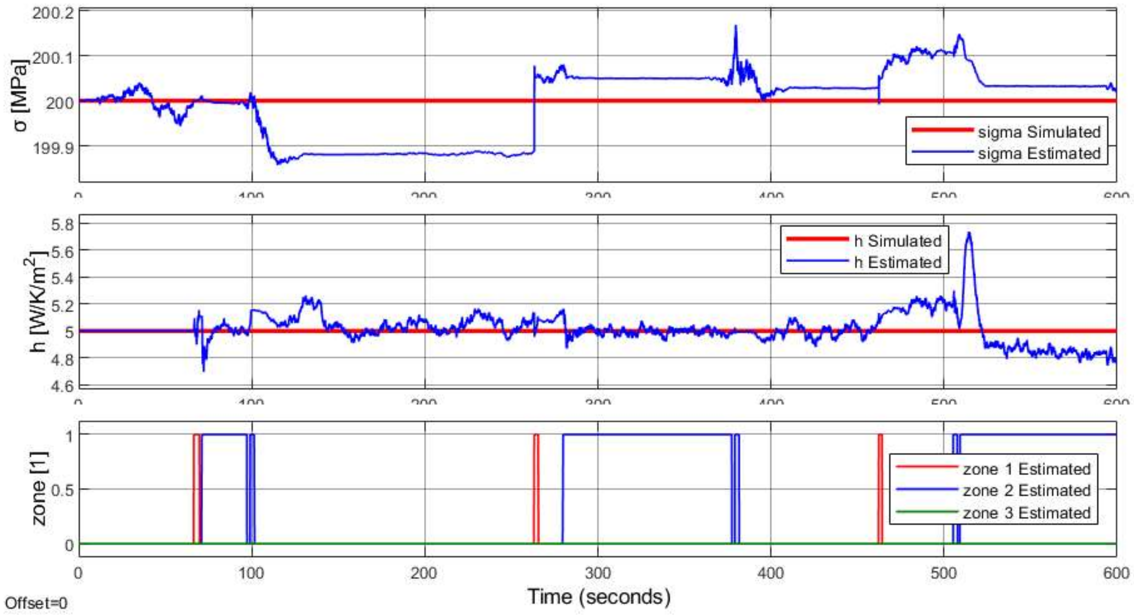


Figure 49 - Unknown parameter estimation on complete cycles with h = 5 W/m²K

In Figure 49 we observe the estimation of the unknown parameters and their simulated values. The estimated values are affected by the presence of the measurement noise, but they stay close to the simulated values; during the simulated timespan the estimated σ has mean value 200 MPa, standard deviation 0.075 MPa and its estimation error has maximum absolute value 0.169 MPa; while the estimated h has mean value 5.02 W/K/m², standard deviation 0.112 W/K/m² and maximum error absolute value 0.7361 W/K/m².

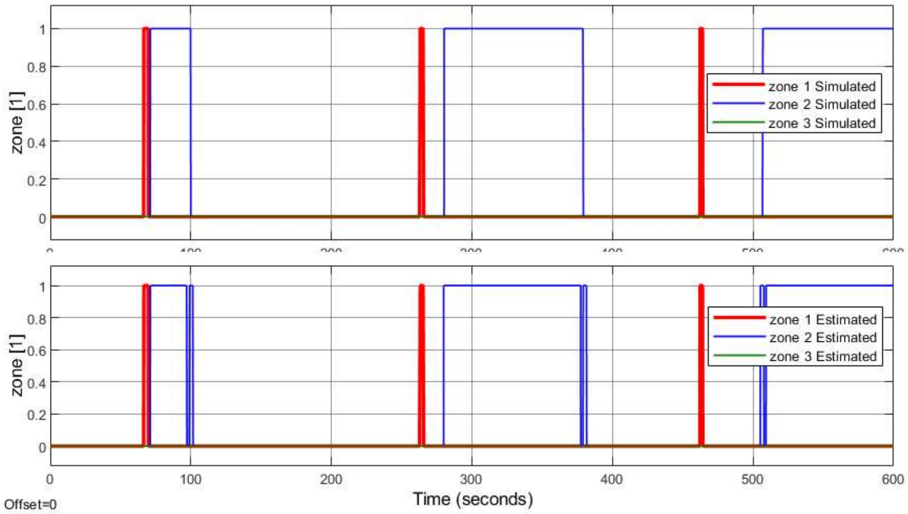


Figure 50 - Transformation zones estimation on complete cycles with h = 5 W/m²K

Finally, in Figure 50 we see the transformation zones, we can notice that the filter exits and reenters quickly zone 2 at 98 s, 378 s and 507 s, however we can verify *a posteriori* by looking at Figure 48 that this zone chattering does not have too many negative consequences on the estimation.

8.1.2 Tests with h = 15

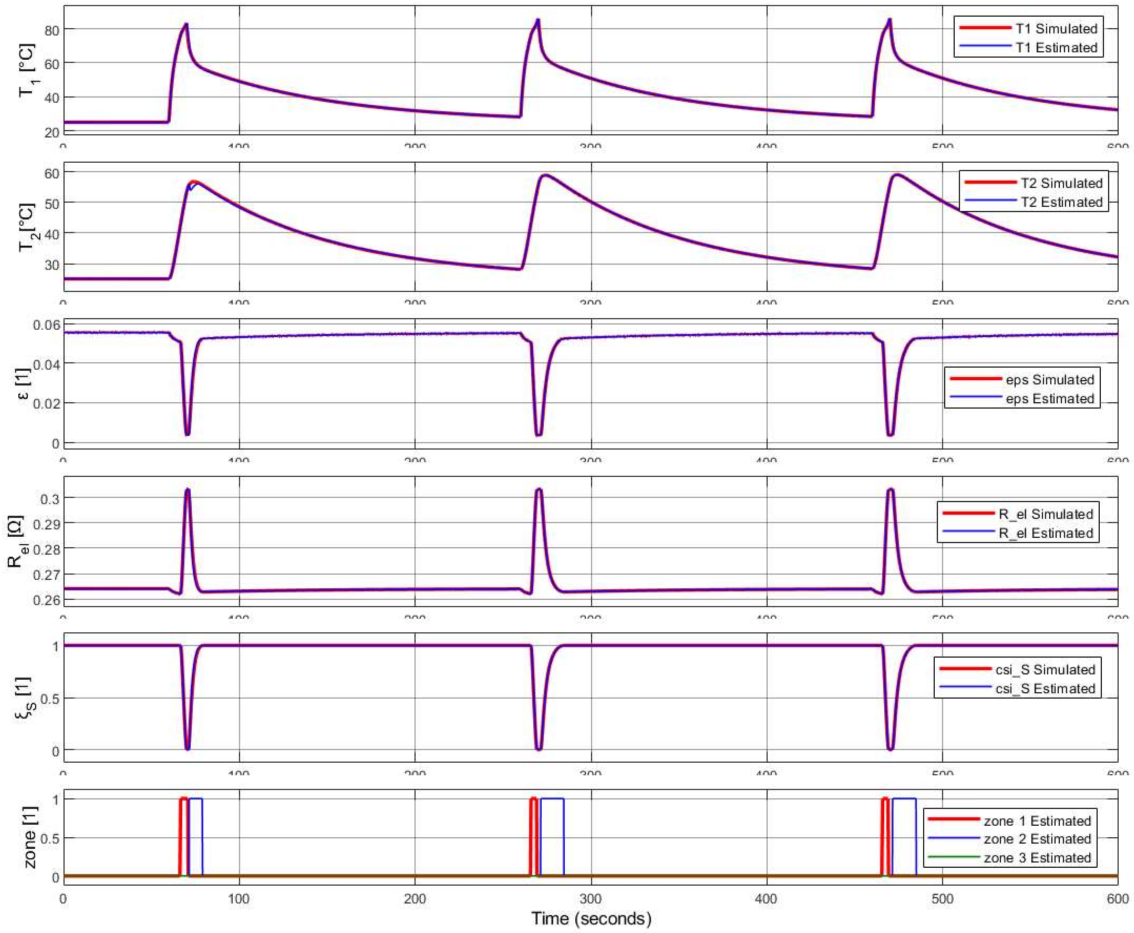


Figure 51 - State variables estimation on complete cycles with h = 15 W/m²K

Looking at the state variables in Figure 51 we can notice again a good filter performance, except for a small error of T_2 at 72 s, at the end of zone 1, which is however recovered by the end of the next zone.

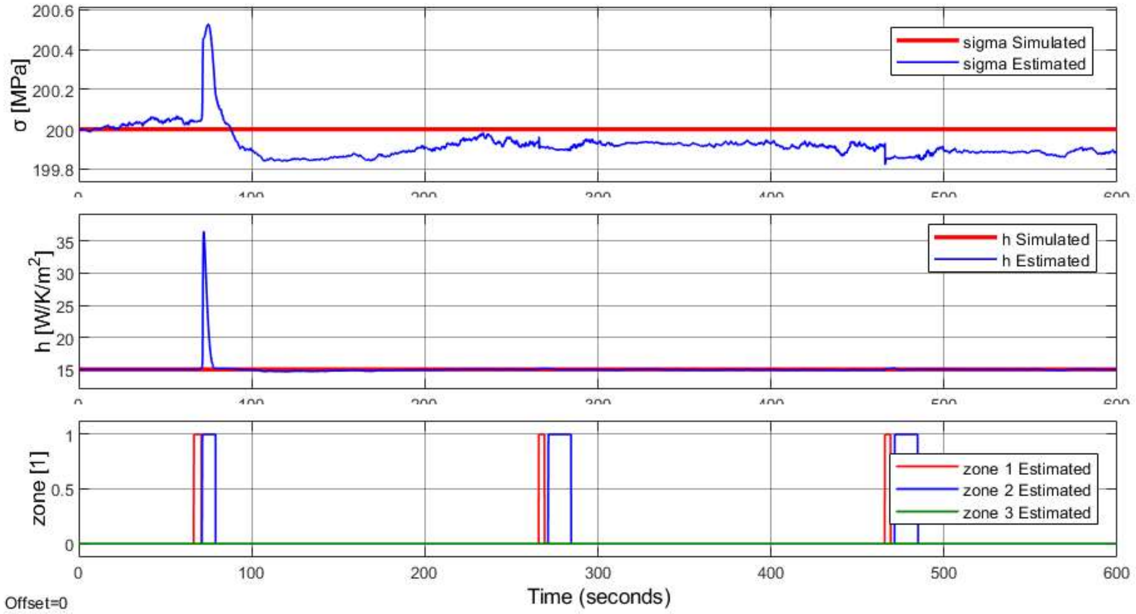


Figure 52 - Unknown parameter estimation on complete cycles with h = 15 W/m²K

We find the cause of the error on T_2 in Figure 52, where we observe two peaks in strain and convective heat transfer coefficient, which get propagated to T_2.

During the simulated timespan the estimated σ has mean value 199.9 MPa, standard deviation 0.079 MPa and its estimation error has maximum absolute value 0.53 MPa; while the estimated h has mean value 15.04 $W/K/m^2$, standard deviation 1.176 $W/K/m^2$ and maximum error absolute value 11.55 $W/K/m^2$.

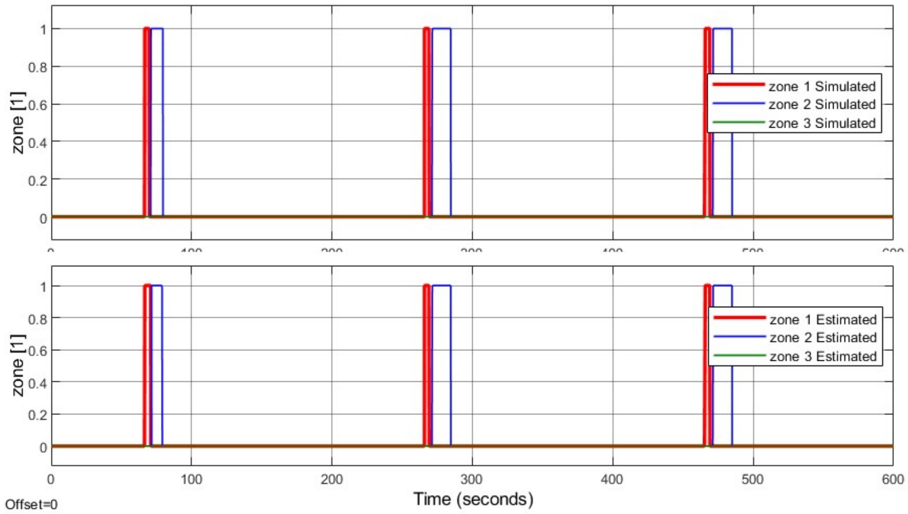


Figure 53 - Transformation zones estimation on complete cycles with h = 15 W/m^2K

The estimated transformation zones in Figure 53 are coherent with the simulated ones.

8.1.3 Tests with h = 35

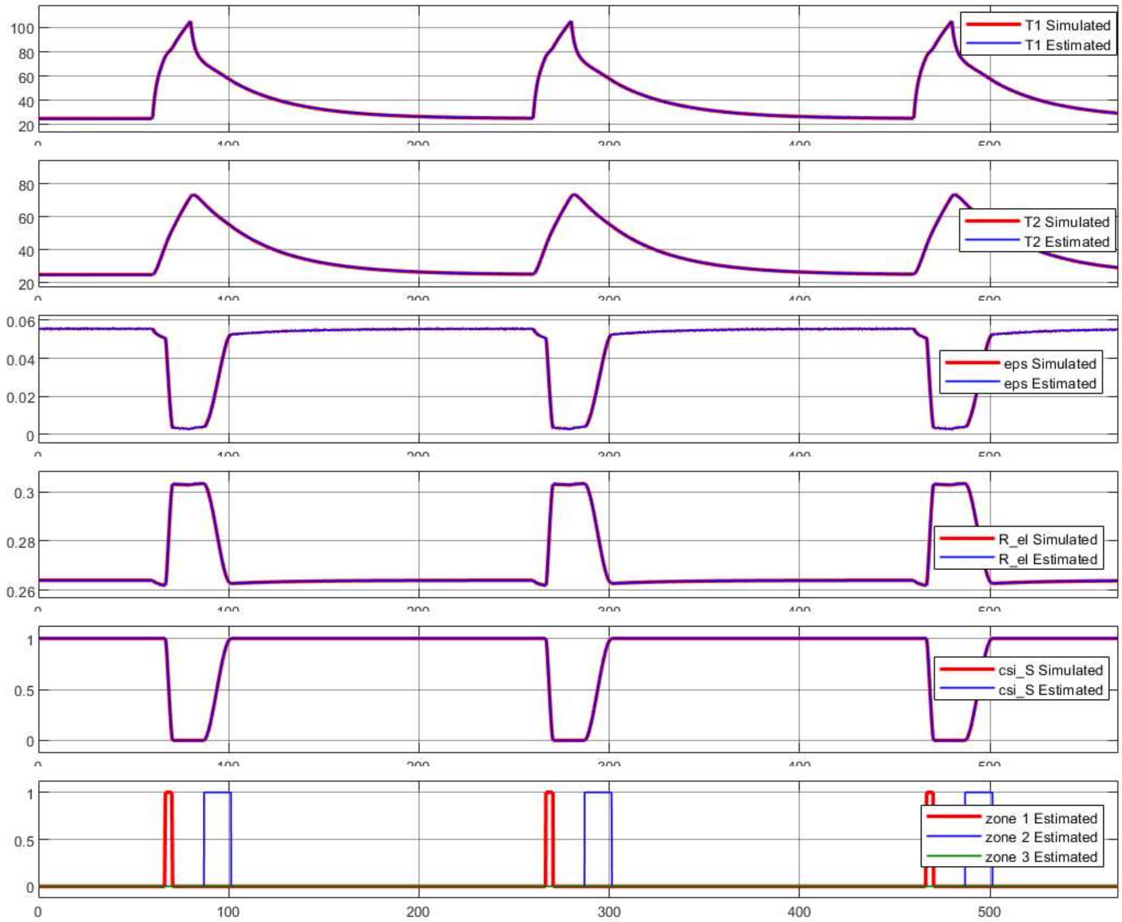


Figure 54 - State variables estimation on complete cycles with h = 35 W/m²K

We see from Figure 54 that the state variable estimations are again in good accordance with the simulated ones.

Below in Figure 55 are the estimated unknown parameters.

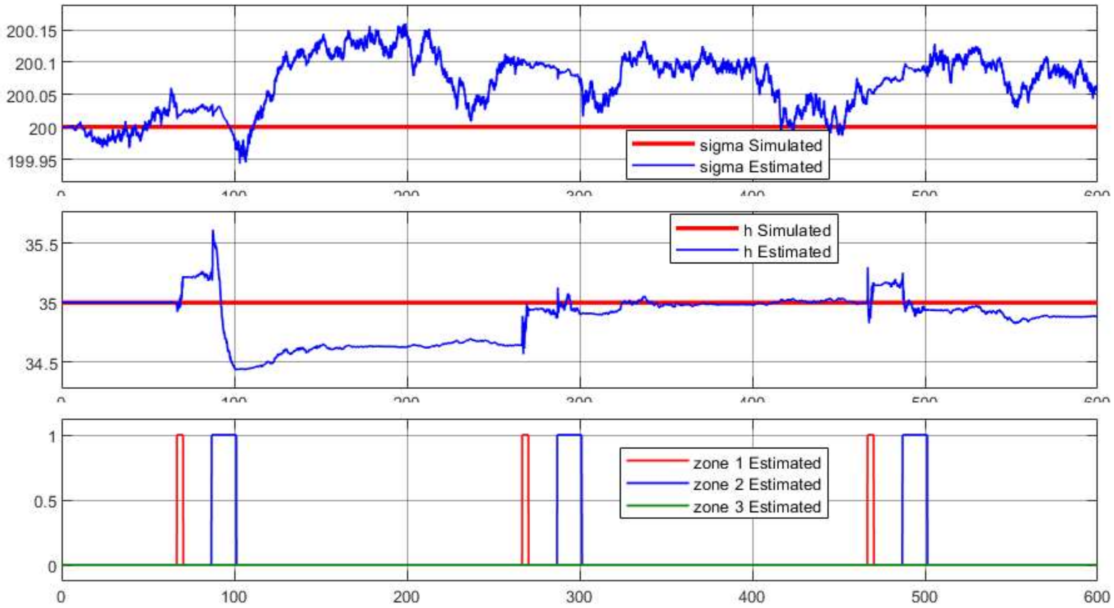


Figure 55 - Unknown parameter estimation on complete cycles with h = 35 W/m²K

During the simulated timespan the convective heat coefficient h has mean value 34.88 W/K/m², standard deviation 0.190 W/K/m² and maximum error absolute value 0.61 W/K/m².
The estimated σ has mean value 200.1 MPa, standard deviation 0,044 MPa and its estimation error has maximum absolute value 0.2 MPa.

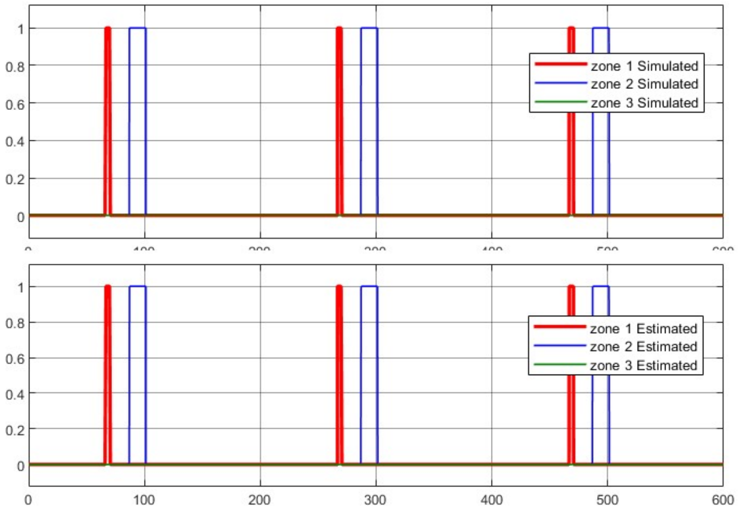


Figure 56 - Transformation zones estimation on complete cycles with h = 35 W/m²K

In Figure 56 we can observe that the estimated transformation zones are coherent with the simulated ones.

8.2 Estimations of closed loop system with correct initial conditions

Having verified the correct performance of the EKF on complete heating cycles, we wanted to test the performance and reactivity of the EKF and of the zone estimation block to variably changing temperatures and voltage by making it estimate the state variables of a closed loop system composed by the reference PID controller and the plant, with the test specifications defined in section 7.6.

8.2.1 Test with h = 5

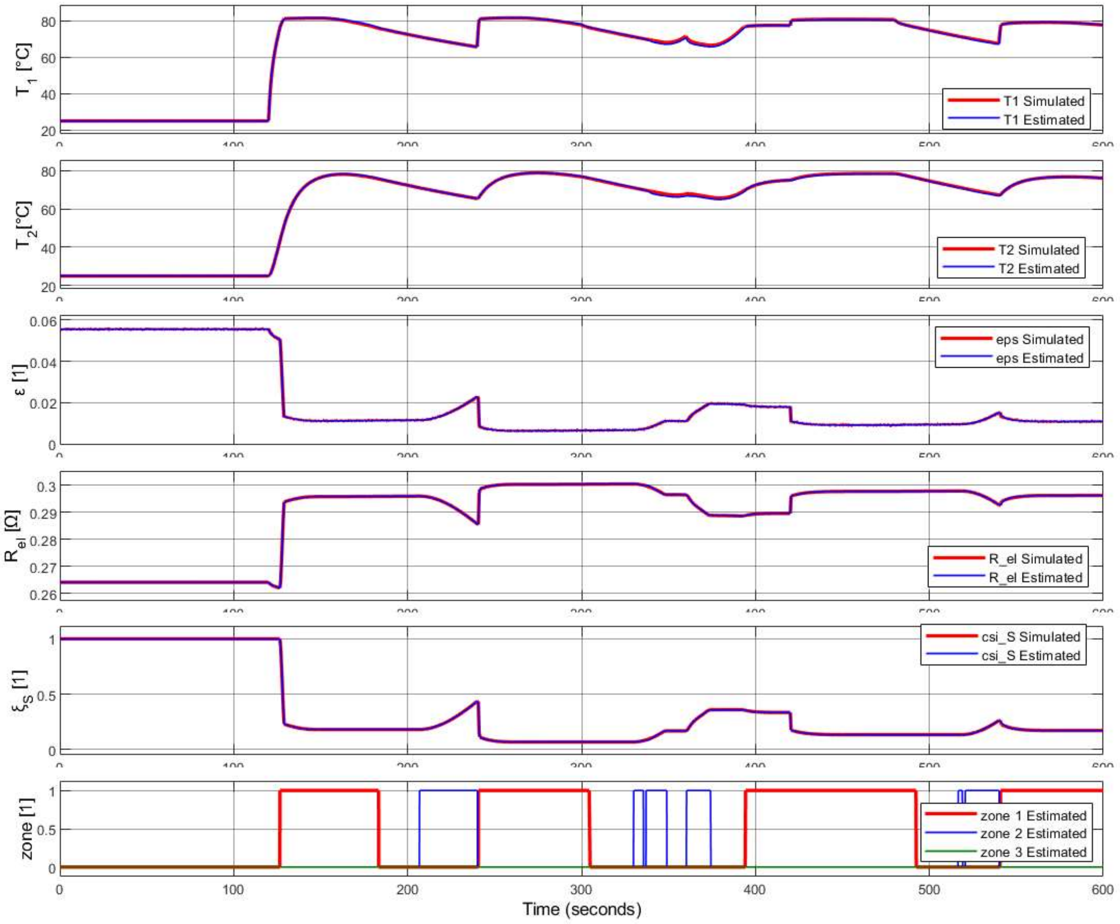


Figure 57 - State variables estimation of closed loop system with h = 5 W/m²K

We see from Figure 57 that the EKF is able to estimate correctly the state variables of the simulates system, the biggest discrepancies are seen during the cooling phases, with errors on T_1 and T_1 of around 1.2 °C.

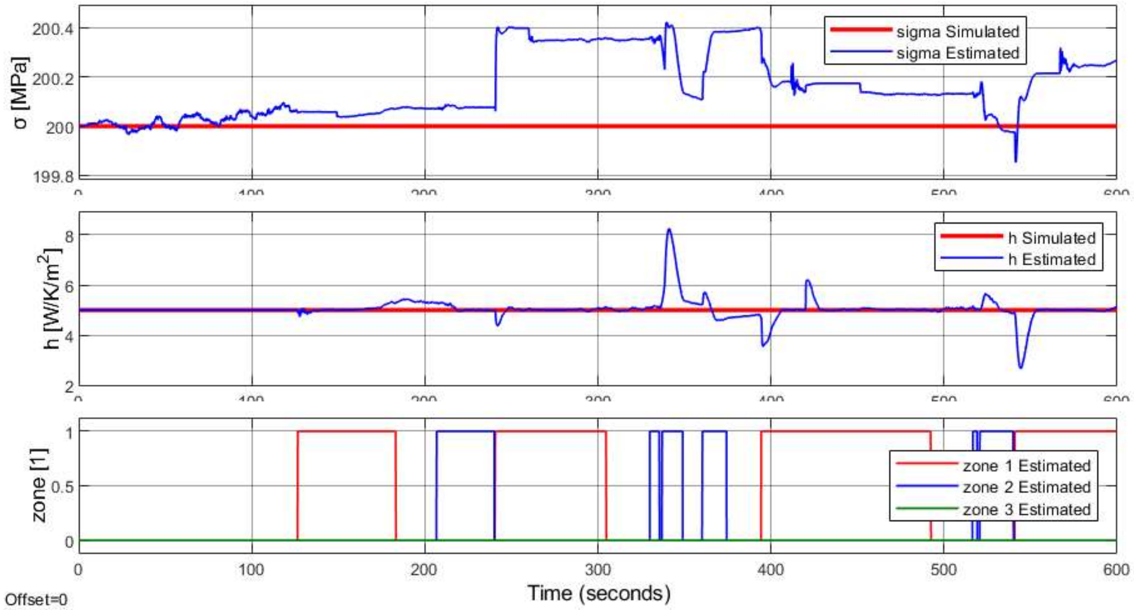


Figure 58 – Unknown parameter estimation of closed loop system with h = 5 W/m²K

In Figure 58 are shown the estimated unknown parameters, we see that the errors peaks in correspondence of the instants where the zones change.

The estimated σ has mean value 200.2 MPa, standard deviation 0.131 MPa and its estimation error has maximum absolute value 0.42 MPa; while the estimated h has mean value 5.04 W/K/m², standard deviation 0.416 W/K/m² and maximum error absolute value 3.25 W/K/m².

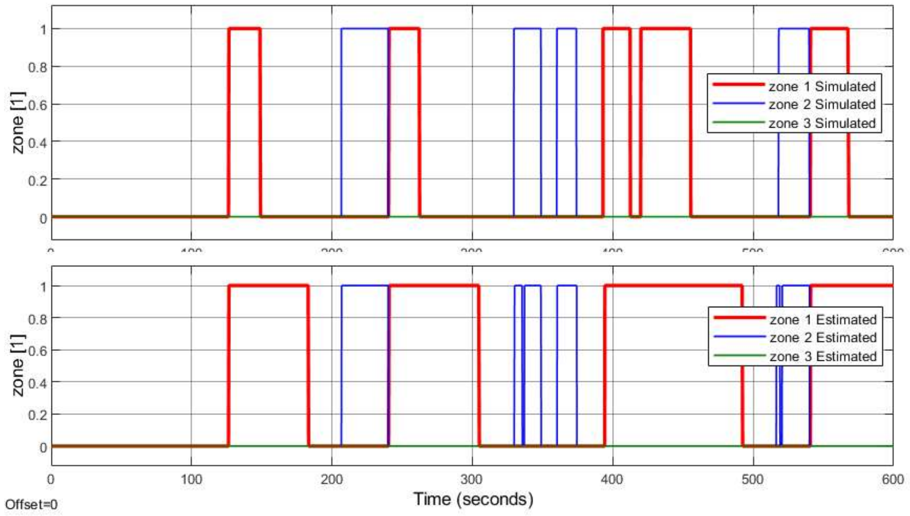


Figure 59 – Transformation zones estimation of closed loop system with h = 5 W/m²K

In Figure 59 we can observe the transformation zones comparison.

We see that the EKF tends to prolong zone 1 and sometimes exits and reenters zone 2 at the beginning of the transformation, however like in section 8.1.1 the error caused by this incorrect zone estimation is quickly corrected.

8.2.2 Test with h = 35

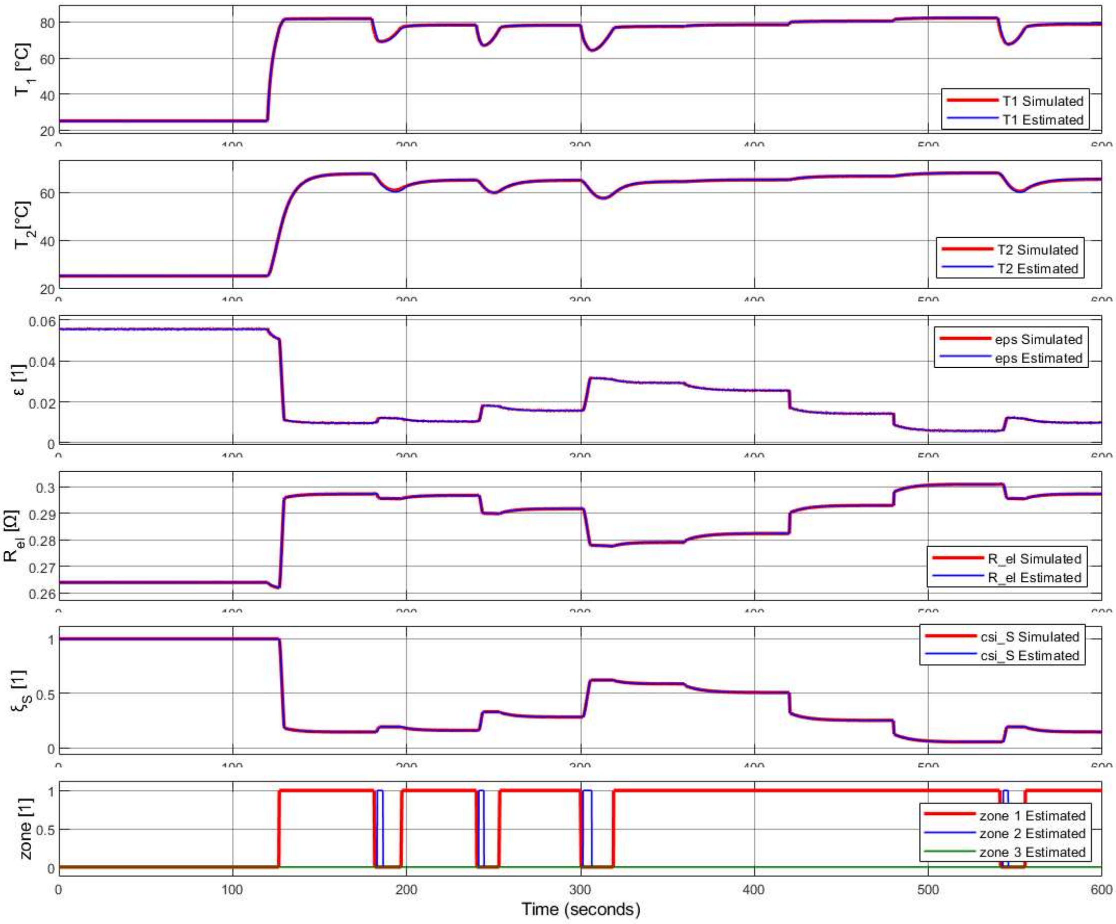


Figure 60 - State variables estimation of closed loop system with h = 35 W/m^2K

We can see from Figure 60 that in this other case as well that the estimated values of the variables follow the simulated values.

Below, in Figure 61, are the estimated unknown parameters and we can see that the estimation of h correctly follows the simulation, it has mean value 35.06 $W/K/m^2$, standard deviation 0.512 $W/K/m^2$ and maximum error absolute value 4.62 $W/K/m^2$ in the error peaks corresponding to zone 2; while the estimation of σ shows a small divergence, which is however recovered in longer simulations as will be shown with the tests in section 8.3, The estimated σ has mean value 200.3 MPa, standard deviation 0.278 MPa and its estimation error has maximum absolute value 1.1 MPa.

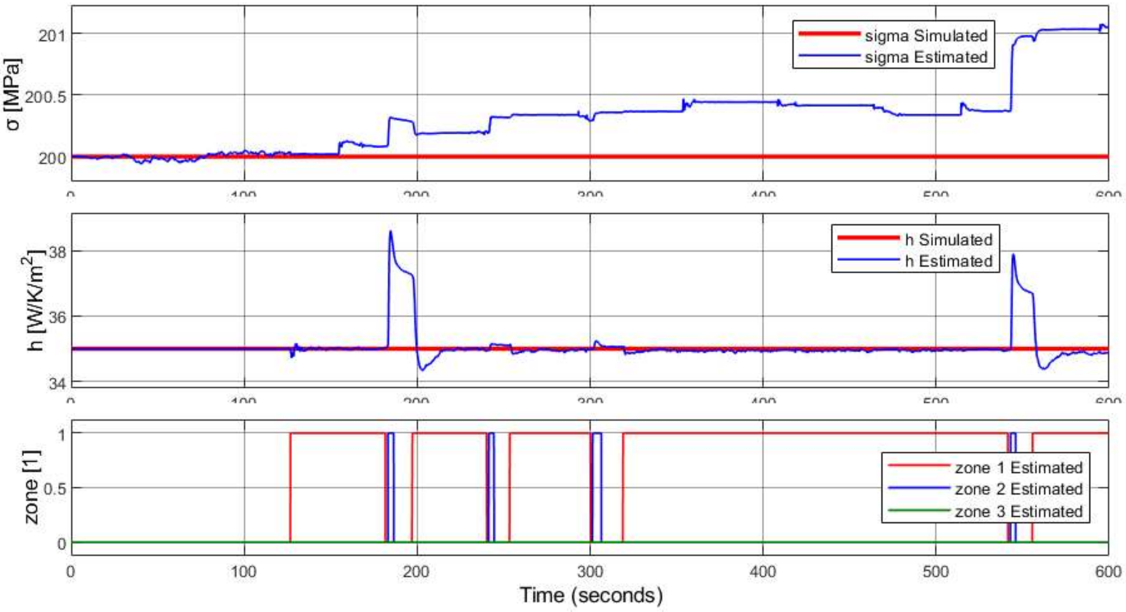


Figure 61 – Unknown parameter estimation of closed loop system with h = 35 W/m²K

In Figure 62 we can observe that the estimated transformation zones are coherent with the simulated ones.

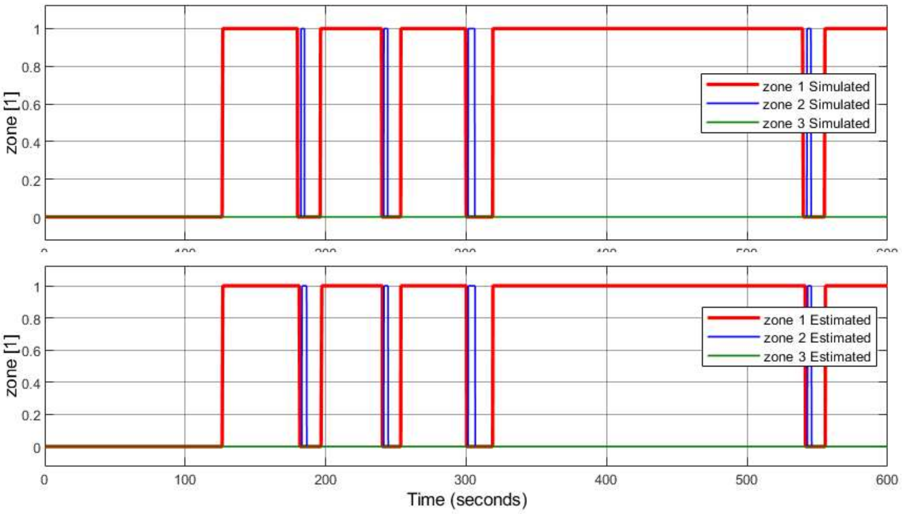


Figure 62 – Transformation zones estimation of closed loop system with h = 35 W/m²K

8.3 Estimations of closed loop system with incorrect initial conditions

The final set of tests we conducted was done to assess the robustness of the EKF relative to errors on the initial conditions by assigning wrong initial conditions to the filter.
h_{0f} and σ_{0f} are the filter initial conditions on the convective heat exchange coefficient and stress respectively and h_0 and σ_0 are the correct initial conditions of the system.
We tested various combinations of wrong initial conditions to better test the filter robustness to the uncertainties on the initial conditions.

8.3.1 Test with $h_{0f} = h_0 + 10$, $\sigma_{0f} = \sigma_0 + 10$

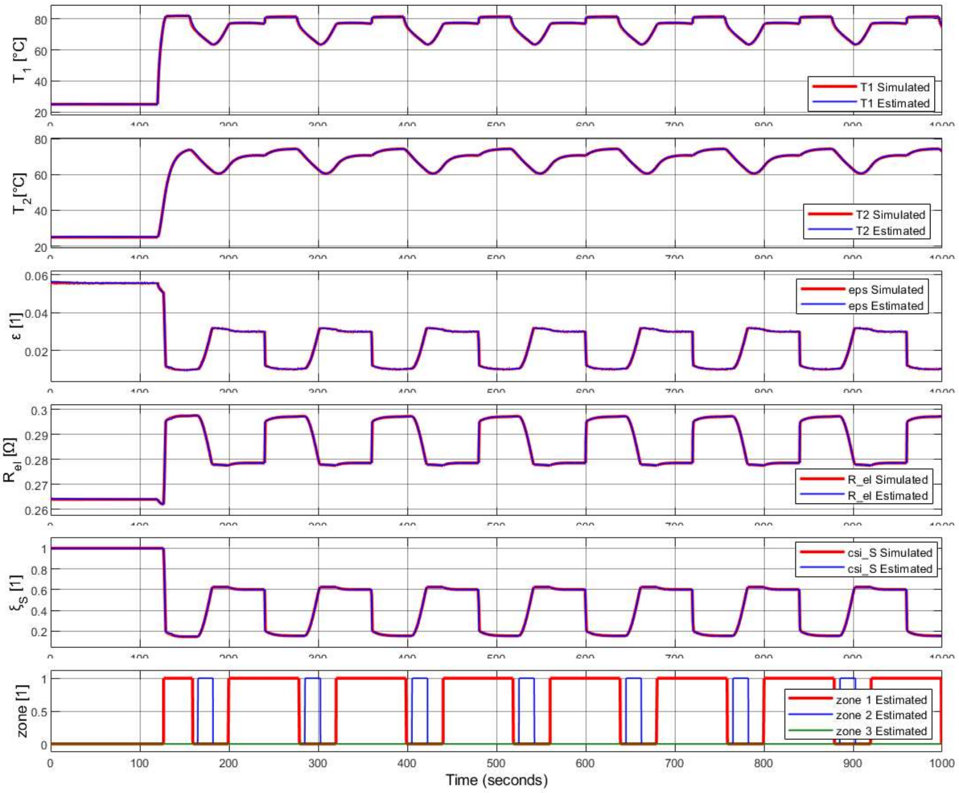


Figure 63 – State variables estimation on closed loop system with $h_{0f} = h_0 + 10$ W/m²K and $\sigma_{0f} = \sigma_0 + 10$ MPa

In Figure 63 we see the estimation of all the state variables of the closed loop system, it can be noted that the estimation keeps being quite correct despite the wrong initial conditions.

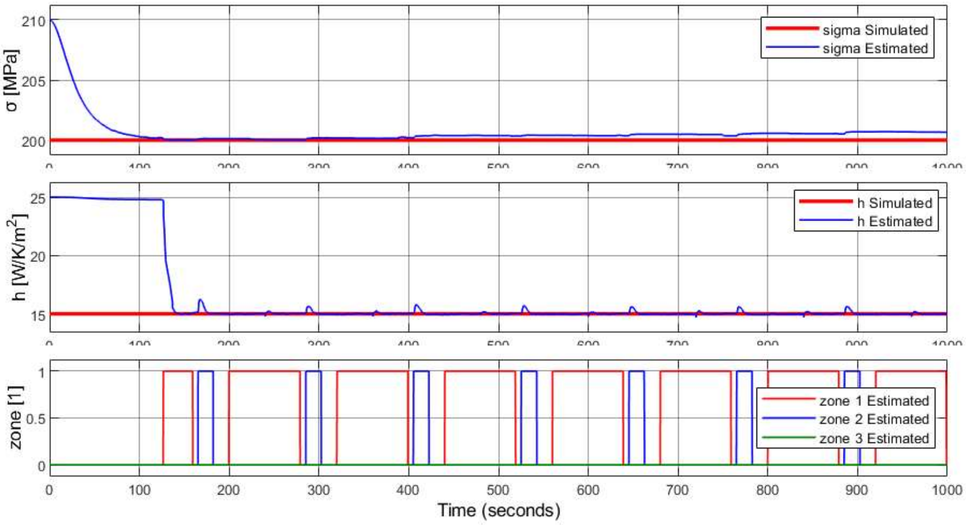


Figure 64 – Unknown parameter estimation on closed loop system with $h_{0f} = h_0 + 10$ W/m^{2}K and $\sigma_{0f} = \sigma_0 + 10$ MPa

In Figure 64 we see the estimation of the unknown parameters; the initial correction of σ occurs by the first transformation, which is also where the correction of h starts.
After the initial transitory, considered of 140 s, the estimated σ has mean value 200.38 MPa, standard deviation 0.195 MPa and its estimation error has maximum absolute value 0.7 MPa; while the estimated h has mean value 15.04 W/K/m^2, standard deviation 0.152W/K/m^2 and maximum error absolute value 1.25 W/K/m^2.

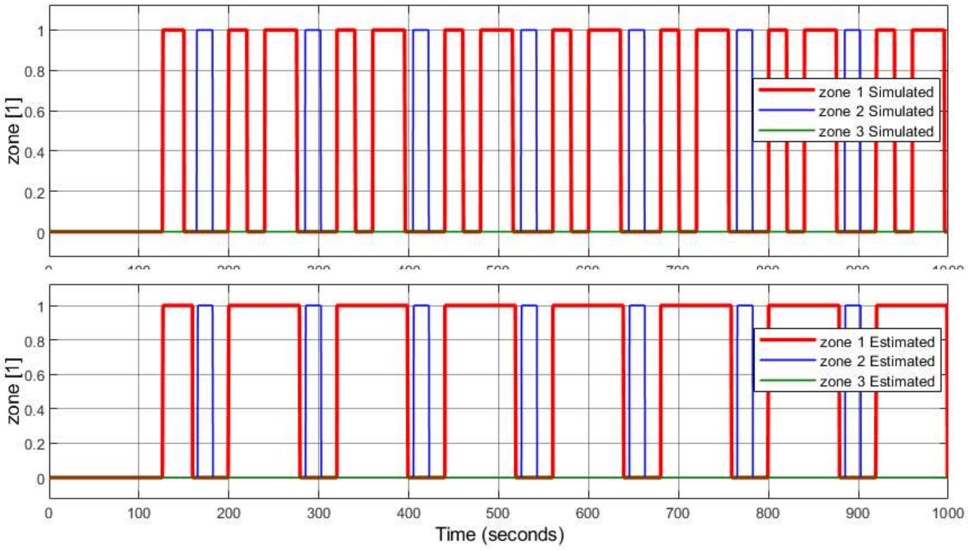


Figure 65 – Transformation zones estimation on closed loop system with $h_{0f} = h_0 + 10$ W/m^{2}K and $\sigma_{0f} = \sigma_0 + 10$ MPa

The comparison between estimated and simulated zones is shown in Figure 65, we can observe that the system merges two consecutive zones 1 into a big one, however this does not appear to be a problem, as in the other graphs it is clear that the maximum errors occur between zone 1 and 2.

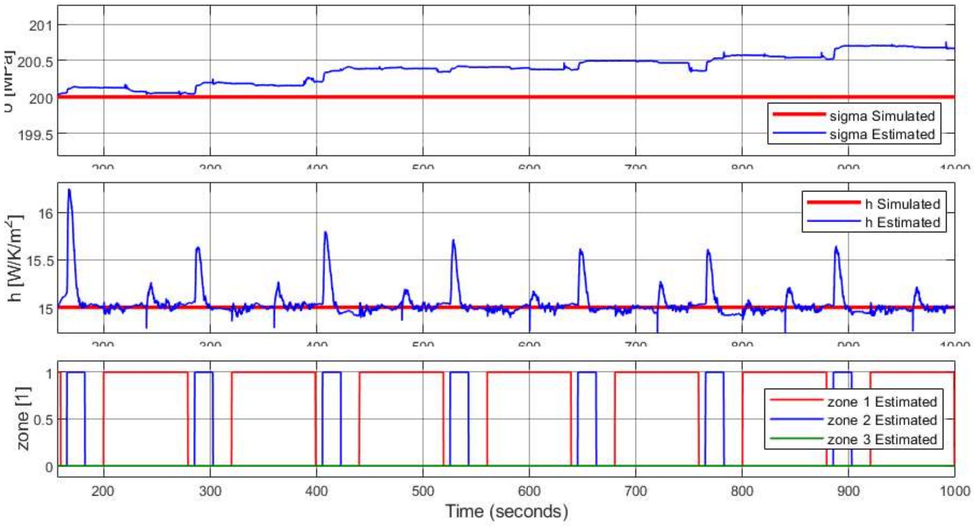


Figure 66 – Wire stress estimation zoom after transient with $h_{0f} = h_0 + 10$ W/m²K and $\sigma_{0f} = \sigma_0 + 10$ MPa

In Figure 66 we see the estimated stress after the error correction transient.

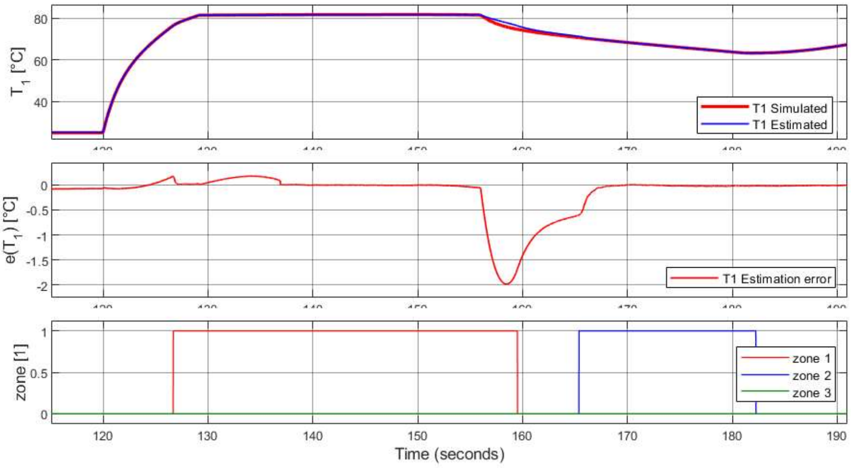


Figure 67 – Wire temperature estimation zoom on zone 1 with $h_{0f} = h_0 + 10$ W/m²K and $\sigma_{0f} = \sigma_0 + 10$

In Figure 67 we see in detail the largest error on T_1 in zone 1, which is also the largest error of this variable overall, and has a magnitude of 2 °C, while in the rest of the simulation the error is much smaller.

We focus on the error of this variable in this particular zone because, as discussed in section 9.1 this variable, and the following ones, will be used in our formulation of an improved PID controller.

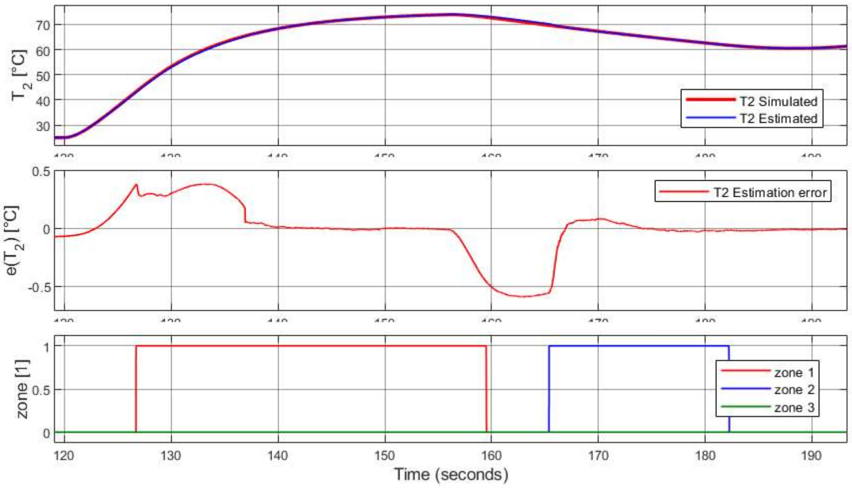


Figure 68 – Silicone temperature estimation zoom on zone 1 with $h_{0f} = h_0 + 10$ W/m²K and $\sigma_{0f} = \sigma_0 + 10$ MPa

In Figure 68 the maximum error of T_2 in zone 1 is shown to be 0.38 °C, which is 0.63% of the signal at the same time instant.

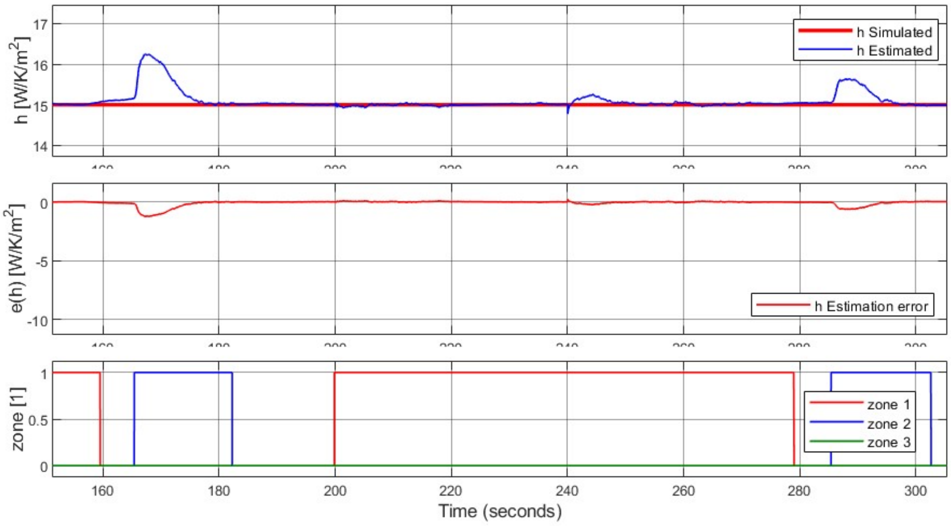


Figure 69 – Convective heat exchange coefficient estimation zoom on zone 1 with $h_{0f} = h_0 + 10$ W/m²K and $\sigma_{0f} = \sigma_0 + 10$ MPa

In Figure 69 we see h in the zones of interest, that is between the end of zone 1 and the beginning of zone 2; here the maximum error is 0.2 W/K/m², the maximum overall error is found in zone 2 but we do not use h to modify the controller in that region.

8.3.2 Test with $h_{0f} = h_0 + 15$, $\sigma_{0f} = \sigma_0 - 20$

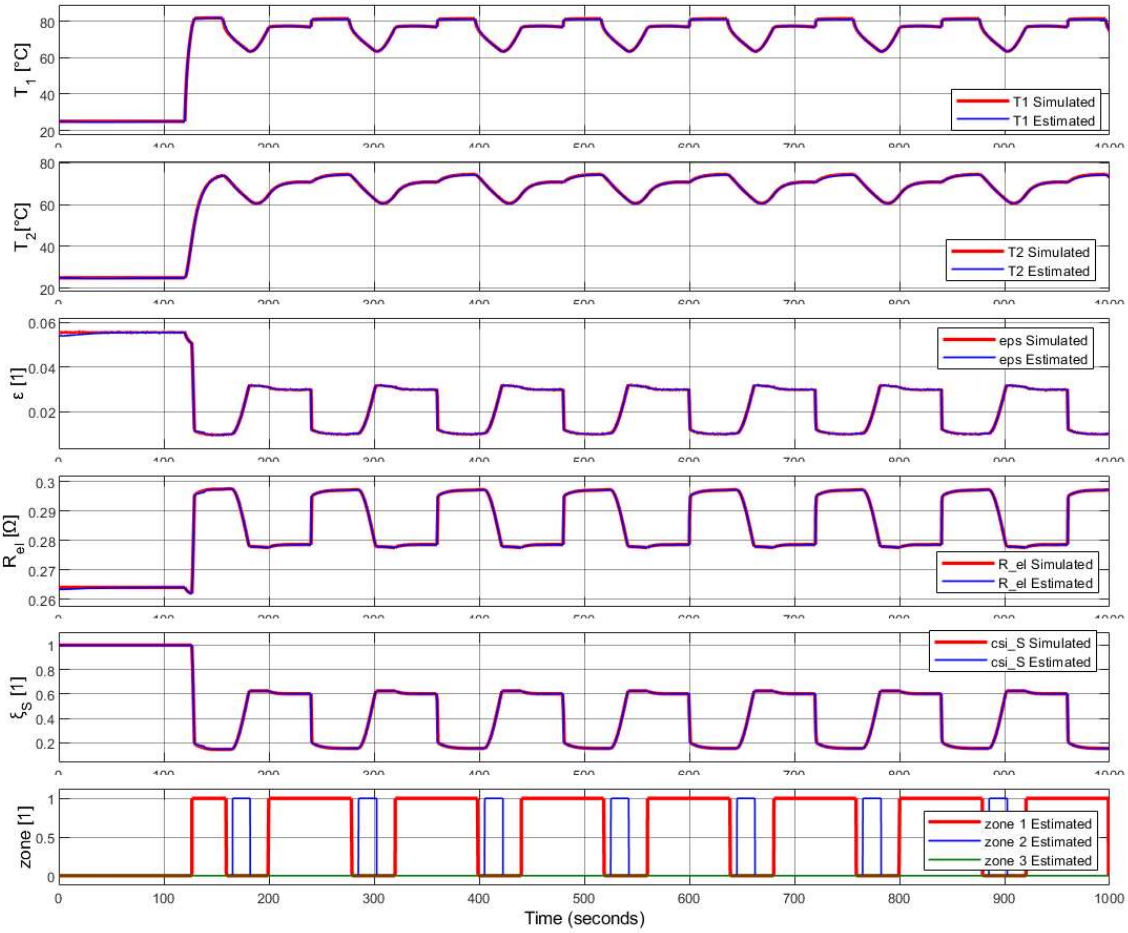


Figure 70 – State variables estimation on closed loop system with $h_{0f} = h_0 + 15$ W/m²K and $\sigma_{0f} = \sigma_0 - 20$ MPa

In Figure 70 we see the estimation of all the state variables of the closed loop system, in this case also the estimation is coherent with the simulations.

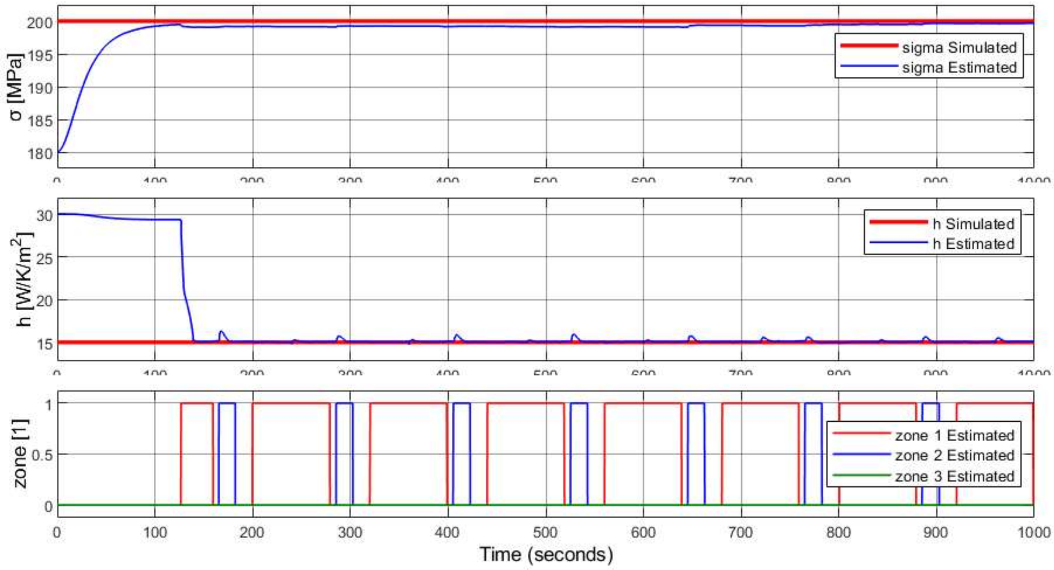


Figure 71 – Unknown parameter estimation on closed loop system with $h_{0f} = h_0 + 15$ W/m²K and $\sigma_{0f} = \sigma_0 - 20$ MPa

In Figure 71 we observe the estimated unknown parameters.

After the initial transitory the estimated σ has mean value 199. 3 MPa, standard deviation 0.182 MPa and its estimation error has maximum absolute value 0.25 MPa; while the estimated h has mean value 15.11 W/K/m^2, standard deviation 0.168 W/K/m^2 and maximum error absolute value 1.32 W/K/m^2.

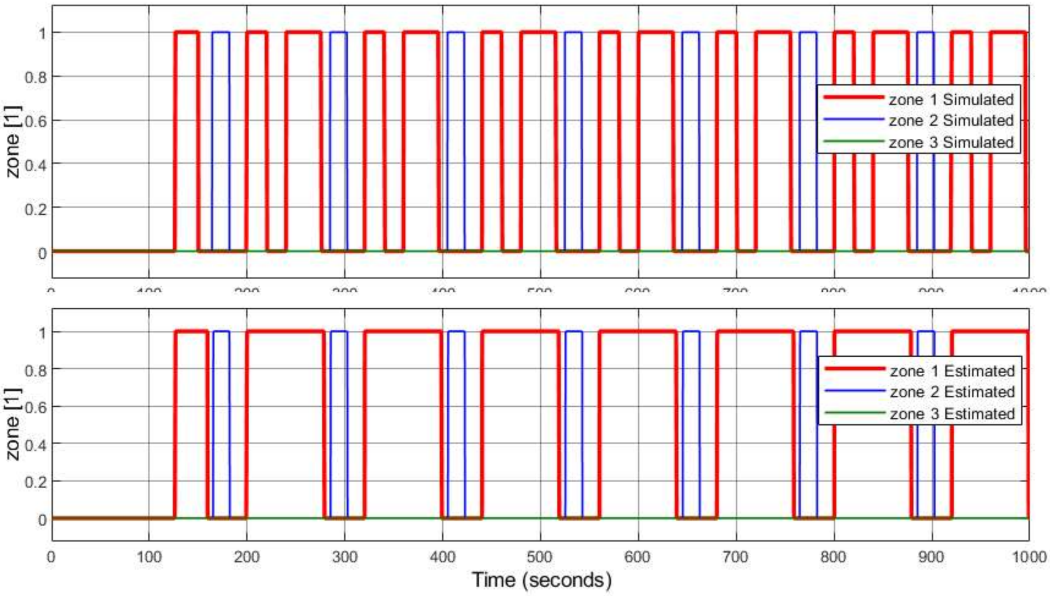


Figure 72 – Transformation zones estimation on closed loop system with $h_{0f} = h_0 + 15$ W/m^{2}K and $\sigma_{0f} = \sigma_0 - 20$ MPa

For Figure 72 are valid the same consideration done for Figure 65 in section 8.3.1.

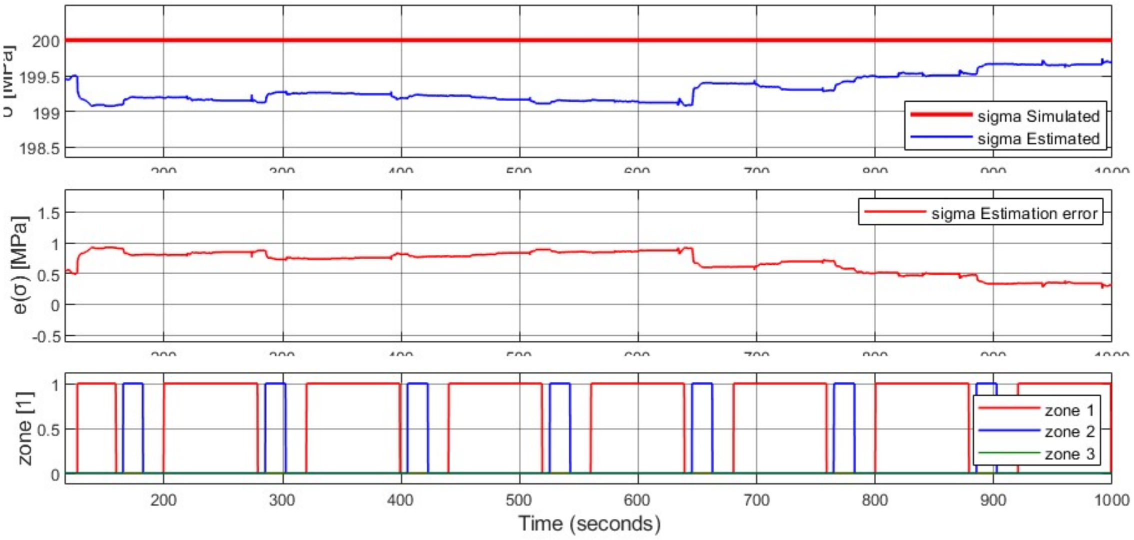


Figure 73 – Wire stress estimation after transient with $h_{0f} = h_0 + 15$ W/m^{2}K and $\sigma_{0f} = \sigma_0 - 20$ MPa

In Figure 73 we see the estimated stress after the error correction.

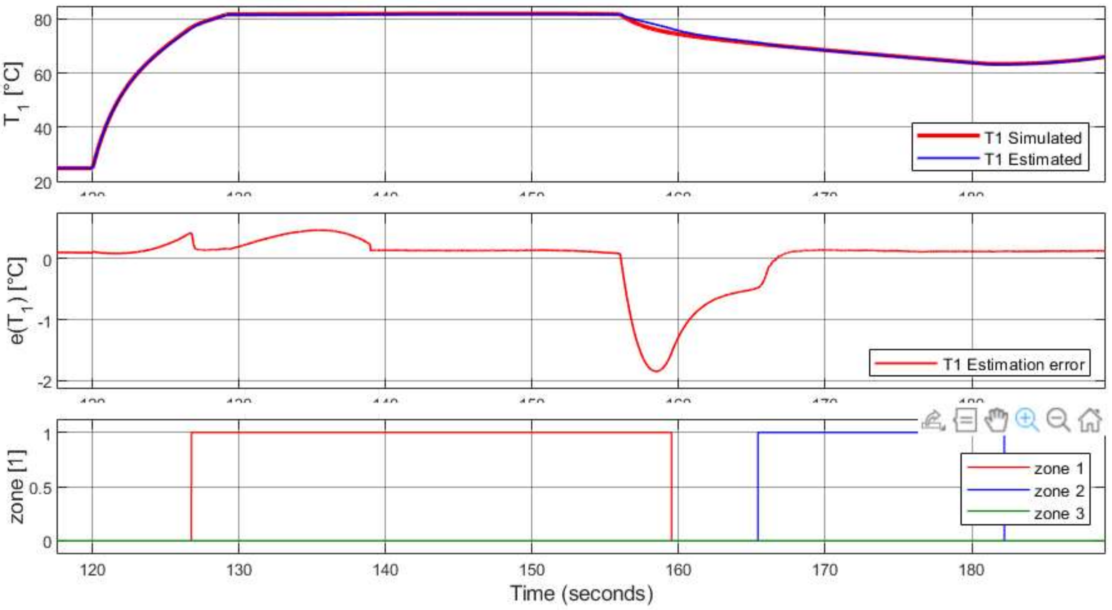


Figure 74 – Wire temperature estimation zoom on zone 1 with $h_{0f} = h_0 + 15$ W/m²K and $\sigma_{0f} = \sigma_0 - 20$ MPa

In Figure 74 we see in detail the behavior of T_1 in zone 1, we can see that the largest error has a magnitude of 1.7 °C.

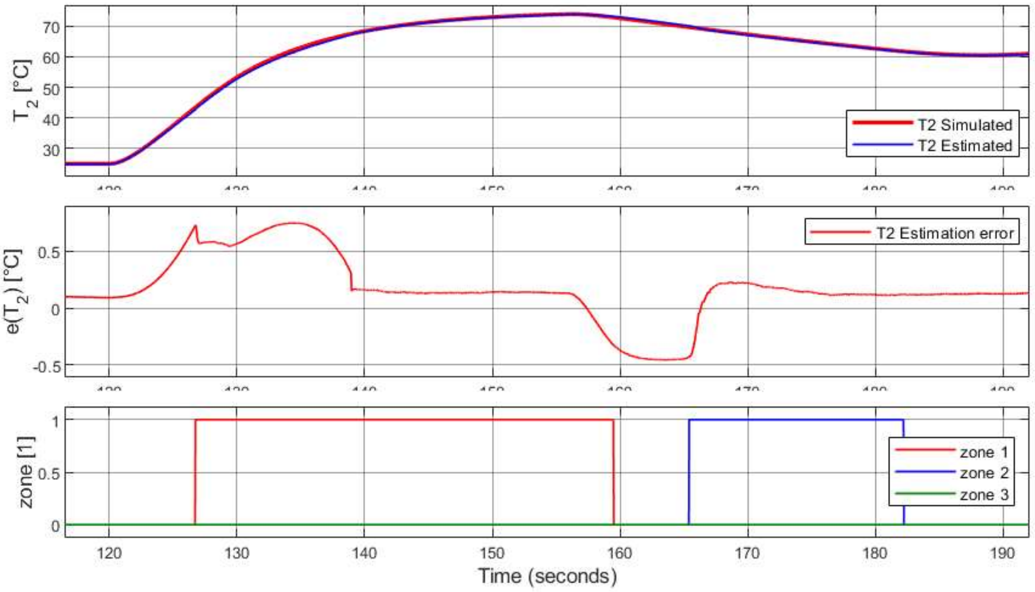


Figure 75 – Silicone temperature estimation zoom on zone 1 with $h_{0f} = h_0 + 15$ W/m²K and $\sigma_{0f} = \sigma_0 - 20$ MPa

In Figure 75 we see a zoom of the behavior of T_2 in zone 1, where the error does not exceed 0.8 °C.

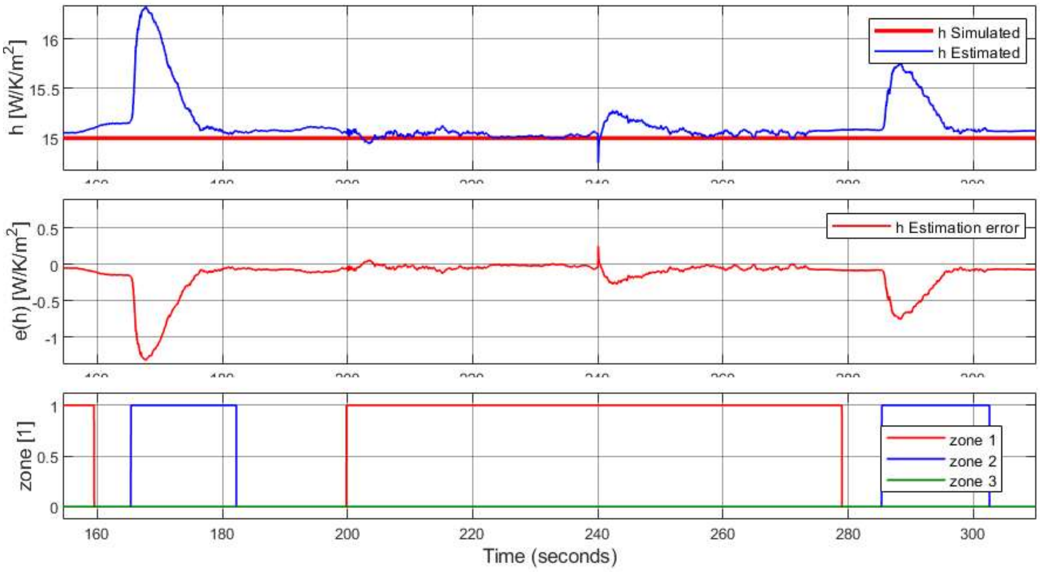


Figure 76 – Convective heat exchange coefficient zoom on zone 1with $h_{0f} = h_0 + 15$ W/m^{2}K and $\sigma_{0f} = \sigma_0 - 20$ MPa

Finally, in Figure 76 we see h between the end of zone 1 and the beginning of zone 2, here the maximum error is 0.2 W/K/m^2.

8.3.3 Test with $h_{0f} = h_0 - 10$, $\sigma_{0f} = \sigma_0 - 20$

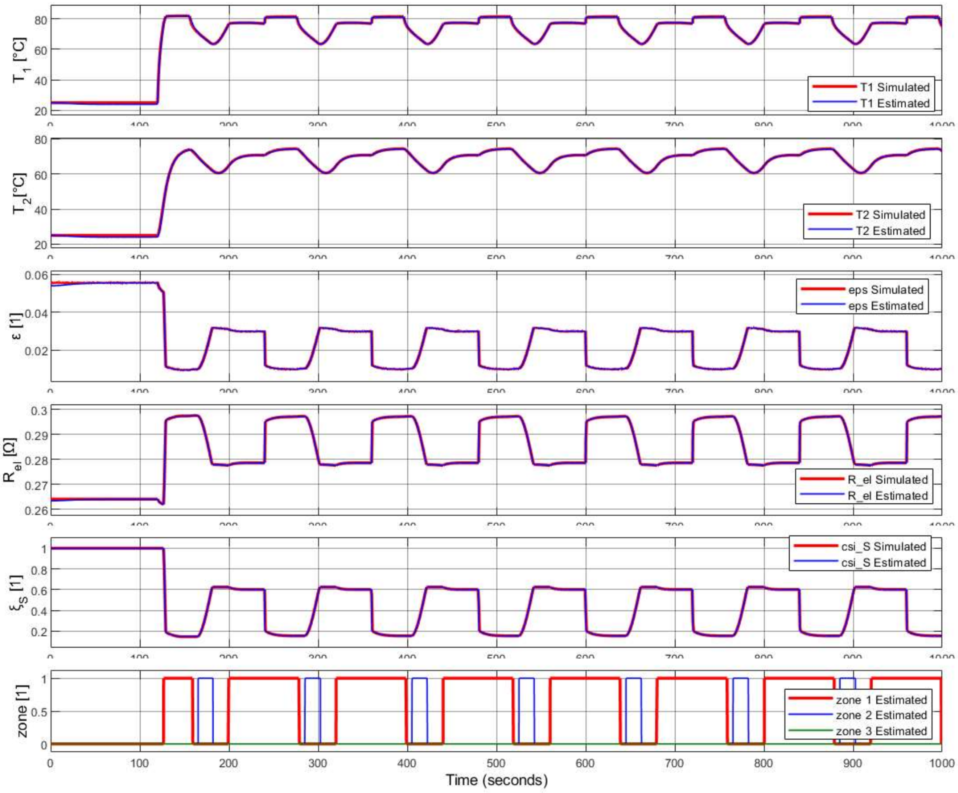


Figure 77 – State variables estimation on closed loop system with $h_{0f} = h_0 - 10$ W/m^{2}K and $\sigma_{0f} = \sigma_0 - 20$ MPa

In Figure 77 we see that the estimation of the state variables of the closed loop system is coherent with the simulated variables.

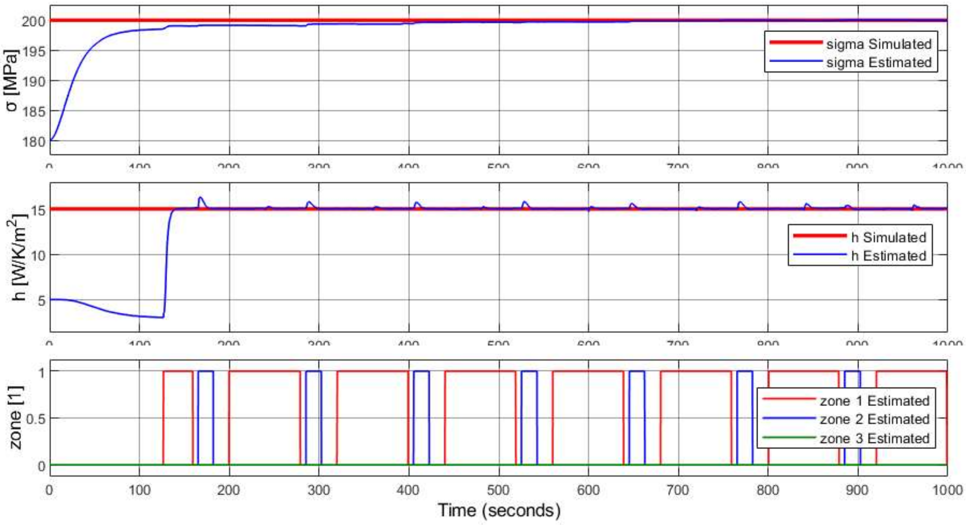


Figure 78 – Unknown parameter estimation on closed loop system with $h_{0f} = h_0 - 10$ W/m^{2}K and $\sigma_{0f} = \sigma_0 - 20$ MPa

In Figure 78 we observe the estimated unknown parameters of this test, we can observe that the correction of σ takes longer than for the other scenarios.

After the initial transitory the estimated σ has mean value 199. 63 MPa, standard deviation 0.323 MPa and its estimation error has maximum absolute value 0.05 MPa; while the estimated h has mean value 15.09 W/K/m^2, standard deviation 0.157 W/K/m^2 and maximum error absolute value 1.30 W/K/m^2.

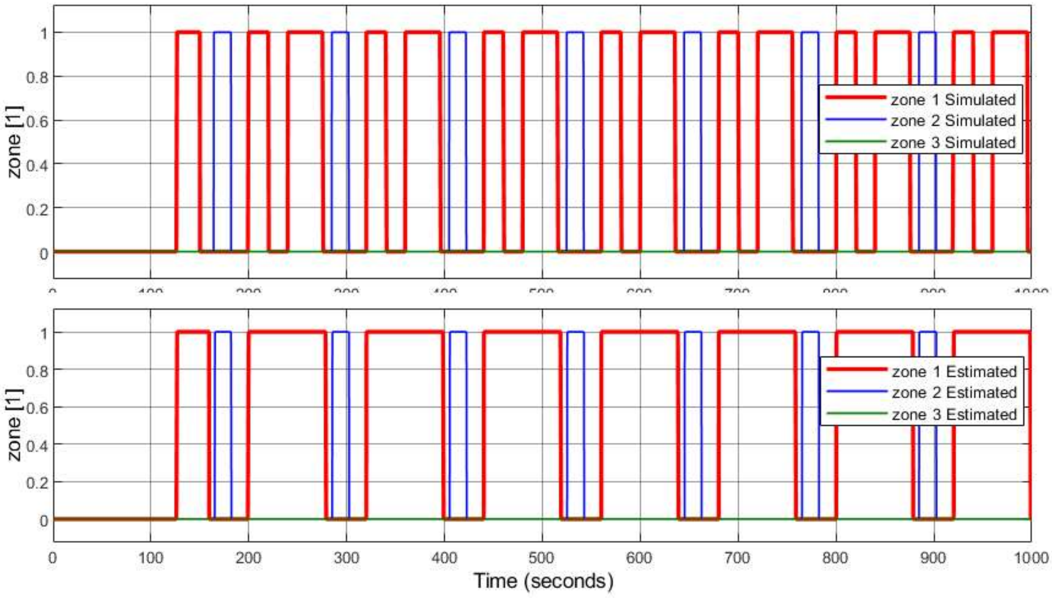


Figure 79 – Transformation zones estimation on closed loop system with $h_{0f} = h_0 - 10$ W/m^{2}K and $\sigma_{0f} = \sigma_0 - 20$ MPa

For Figure 79 are valid the same consideration done for Figure 65 in section 8.3.1.

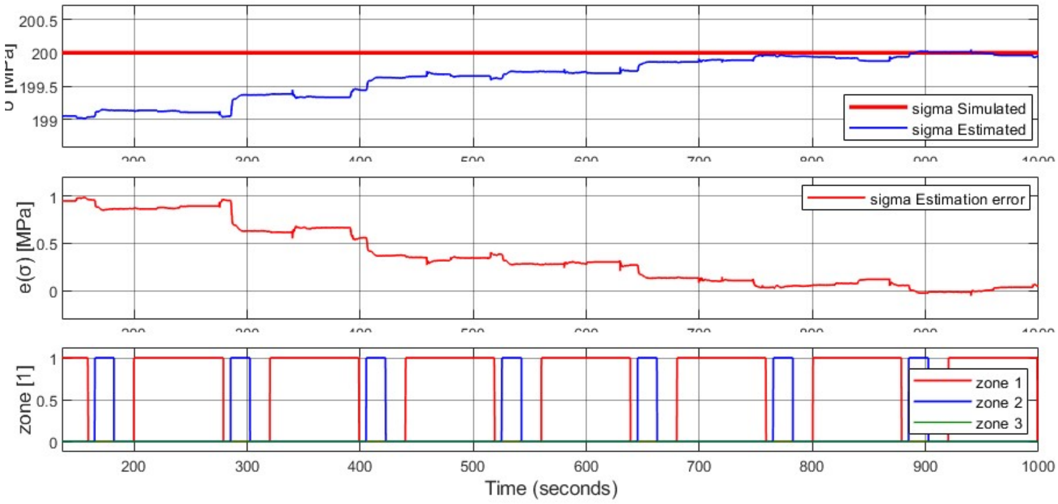


Figure 80 – Wire stress estimation after transient with $h_{0f} = h_0 - 10$ W/m²K and $\sigma_{0f} = \sigma_0 - 20$ MPa

In Figure 80 we see the estimated stress after the error correction.

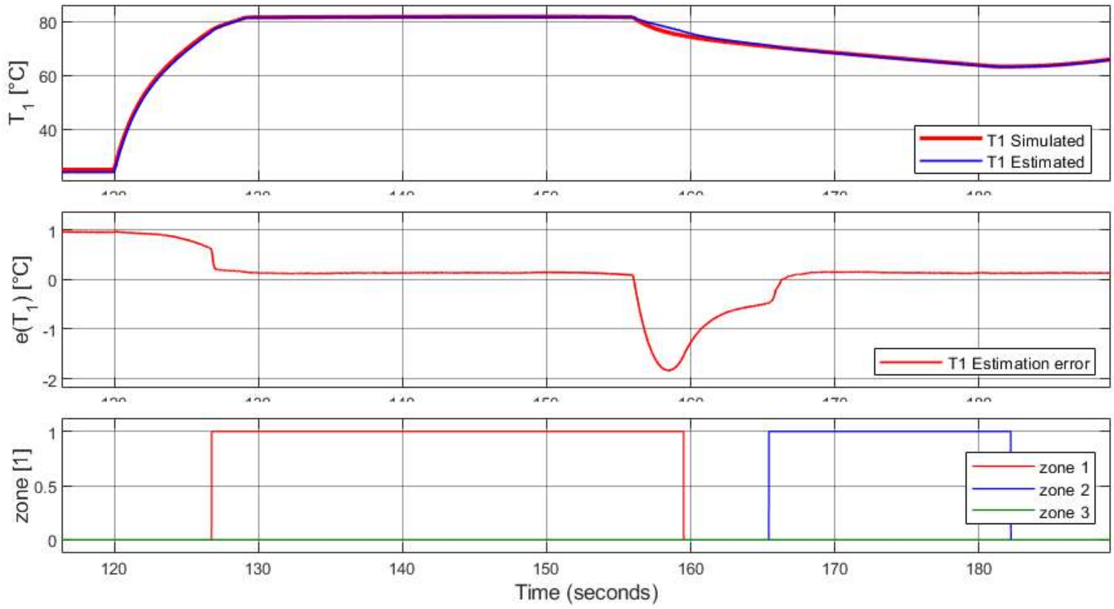


Figure 81 – Wire temperature estimation zoom on zone 1 with $h_{0f} = h_0 - 10$ W/m²K and $\sigma_{0f} = \sigma_0 - 20$ MPa

In Figure 81 we see in detail the behavior of T_1 in zone 1, where the largest error has a magnitude of 1.8 °C.

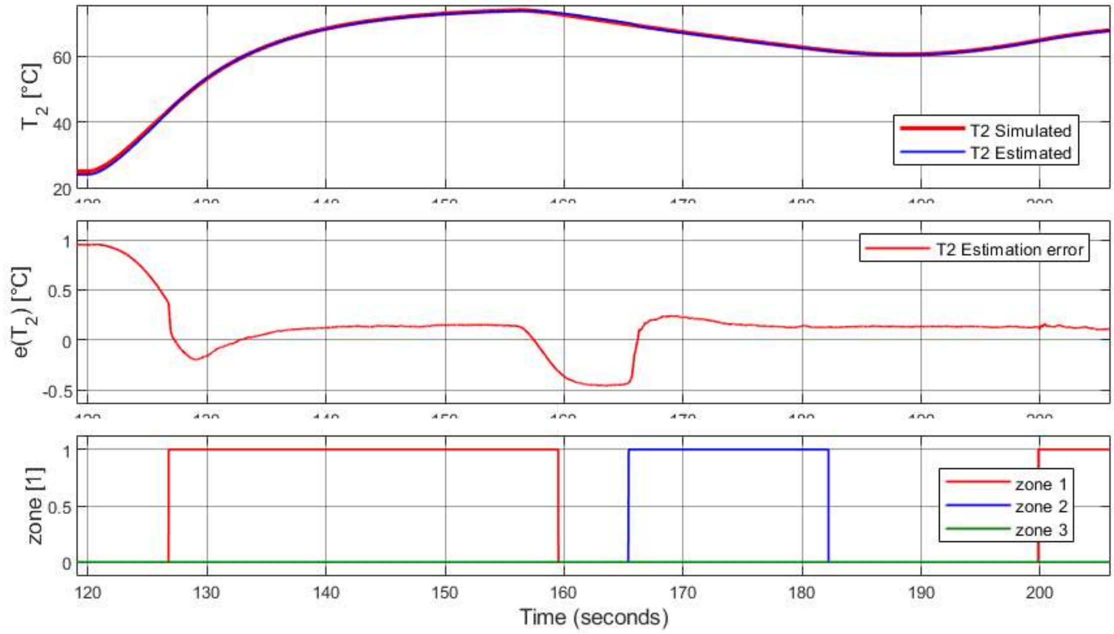


Figure 82 – Silicone temperature estimation zoom on zone 1 with $h_{0f} = h_0 - 10$ W/m^{2}K and $\sigma_{0f} = \sigma_0 - 20$ MPa

In Figure 82 we see the behavior of T_2 zoomed in zone 1, here the error does not exceed 0.4 °C.

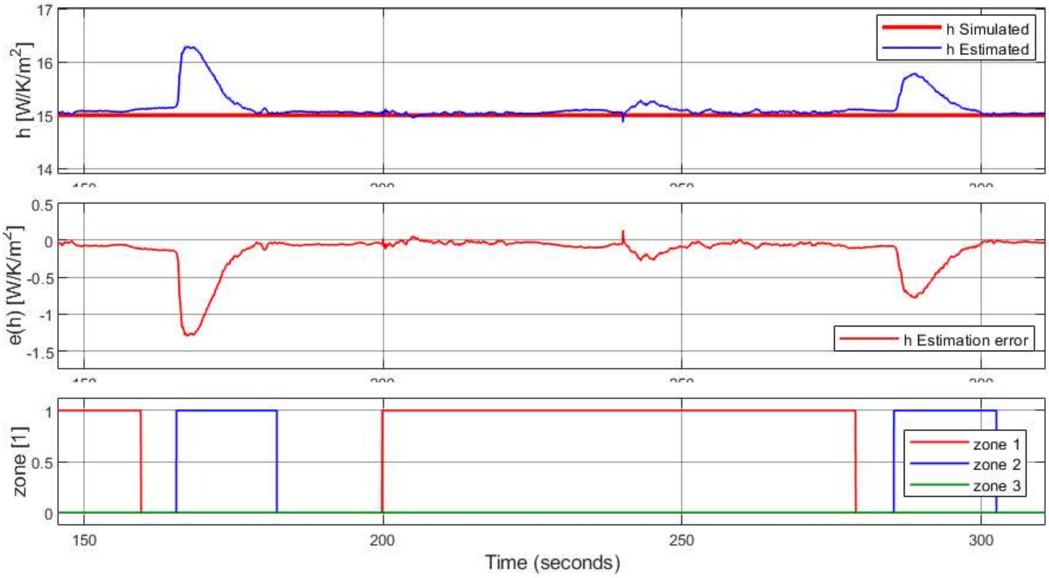


Figure 83 – Convective heat exchange coefficient zoom on zone 1 with $h_{0f} = h_0 - 10$ W/m^{2}K and $\sigma_{0f} = \sigma_0 - 20$ MPa

Finally, in Figure 83 we see h in the zone of interest, here the maximum error is again bound by 0.2 W/K/m^2.

8.3.4 Test with $h_{0f} = h_0 - 15$, $\sigma_{0f} = \sigma_0 + 10$

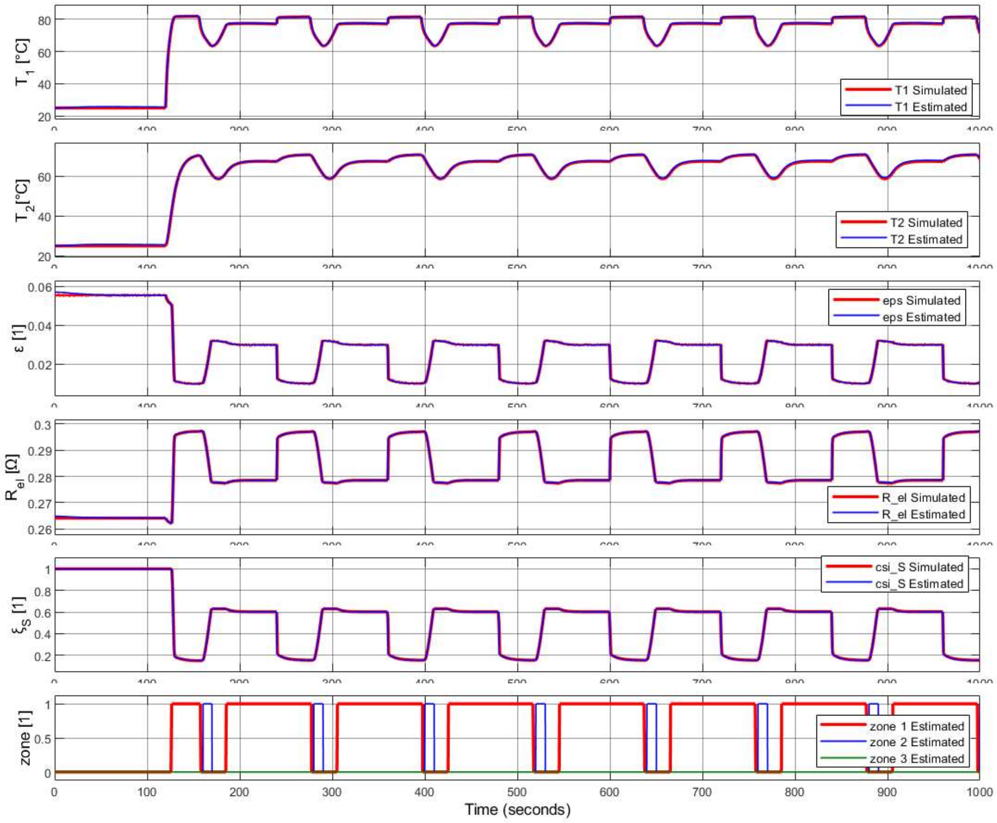


Figure 84 – State variables estimation on closed loop system with $h_{0f} = h_0 - 15$ W/m²K and $\sigma_{0f} = \sigma_0 + 10$ MPa

Figure 84 shows that the estimation of the state variables of the closed loop system is coherent with the simulated variables.

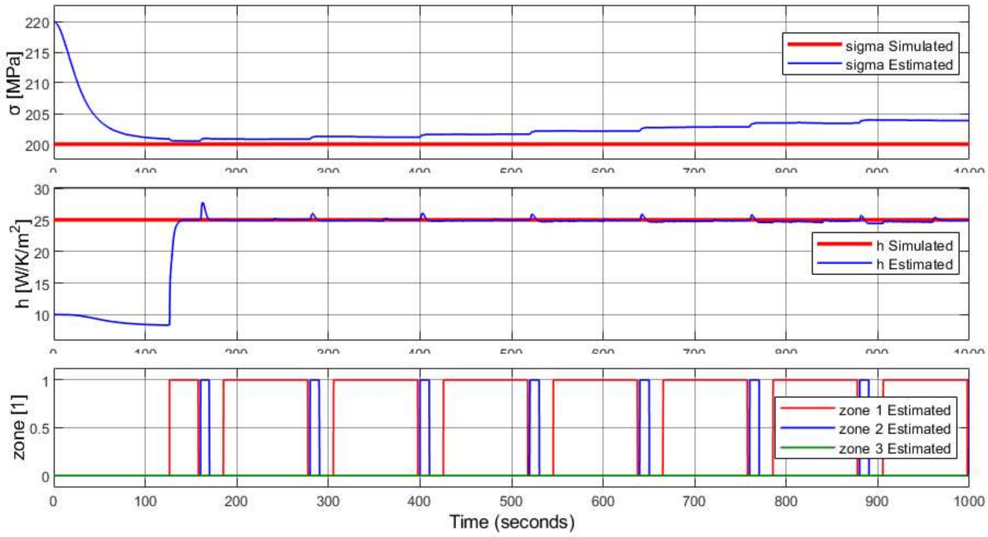


Figure 85 – Unknown parameter estimation on closed loop system with $h_{0f} = h_0 - 15$ W/m²K and $\sigma_{0f} = \sigma_0 + 10$ MPa

In Figure 85 we observe the estimated unknown parameters for this case.

After the initial transitory the estimated σ has mean value 202.21 MPa, standard deviation 1.093 MPa and its estimation error has maximum absolute value 4.00 MPa; while the estimated h has mean value 24.89 W/K/m^2, standard deviation 0.272 W/K/m^2 and maximum error absolute value 2.77 W/K/m^2.

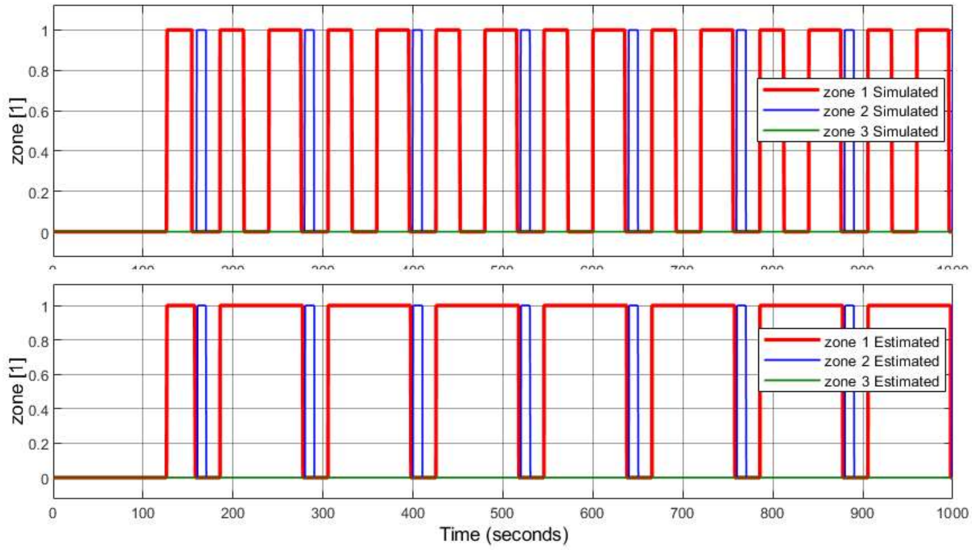


Figure 86 – Transformation zones estimation on closed loop system with $h_{0f} = h_0 - 15$ W/m^{2}K and $\sigma_{0f} = \sigma_0 + 10$ MPa

In Figure 86 we see the simulated and estimated transformation zones, here are valid the same consideration done for Figure 65 in section 8.3.1.

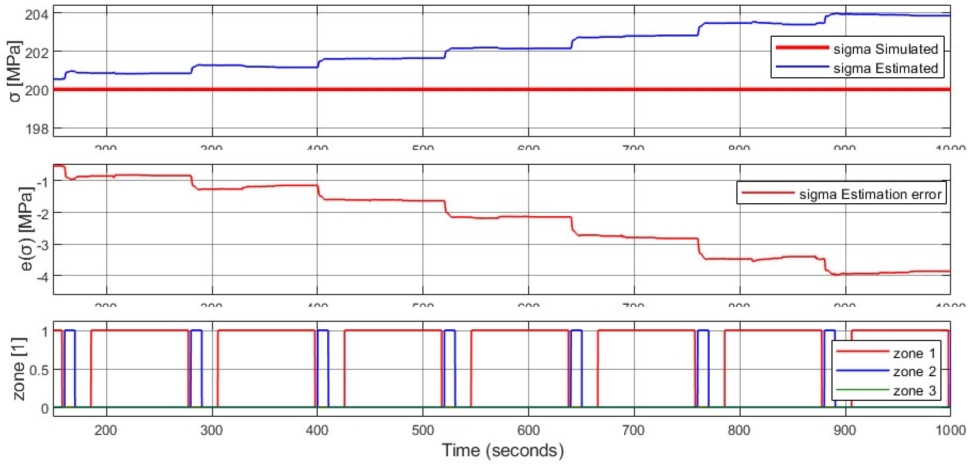


Figure 87 – Wire stress estimation after transient with $h_{0f} = h_0 - 15$ W/m^{2}K and $\sigma_{0f} = \sigma_0 + 10$ MPa

In Figure 87 we see the estimated stress after the error correction.

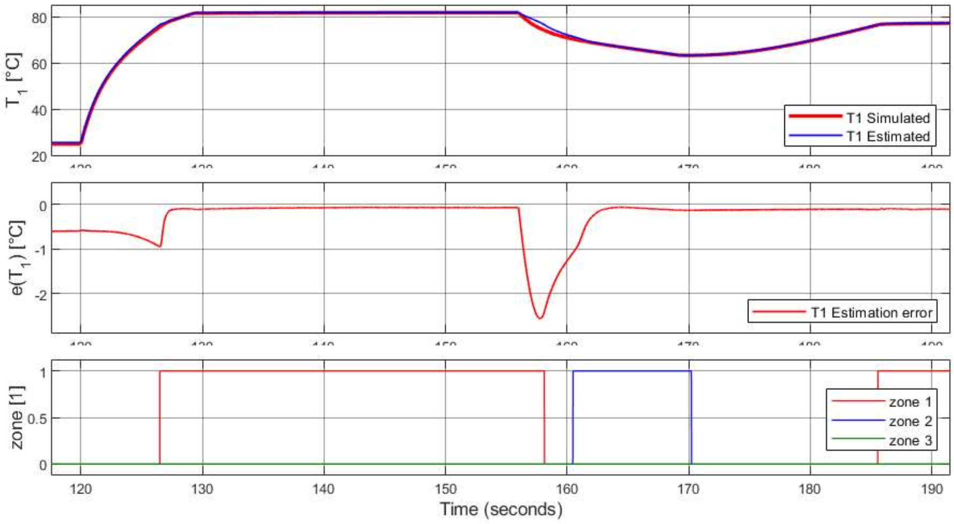


Figure 88 – Wire temperature estimation zoom on zone 1 with $h_{0f} = h_0 - 15$ W/m^{2}K and $\sigma_{0f} = \sigma_0 + 10$ MPa

In Figure 88 we see the estimated wire temperature and its estimation error in the zone of interest for the proposed controller, where the maximum error is 2.6 °C.

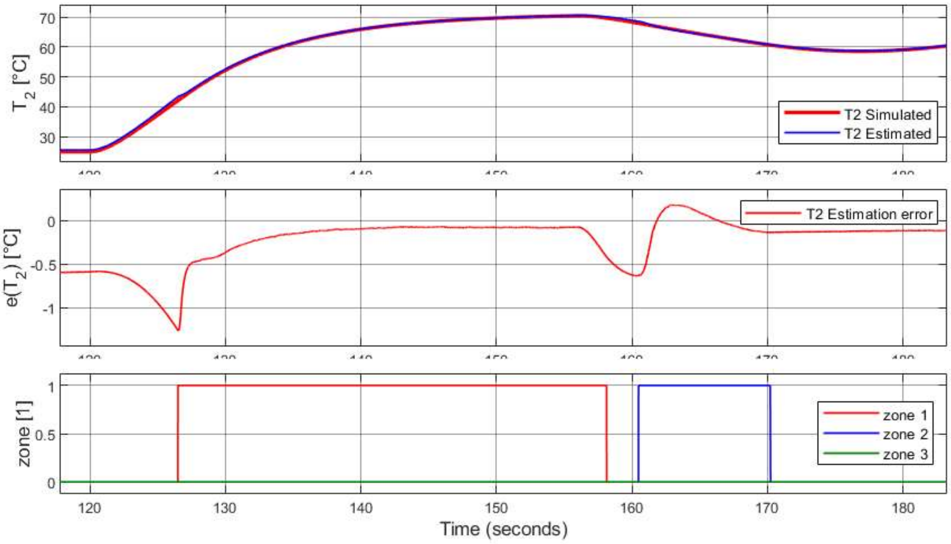


Figure 89 – Silicone temperature estimation zoom on zone 1 with $h_{0f} = h_0 - 15$ W/m^{2}K and $\sigma_{0f} = \sigma_0 + 10$ MPa

In Figure 89 we can see the estimated silicone temperature, its maximum estimation error in the zone of interest for the proposed controller is bound by 1.3 °C.

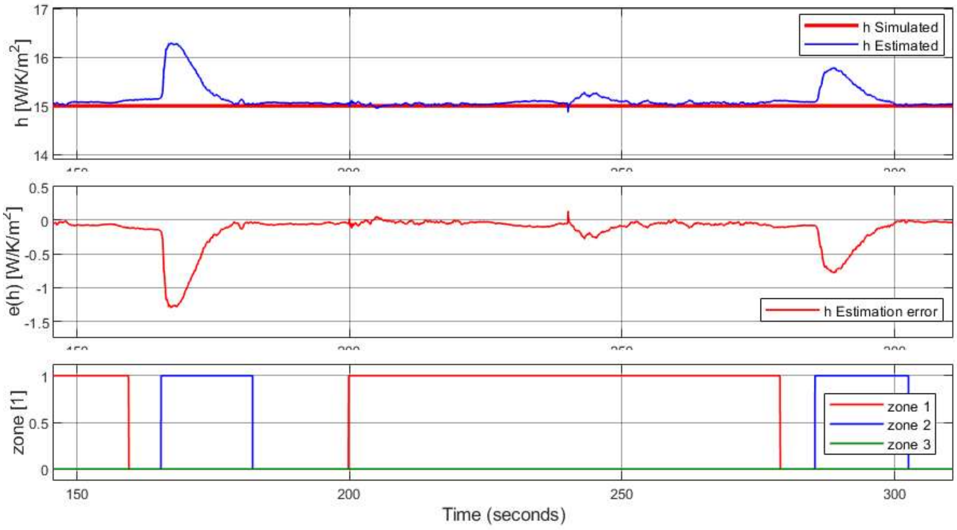


Figure 90 – Convective heat exchange coefficient estimation zoom on zone 1 with $h_{0f} = h_0 - 15$ W/m^{2}K and $\sigma_{0f} = \sigma_0 + 10$ MPa

Finally, in Figure 90 we see h in the zone of interest, where the maximum error is bound by 0.2 W/K/m^2 at 240 s, and peaks outside of it.

8.3.5 Parameter estimation sensitivity on the initial delay

During our initial tests, we noticed that the parameter estimation yielded unsatisfactory results when given wrong initial parameters. This is due to the difficulty in correcting a wrong initial stress value.

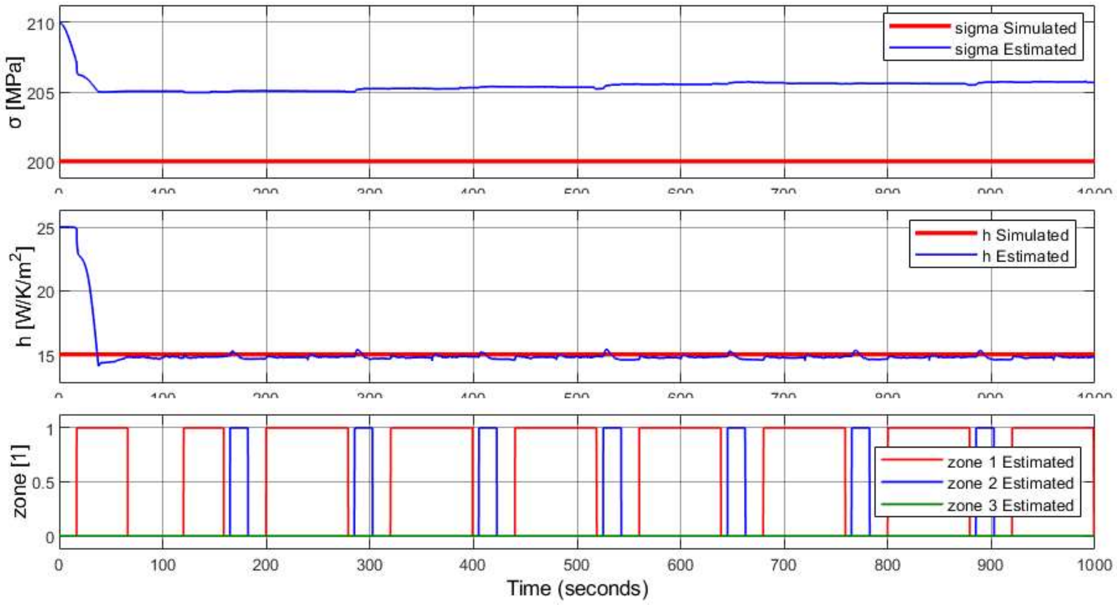


Figure 91 - Unknown parameter estimation on closed loop system with $h_{0f} = h_0 - 15$ W/m²K and $\sigma_{0f} = \sigma_0 + 10$ MPa with 10 s parameter conditioning

In Figure 91 we see that the stress presents a constant offset, which diminishes at the beginning of the actuation but then stays almost constant when the transformations start; this is due do the fact that the EKF can fix the error on the initial condition of σ when there is no transformation because the system equations become linear; on the other hand the value of h can be fixed only during the transformations, because the state variables depend on it only through the variation of T_2, which starts to vary only after the first cycle.

To improve the EKF performance, we defined a parameter conditioning phase before the actuation, where no change occurs in the system and the filter can correct the estimation of σ and later of h, during the phase change.

In Figure 92 we gave a long initial conditioning delay of 200 s.

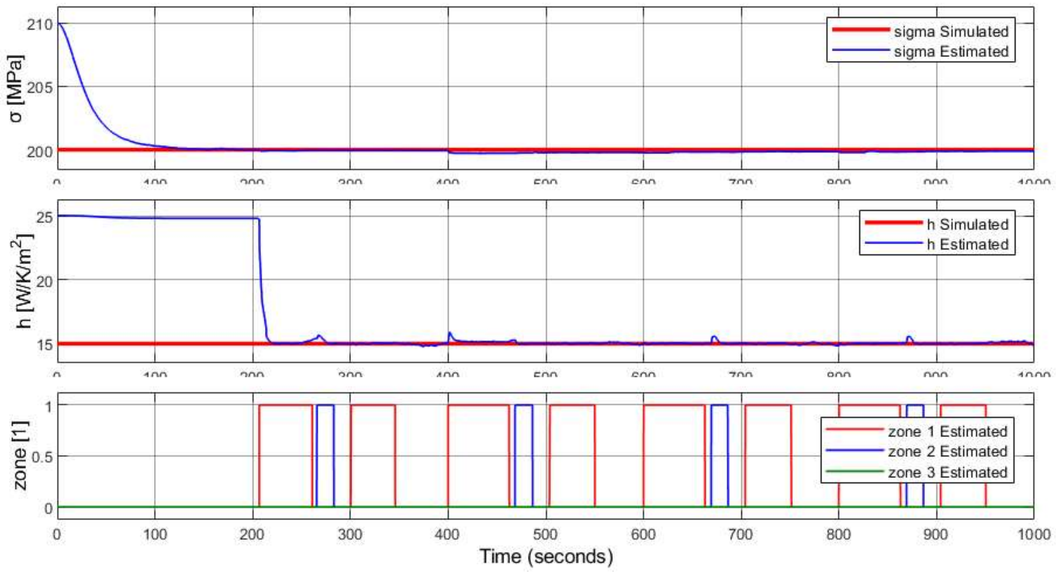


Figure 92 - Unknown parameter estimation on closed loop system with $h_{0f} = h_0 – 15$ W/m^{2}K and $\sigma_{0f} = \sigma_0 + 10$ MPa without 200 s parameter conditioning

We can see that the stress estimation drastically improved but having to wait 200 s before actuation could be impractical. After some testing, we decided to adopt an initial delay of 120 s, which is the one we used for all the tests presented in the following chapters.

Figure 93 below is identical to Figure 64 but we copied it here to better appreciate the stress correction compared to Figure 91.

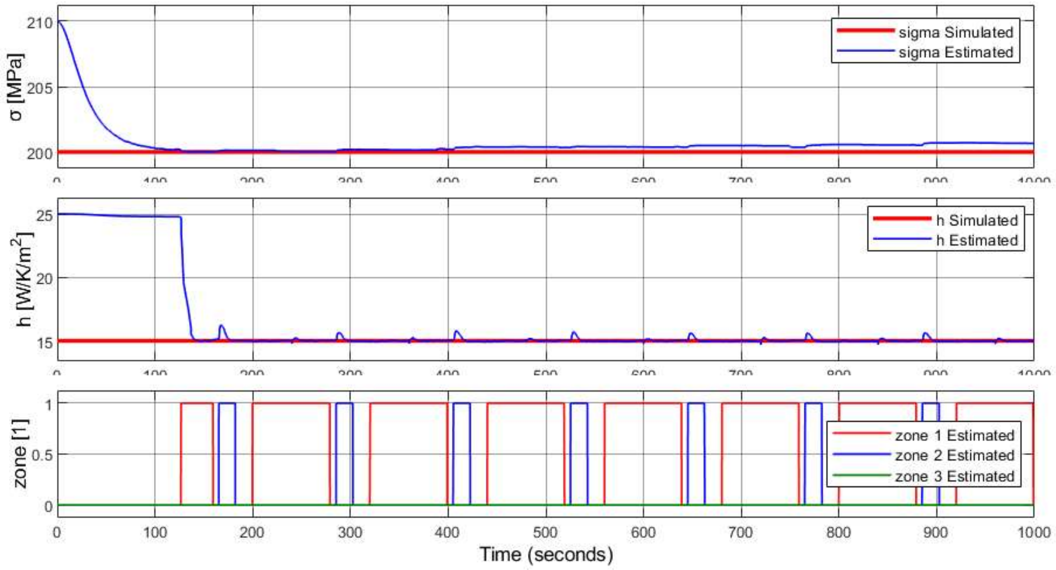


Figure 93 - Unknown parameter estimation on closed loop system with $h_{0f} = h_0 – 15$ W/m^{2}K and $\sigma_{0f} = \sigma_0 + 10$ MPa without enough parameter conditioning

9 Preliminary controller implementation and tests

In this chapter we provide some initial results and implementation of a state variable dependent controller introduced in chapter 6.

The preliminary nature of the current implementation depends on the fact that the rules for PID gain modulation were established empirically based on our observations of the controlled system behavior in the different zones. With respect to the zones already defined for the implementation of the EKF, it was found useful to refine the zone definitions by adding a new set of conditions to identify a region of "holding position" as opposed to just heating or cooling. In this thesis the holding zone was implemented only for the heating ("Austenite holding") but a similar zone could be defined for the cooling direction as well.

9.1 Variable PID block implementation

This block contains the PID gains for each zone, which are contained inside enabled Simulink subsystems, each activated by a different control zone signal; the outputs of all these subsystems are then connected to a Merge block that selects the active output.

Here are the PID gain associated to each control zone:

- **Zone a – Austenite heating**

  In this zone we want the controller to arrive quickly at the reference possibly with a little overshoot, this overshoot creates a small excess of austenite, which in turn will make the wire shorter and thus it will be possible to reach the reference strain without maintaining the wire in the transformation zone and saving up some energy.

  The gains are $K_P = -1000$, $K_I = -200$ and $K_D = 0$, the integral action is saturated between 0 and -10, to avoid overshoots due to the integration of a positive error since our controller can only work by shortening the wire heating it, thus only on negative errors.

  The PID coefficient are helped by additional terms depending on the state variables.

  The integral action is linearly modulated by the estimated convective heat exchange coefficient, it is $K_I = -200$ with high convection and $K_I = -190$ with low convection, this is because the control action needs to be weaker with low thermal exchange with the environment, when more energy is stored inside the silicone.

  Both the K_P and K_I coefficients are also modulated by T_2 for a similar reason, that is when the silicone is hotter the control action can be weaker, to save energy, thus the coefficient are multiplied by 1 when $T_2 = 90$ °C and by 3 when $T_2 = 20$ °C.

- **Zone b – Martensite Stress Cooling**

  Here we have a fast wire elongation due to the martensitic transformation that happens while cooling. The scope is to slow the elongation in order to limit the overshoot. The gains $K_P = -300$, $K_I = -20$ and $K_D = 0$, the integral action is saturated between 0 and -10. Until the strain error become negative the controller remains in standby, and then the voltage is gradually increased to stop the elongation.

- **Zone c – Martensite Thermal Cooling**

  In this zone the behavior is the same as for Zone b.

- **Zone d – Free heating**

  Here the system changes its length only because of thermal elongation, thus the control action can be very high in order to reach zone 1 as fast as possible thus the controller reduces to a P controller with $K_P = -10000$ to ensure always the maximum voltage.

- **Zone e – Free cooling**

Here as well the wire length varies only because of thermal elongation. The gains are $K_P = -200$, $K_I = -20$ and $K_D = 0$, the integral action is saturated between 0 and -10. With this gains the controller can act correctly in case of small changes in the reference. In case of larger changes in the reference, the error is always positive, and the controller in inactive permitting the system to reach zone 2 or 3 as fast as possible. With the control strategy implemented, and in particular the introduction of the Holding Zone h, the initial temperature is generally quite close to the boundaries of Zones b or c. Hence the permanence in Zone e is limited to a short time after exiting the heating or holding phases.

- **Zone f – 100% Austenite**

  The controller is in this zone when the martensitic fraction is 0 but the reference strain is still smaller than the present one, thus the controller would continue to inject electrical power into the wire indefinitely if this zone was treated as the heating free zone and the reference could not be reached. To avoid this, we impose smaller PID gains and we stop the control action when the system reaches a maximum temperature T_{MAX}.

- **Zone g - Martensite 100% zone**

  Here the martensitic fraction is 1 but the wire needs to stretch even further to reach the reference strain.

  The prolonged presence of the controller in this zone could mean that the reference strain is too large to be reached in reasonable time by letting the wire cool naturally or that the required wire temperature is below ambient temperature.

  To avoid the last case, we impose a threshold strain ε_{MAX} above which the controller starts to give again electrical power to stop the elongation of the wire.

- **Zone h – Austenite Holding**

  This zone holds the strain reference, it is activated when the system is in Zone a and the strain error reaches 0, we exit from this zone when the reference strain changes or Zones b or c are active.

  This zone shares the same modulation of K_P based on the state variables as Zone a plus inversely proportionally on the distance between T_{AUST} and $M_S - 2$, multiplied by $K_T = 0.5$, the integral and derivative parts are absent.

9.2 Test protocols for the control system

In the literature for SMA systems control we find many recurring wave forms used to check the controller performance, thus we decided accordingly for our tests on the new controller, keeping in mind the operating conditions, which in a real scenario would cause plasticization and damage irreversibly the SMA actuator.

We will test the proposed controller tracking a square strain reference, with period 60s, on time 20s and amplitude between 1% and 3% of strain, stress 200 MPa and convective heat coefficient 35 W/K/m^2, and a random step strain reference within the same range of strains and stress but convective heat coefficient 20 W/K/m^2.

9.3 Preliminary results of the proposed controller performance

In this section we see some introductory results of our proposed modified PID controller.

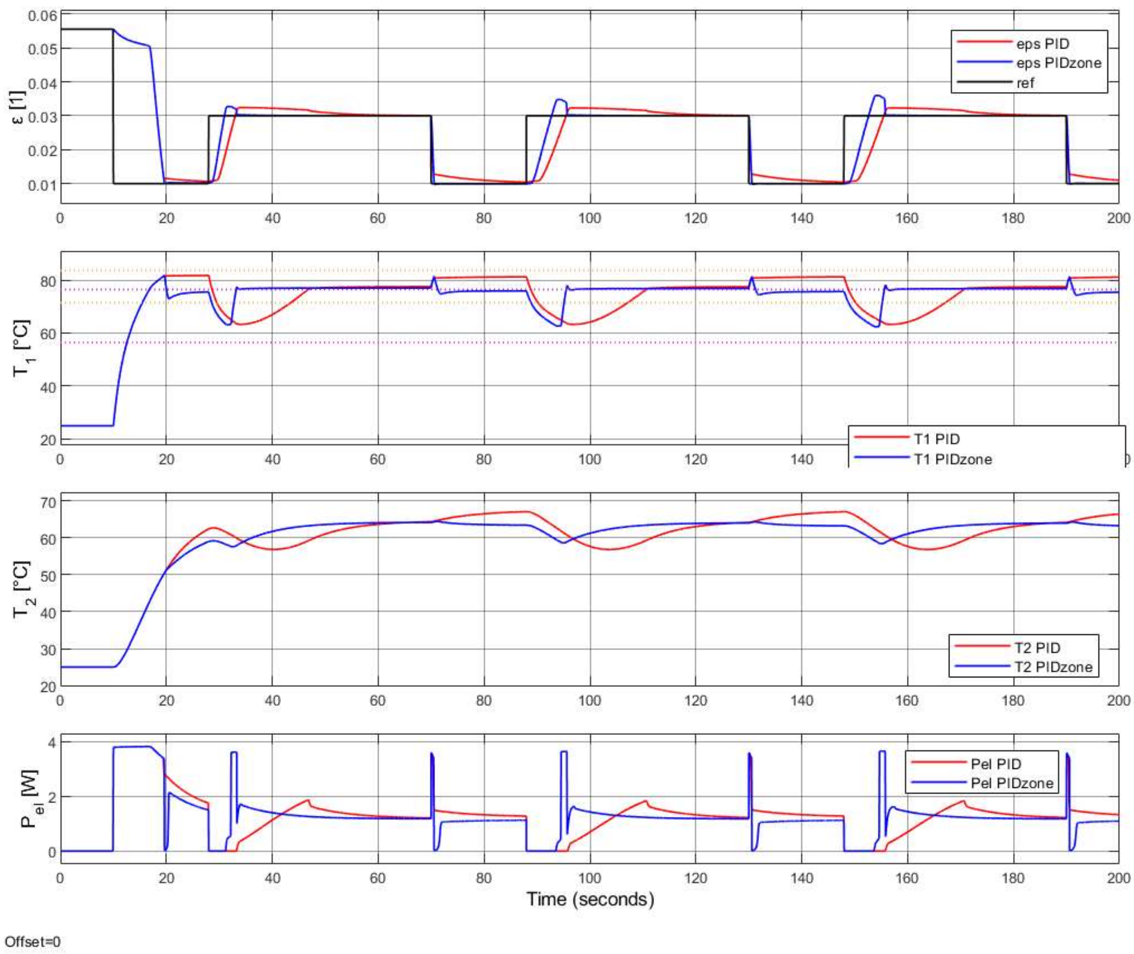


Figure 94 – Performance of the proposed controller with square reference, σ = 200MPa and h = 35 W/K/m^2

In Figure 94 we see the variable PID response to a square wave of amplitude 2% of strain.

During the first heating phase, the two controllers have a similar performance, while on the following heating zones our proposed controller is slightly faster than the reference PID and manages to converge more quickly to the reference and most importantly manages to keep the wire and silicone cooler; this is crucial during the cooling phase, where our variable PID controller starts before the reference one.

At the end of the cooling phase we observe an unwanted overshoot, after which the system quickly converges to the strain reference, faster than the reference controller.

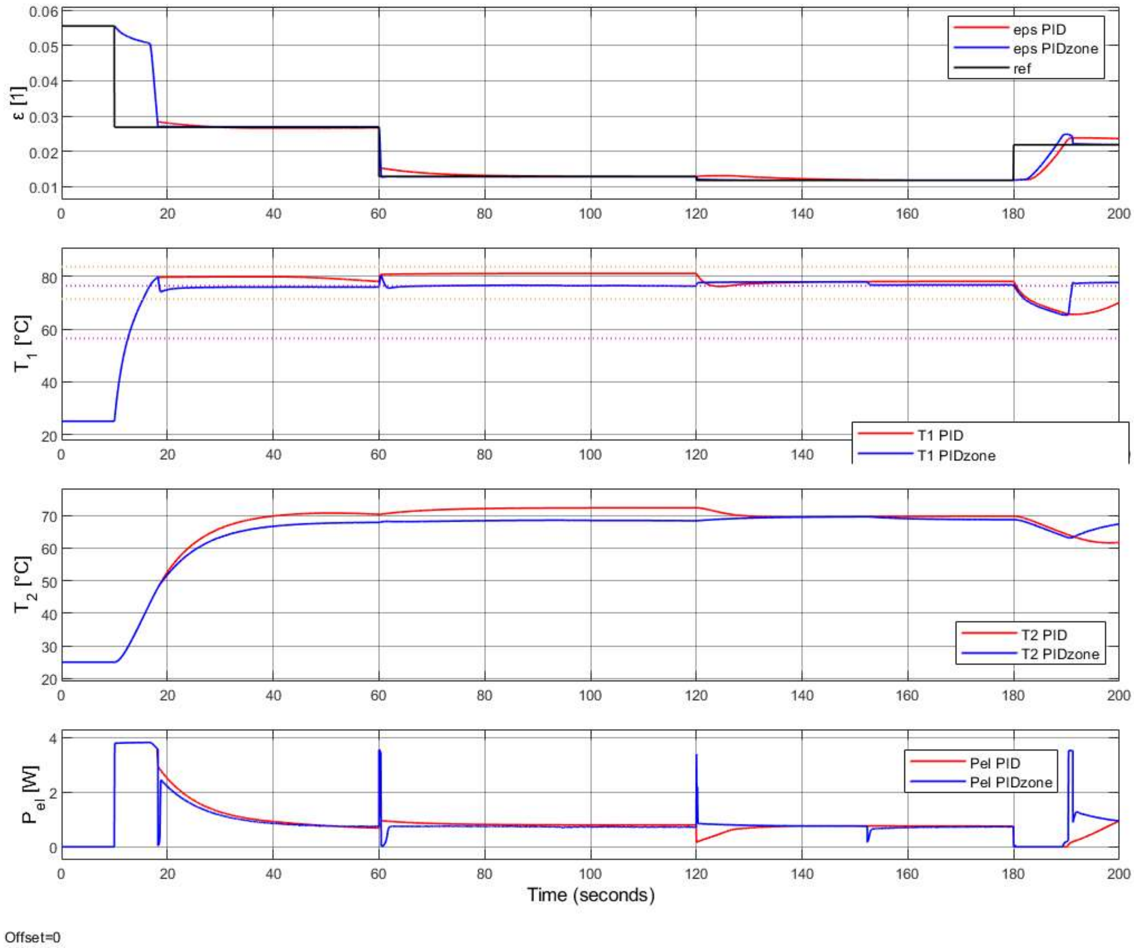


Figure 95 – Performance of the proposed controller with random steps reference, σ = 200MPa and h = 20 W/K/m²

In Figure 95 we can observe the tracking of random strain steps.

It can be shown that our controller is faster than the reference one both during the heating phase, without causing large overshoots, and the cooling phase, thanks to the lower temperatures.

In the third random step, we see that at the beginning the two controllers maintain the wire at the same temperature, however in the last part of the step our controller manages to maintain the reference while lowering the temperatures with respect to the reference controller.

10 Discussion and Conclusions

In this last chapter we discuss the obtained results and draw conclusions about our work.

10.1 Discussion on the modified and extended SMA model

In this thesis, we designed a model of the SMA thermomechanical behavior based on the model from (Brinson, 1993), which was selected since it is already widely used, with plenty of examples from the literature, and one of the most complete phenomenological models.

We chose a phenomenological model because our goal of developing an observer for control application requires a relatively simple model and we need a good macroscopic description of the austenitic and martensitic phase changes.

We did some modifications to the original model to include the main effect of TWSME on wire length when cold (partition coefficient between twinned and detwinned Martensite at low temperatures); fit better the SMA behavior at low stresses and temperatures (splitting of zone 5); include partial cycles and avoid discontinuities when changing transformation zone (updating of the variables at transformation onset).

We were able to adjust such wire model parameters to replicate measured curves of complete heating cycles from experiments on the real wire; future works could also compare the performance of the model with partial heating cycles to improve their modeling.

Finally, an important and innovative addition to the model was the coupling with a second material, i.e., a silicone rubber sheath implementing a surface liner for the wire. We modeled though suitable lumped parameter heat transfer equations the relationship between the wire and silicone temperatures, influenced by Joule heating, latent heat of transformation and heat exchange between wire and silicone and the environment.

A comparison of simulations of the bare and silicone-lined wire show differences in behavior, which can be explained considering the heat storage and transport properties of the silicon rubber sheath. This comparison evidenced opportunities for the development of a controller able to exploit the beneficial effect of the cool silicone mass on wire cooling, and to conserve a low temperature of the silicone during repeated actuation cycles.

10.2 Discussion on the behavior of the hybrid actuator

In subchapter 5.4.4 we compared the heat exchange of a naked and coated wire, showing that the presence of the silicone rubber can improve the controller performance in the cooling phase, as long as the silicone temperature T_2 is kept low.

To this end, it is very important to have an energy-efficient control, which is one of the main objectives we tried to obtain with the preliminary controller from chapter 6. In particular, this observation led to the introduction of a special control zone (Zone h) for the position holding phases.

10.3 Discussion on the EKF performance

The introduction of a state estimator, in our case an Extended Kalman Filter, was proposed in order to be able to implement a variable controller, exploiting the knowledge about the working zone, as well as wire and silicone temperatures during operation. This approach proved feasible, based on the ability of the EKF to correct for imprecise initial values of external stress and convection, and to provide estimates about the current working conditions.

The first tests on the EKF helped us verify its performance in ideal conditions (correct initial values, full cycles) and its reactivity, since we were supplying the simulated wire with the maximum voltage.

These tests confirmed the correctness of the EKF formulation. We also observed that the biggest estimation errors are found in correspondence of the zone commutations.

With the second set of tests, we checked the handling of more complicated zone commutations by the EKF, in different stress and convection conditions, with associated closed loop control of the wire, and partial cycles. They demonstrated that the system can always estimate correctly inside the operating conditions.

During our test we found cases in which the zones where wrongly estimated. An example is displayed in Figure 59, where we observe that in some instances the zones are prolonged, while in others they are interrupted and reentered.

These cases are due to the correction of the estimated values by the EKF, which can sometimes wrongly modify the values of the other variables, like T_1, which in turn affect the zone estimation.

When the estimated zone is prolonged with respect to the simulated one, it means that the variation of T_1 is very small or null. Since the EKF is less sensitive to small variations with respect to our model, the zone does not change but at the same time no transformation happens, thus no unexpected change in the variables and no further estimation error is added.

Instead, when the estimated zone is prematurely interrupted, it increases the estimation errors on T_1 and σ, which in turn are quickly corrected by the EKF by reentering the correct transformation zone. This can be seen because the incorrect interruption of a zone is always followed by its quick recovery and no zone chattering is observed.

Finally, the tests in section 8.3 show how the EKF would behave with a realistic system, where the information on the initial state could be inaccurately estimated or measured. They demonstrated the robustness of the EKF to this type of error. In particular, we found that when the initial value of σ was bigger than the real one, the stress estimate converged rapidly to the correct value only to depart from it soon afterwards and accumulate larger drifts. On the other hand, when the initial value is smaller than the real one, the EKF takes more time to converge to the correct value, but the convergence is then stable. It is therefore better to underestimate the initial stress, in order to reduce the estimation error.

The convective heat transfer coefficient h has a higher variability than the other states, however its influence on the other states is very limited, it is much more important to correctly estimate σ.

In the end, we showed *a posteriori* the importance of the initial delay for the correction of the error on the initial state σ_0; we also observed that the system adjusts the estimation of h in the transformation zone only and produces the biggest errors on h during free cooling.

10.4 Preliminary observations on the proposed controller

In section 9.1 we introduced and explained a preliminary idea of how to improve a PID controller with our estimated state variables. Briefly, we intend to use knowledge about the working zones, and the estimated state to modulate the value of the PID controller, effectively implementing a variable PID. The most important state variables in this respect appear to be the wire temperature, the silicone temperature, the convective coefficient, besides the current working zone.

In the light of this, in section 8.3 we showed in some detail the behavior of the selected state variables and we observed that these variables are estimated with negligible errors, thus we can be confident on their values.

The images of section 9.3 show some preliminary tests of this controller.

During the heating phase the variable PID can be more aggressive than the simple PID and reach close to the reference faster. Once reached, the Austenite holding zone starts, which through a control loop closed around T_1, is able to keep the temperatures of the system lower than the reference PID.

This temperature control is done to avoid transforming back into Martensite and to keep le temperature T1 as low as possible, holding the position but ready to cool rapidly below M_S if necessary, this fact is of capital importance, since it could allow us to save energy and be quicker on reaching the martensitic transformation in the next cooling phase.

A similar separated holding zone with a temperature control for the martensitic cooling could avoid the overshoot after the cooling phase and limit the requested energy to maintain the reference; this could be the subject of future work.

10.5 Conclusions

In this thesis, we implemented the Brinson model for the thermomechanical behavior of a SMA and a lumped parameter dynamic model of the heat exchange with the silicone, obtaining a Simulink model which allowed us to simulate the behavior of the real wire actuator, and to conduct meaningful *in silico* experiments without a physical set-up.

To model the SMA hysteretic and conditional behavior in Simulink, we introduced a zone selection block which changes the model equations based on the current wire temperature, stress and their derivatives.

We also produced a state observer for the model in the form of an EKF, which is able to estimate reliably all the unmeasured state variables and at the same time estimate the unknown parameters wire stress σ and convective heat coefficient h, despite the complexity of the plant, including the hysteresis and the conditional behavior of the model.

In parallel with the EKF, works another zone selection block, based on the estimated values of T_1 and σ, which is fundamental for the implementation of the proposed Variable PID controller scheme.

Finally, we discussed how this added knowledge of the system could be used for improving a model-based controller especially with the future aim to reduce the requested electrical power for the actuation, efficiency being one of the biggest limiting factors for SMA based actuators. We were able to provide preliminary example of the proposed controller implementation and testing.

Future research could investigate further model-based control methods in conjunction with the proposed EKF.

References

Ashrafiuon, H., & Elahinia, M. (2015). Control of SMA actuators. In *Shape Memory Alloy Actuators: Design, Fabrication, and Experimental Evaluation.*

Auricchio, F., Marfia, S., & Sacco, E. (2003). Modelling of SMA materials: Training and two way memory effects. *Computers and Structures, 81*, 2301-2317.

Brinson, L. (1993). One-Dimensional Constitutive Behavior of Shape Memory Alloys: Thermomechanical Derivation with Non-Constant Material Functions and Redefined Martensite Internal Variable. *JOURNAL OF INTELLIGENT MATERIAL SYSTEMS AND STRUCTURES, 4*, 229-242.

Cisse, C., Zaki, W., & Ben Zineb, T. (2016). A review of constitutive models and modeling techniques for shape memory alloys. *International Journal of Plasticity, 76*, 244-284.

Copaci, D., Blanco, D., & Moreno, L. (2019). Flexible Shape-Memory Alloy-Based Actuator: Mechanical Design Optimization According to Application. *Actuators, 8*(3).

Daohui, Z., Xingang, Z., Jianda, H., Xiaoguang, L., & Bi, Z. (2019). Active Modeling and Control for. *IEEE Access, 7*.

Elahinia, M., & Ashrafiuon, H. (2002). Nonlinear Control of a Shape Memory Alloy Actuated Manipulator. *J. Vib. Acoust., 124*(4), 566-575.

Elahinia, M., Esfahani, E., & Wang, S. (2010). Control of SMA systems: Review of the state of the art. *Shape Memory Alloys: Manufacture, Properties and Applications*, 49-68.

Franceschinis, S., & Vergani, M. (2015). *Modeling of Shape Memory Alloy components with application to a micro-actuator.* Politecnico di Milano, Scuola di Ingegneria Industriale e dell'Informazione.

Fried, E., & Gurtin, M. (1994, May). Dynamic solid-solid transitions with phase characterized by an order parameter. *Physica D (Nonlinear Phenomena), 72*(4), 287-308.

Ivshin, Y., & Pence, T. (1994, April). A constitutive model for hysteretic phase transition behavior. *International Journal of Engineering Science, 32*(4), 681-704.

J.B., H., & J.W., L. (2017). Finite temperature properties of NiTi from first principles simulations: Structure, mechanics, and thermodynamics. *Journal of Applied Physics, 121*(20).

J.C., C., R., G., K., O., & D., B. (2019). Advances in Selective Laser Melting of Nitinol Shape Memory Alloy Part Production. *Materials, 12*(5).

Jani, J., Leary, M., Subic, A., & Gibson, M. (2014). A review of shape memory alloy research, applications and opportunities. *Materials and Design, 56*, 1078-1113.

Khandelwal, A., & Buravalla, V. (2009, December). Models for Shape Memory Alloy Behavior: An overview of modeling approaches. *International journal of structural changes in solids, 1*(1), 111-148.

Kheirikhah, M., Samaneh, R., & Edalat, M. (s.d.). A Review of Shape Memory Alloy Actuators in Robotics. *Lecture Notes in Computer Science, 6556*, 206-217.

Kim, H., Han, M., Song, S., & Ahn, H. (2016). Soft morphing hand driven by SMA tendon wire. *Composites Part B, 105*, 138-148.

Lange, G., Lachmann, A., Rahim, A., Ismail, M., & Low, C. (2015). Shape Memory Alloys as Linear Drives in Robot Hand Actuation. *Procedia Computer Science, 76*, 168-173.

Lee, J., Chung, Y., & Rodrigue, H. (2019). Long Shape Memory Alloy Tendon-based Soft Robotic Actuators and Implementation as a Soft Gripper. *Scientific Reports, 9*.

Lee, J., Jin, M., & Ahn, K. (2013). Precise tracking control of shape memory alloy actuator systems using hyperbolic tangential sliding mode control with time delay estimation. *Mechatronics, 23*(3), 310-317.

Lee, S., & Kim, S. (2019). Improved position control of shape memory alloy using the self-sensing model. *Sensors and Actuators A: Physical, 297*.

Liang, C., & Rogers, C. (1997, April). One-dimensional thermomechanical constitutive relations for shape memory materials. *Journal of Intelligent Materials Systems and Structures, 8*(4), 285-302.

materials, A. (s.d.). *Silicone rubber.* Tratto da https://www.azom.com/properties.aspx?ArticleID=920

Mattey, J. (s.d.). *Nitinol technical properties*. Tratto da https://matthey.com/en/products-and-services/medical-components/resource-library/nitinol-technical-properties

Ortin, J. (1992). Preisach modeling of hysteresis for a pseudoelastic Cu-Zn-Al single crystal. *Journal of Applied Physics, 71*(1454).

Ortin, J., & Delaey, L. (2002, December). Hysteresis in shape-memory alloys. *International Journal of Non-Linear Mechanics, 37*(8), 1275-1281.

Otsuka, K., & Ren, X. (2005). Physical metallurgy of Ti–Ni-based shape memory alloys. *Progress in Materials Science, 50*, 511-678.

Pause, B. (2007). Application of phase change and shape memory materials in medical textiles. In *Smart Textiles for Medicine and Healthcare* (p. 74-87). Woodhead Publishing.

Pittaccio, S., & Garavaglia, L. (2014). Electric resistance monitoring as a method for controlling shape memory alloy characteristics during shape-setting treatments in the furnace. *Materials Science and Engineering: A, 599*, 92-104.

Rodrigue, H., Wang, W., Bhandari, B., Han, M., & Ahn, S. (2014). Cross-Shaped Twisting Structure Using SMA-Based Smart Soft Composite. *INTERNATIONAL JOURNAL OF PRECISION ENGINEERING AND MANUFACTURING-GREEN TECHNOLOGY, 1*(2), 153-156.

Rodrigue, H., Wang, W., Kim , D., & Ahn, S. (2017). Curved shape memory alloy-based soft actuators and application to soft. *Composite Structures, 176*, 398-406.

Sayyadi, H., Zakerzadeh, M., & Salehi, H. (2012, April). A comparative analysis of some one-dimensional shape memory alloy constitutive models based on experimental tests. 249-257.

Scarpa, F., & De Rosa, M. (2016, August). Transient heat conduction in wires with heat sources; lumped and distributed solution techniques. *Heat Transfer Research, 47*(8), 753-765.

Shameli, E., Alasty, A., & Salaarieh, H. (2005). Stability analisis and nonlinear control of a minature shape memory alloy for precise applications. *Mechatronics, 15*, 471-486.

Shaw, J. (s.d.). *TIPS AND TRICKS FOR CHARACTERIZING SHAPE MEMORY ALLOY WIRE: PART 1—DIFFERENTIAL SCANNING CALORIMETRY AND BASIC PHENOMENA*. Tratto da https://onlinelibrary.wiley.com/doi/10.1111/j.1747-1567.2008.00410.x

Smith, R. (2005). *Smart Material Systems Model developement.* SIAM.

Srinivasa, A., & Rajagopal, K. (1999). On the thermomechanics of shape memory wires. *Zeitschrift für angewandte Mathematik und Physik, 50*, 459-496.

Tanaka, K. (1986). A thermomechanical sketch of shape memory effect: one dimensional tensile behavior. *Res Mechanica, 18*(3), 251-263.

Villoslada, A., Flores, A., Copaci, D., Blanco, D., & Moreno, L. (2015). High-displacement flexible Shape Memory Alloy actuator for soft. *Robotics and Autonomous Systems*(73), 91-101.

Zak, A., Cartmell, M., Ostachowicz, W., & Wiercigroch, M. (2003). One-dimensional shape memory alloy models for use with reinforced composite structures. *Smart Materials and Structures, 12*, 338-346.

A. Appendix – EKF equations

Here are presented all the equations used in the formulation of our EKF.

A.1 Filter equations

In this section the state variable equations of our EKF are shown, in order to shorten the equations, we left some terms implicit.
First, here are the continuous time equations.

$$\dot{T}_1 = \frac{1}{m_1 c_1}\left(\frac{V^2}{R_{el}(\xi)} - \frac{T_1 - T_2}{R_{t1}} + m_1 L_h \dot{\xi}\right) \tag{A.1}$$

$$\dot{T}_2 = \frac{1}{m_2 c_2}\left(\frac{T_1 - T_2}{R_{t1}} - \frac{T_2 - T_\infty}{R_{t2}(h)}\right) \tag{A.2}$$

$$\dot{\xi}_S = \Xi_{ST}(T_1, \sigma)\dot{T}_1 + \Xi_{SNT}(T_1, \sigma) \tag{A.3}$$

$$\dot{\xi}_T = \Xi_{TT}(T_1, \sigma)\dot{T}_1 + \Xi_{TNT}(T_1, \sigma) \tag{A.4}$$

$$\dot{\varepsilon} = \frac{\dot{\sigma} + (D_M - D_A)[\varepsilon_L \xi_S - \varepsilon]\dot{\xi} + \varepsilon_L[D_A + \xi(D_M - D_A)]\dot{\xi}_S - \theta \dot{T}_1}{D_A + \xi(D_M - D_A)} \tag{A.5}$$

$$\dot{R}_{el} = \frac{R_\xi \dot{\xi} + R_\varepsilon \dot{\varepsilon}}{A_0(1 - \nu_{NiTi}\varepsilon)^3} \tag{A.6}$$

$$\dot{\sigma} = 0 \tag{A.7}$$

$$\dot{h} = 0 \tag{A.8}$$

However, the EKF needs discrete time equations to work, thus we discretized them using the forward Euler method.
For simplicity, we kept the same notation for the auxiliary equations.

$$T_1(k+1) = T_1(k) + \frac{t_S}{m_1 c_1}\left(\frac{V^2}{R_{el}(k)} - \frac{T_1(k) - T_2(k)}{R_{t1}} + m_1 L_h \dot{\xi}\right) \tag{A.9}$$

$$T_2(k+1) = T_2(k) + \frac{t_S}{m_2 c_2}\left(\frac{T_1(k) - T_2(k)}{R_{t1}} - \frac{T_2(k) - T_\infty}{R_{t2}(h)}\right) \tag{A.10}$$

$$\xi_S(k+1) = \xi_S(k) + t_S\left[\Xi_{ST}(T_1, \sigma)\dot{T}_1 + \Xi_{SNT}(T_1, \sigma)\right] \tag{A.11}$$

$$\xi_T(k+1) = \xi_T(k) + t_S\left[\Xi_{TT}(T_1, \sigma)\dot{T}_1 + \Xi_{TNT}(T_1, \sigma)\right] \tag{A.12}$$

$$\varepsilon(k+1) = \varepsilon(k) + t_S \frac{\dot{\sigma} + (D_M - D_A)[\varepsilon_L \xi_S - \varepsilon(k)]\dot{\xi} + \varepsilon_L[D_A + \xi(k)(D_M - D_A)]\dot{\xi}_S - \theta \dot{T}_1}{D_A + \xi(k)(D_M - D_A)} \tag{A.13}$$

$$R_{el}(k+1) = R_{el}(k+1) + t_S \frac{R_\xi \dot{\xi} + R_\varepsilon \dot{\varepsilon}}{A_0\left(1 - \nu_{NiTi}\varepsilon(k)\right)^3} \tag{A.14}$$

$$\sigma(k+1) = \sigma(k) \tag{A.15}$$

$$h(k+1) = h(k) \tag{A.16}$$

Where t_s is the sampling time of the simulation.

A.2 Auxiliary equations

Here we define all those auxiliary equations we used in order to keep as compact as possible the final form of the system equations.
Some of the auxiliary equations are valid for all operating conditions while other are only valid when a certain control zone is active.

A.2.1 General auxiliary equations

In this section are shown the auxiliary equations valid for all operating conditions.

$$R_\xi(\varepsilon) = l_0(1+\varepsilon)(\rho_M - \rho_A)(1 - \nu_{NiTi}\varepsilon) \tag{A.17}$$

$$R_\varepsilon(\xi, \varepsilon) = l_0[\rho_A + \xi(\rho_M - \rho_A)](1 + 2\nu_{NiTi} + \nu_{NiTi}\varepsilon) \tag{A.18}$$

$$\frac{\partial R_\xi}{\partial \varepsilon} = l_0(\rho_M - \rho_A)(1 - \nu_{NiTi} - 2\nu_{NiTi}\varepsilon) \tag{A.19}$$

$$\frac{\partial R_\varepsilon}{\partial \xi_S} = l_0(\rho_M - \rho_A)(1 - \nu_{NiTi}\varepsilon)(1 + 2\nu_{NiTi} + \nu_{NiTi}\varepsilon) \tag{A.20}$$

$$\frac{\partial R_\varepsilon}{\partial \xi_T} = l_0(\rho_M - \rho_A)(1 - \nu_{NiTi}\varepsilon)(1 + 2\nu_{NiTi} + \nu_{NiTi}\varepsilon) \tag{A.21}$$

$$\frac{\partial R_\varepsilon}{\partial \varepsilon} = -l_0 2{\nu_{NiTi}}^2[\rho_A + \xi(\rho_M - \rho_A)] \tag{A.22}$$

$$\dot{\xi} = \dot{\xi}_S + \dot{\xi}_T \tag{A.23}$$

A.2.2 Zone specific auxiliary equations.

In this section are shown the auxiliary equations whose validity depends on the EKF zone and the respective zone.
While the controller uses different coefficients for all zones, the EKF is only interested in the transformation zones while treating all the neutral zones in the same way.

- Zone a:

$$\Xi_{ST}(T_1, \sigma) = -\frac{\xi_{S_i} a_A}{2} sin\left\{a_A\left[T_1 - A_S - \frac{\sigma}{C_A}\right]\right\} \tag{A.24}$$

$$\Xi_{TT}(T_1, \sigma) = -\frac{\xi_{T_i} a_A}{2} \sin\left\{a_A\left[T_1 - A_S - \frac{\sigma}{C_A}\right]\right\} \tag{A.25}$$

$$\frac{\partial \Xi_{ST}}{\partial T_1} = -\frac{\xi_{S_i} {a_A}^2}{2} cos\left\{a_A\left[T_1 - A_S - \frac{\sigma}{C_A}\right]\right\} \tag{A.26}$$

$$\frac{\partial \Xi_{TT}}{\partial T_1} = -\frac{\xi_{T_i} {a_A}^2}{2} cos\left\{a_A\left[T_1 - A_S - \frac{\sigma}{C_A}\right]\right\} \qquad \text{A.27}$$

$$\frac{\partial \Xi_{ST}}{\partial \sigma} = \frac{\xi_{S_i} {a_A}^2}{2C_A} cos\left\{a_A\left[T_1 - A_S - \frac{\sigma}{C_A}\right]\right\} \qquad \text{A.28}$$

$$\frac{\partial \Xi_{TT}}{\partial \sigma} = \frac{\xi_{T_i} {a_A}^2}{2C_A} cos\left\{a_A\left[T_1 - A_S - \frac{\sigma}{C_A}\right]\right\} \qquad \text{A.29}$$

$$\Xi_{SNT}(T_1,\sigma) = \Xi_{TNT}(T_1,\sigma) = \frac{\partial \Xi_{SNT}}{\partial T_1} = \frac{\partial \Xi_{TNT}}{\partial T_1} = \frac{\partial \Xi_{SNT}}{\partial \sigma} = \frac{\partial \Xi_{TNT}}{\partial \sigma} = 0 \qquad \text{A.30}$$

- Zone b:

$$\Xi_{ST}(T_1,\sigma) = \frac{(1-\xi_{S_i}) a_\Sigma C_M}{2} \sin\{a_\Sigma[\sigma - C_M(T_1 - M_S) - \sigma_{CR}^F]\} \qquad \text{A.31}$$

$$\Xi_{TT}(T_1,\sigma) = -\frac{\xi_{T_i} a_\Sigma C_M}{2} \sin\{a_\Sigma[\sigma - C_M(T_1 - M_S) - \sigma_{CR}^F]\} \qquad \text{A.32}$$

$$\Xi_{TNT}(T_1,\sigma) = \frac{\xi_{T_i}}{1-\xi_{S_i}} \qquad \text{A.33}$$

$$\frac{\partial \Xi_{ST}}{\partial T_1} = -\frac{(1-\xi_{S_i}) {a_\Sigma}^2 {C_M}^2}{2} cos\{a_\Sigma[\sigma - C_M(T_1 - M_S) - \sigma_{CR}^F]\} \qquad \text{A.34}$$

$$\frac{\partial \Xi_{TNT}}{\partial T_1} = \frac{\xi_{T_i} {a_\Sigma}^2 {C_M}^2}{2} cos\{a_\Sigma[\sigma - C_M(T_1 - M_S) - \sigma_{CR}^F]\} \qquad \text{A.35}$$

$$\frac{\partial \Xi_{ST}}{\partial \sigma} = \frac{(1-\xi_{S_i}) {a_\Sigma}^2 C_M}{2} cos\{a_\Sigma[\sigma - C_M(T_1 - M_S) - \sigma_{CR}^F]\} \qquad \text{A.36}$$

$$\frac{\partial \Xi_{TNT}}{\partial \sigma} = -\frac{\xi_{T_i} {a_\Sigma}^2 C_M}{2} cos\{a_\Sigma[\sigma - C_M(T_1 - M_S) - \sigma_{CR}^F]\} \qquad \text{A.37}$$

$$\Xi_{SNT}(T_1,\sigma) = \frac{\partial \Xi_{SNT}}{\partial T_1} = \frac{\partial \Xi_{TT}}{\partial T_1} = \frac{\partial \Xi_{SNT}}{\partial \sigma} = \frac{\partial \Xi_{TT}}{\partial \sigma} = 0$$

- Zone c:

$$\Xi_{ST}(T_1,\sigma) = -\chi a_{TM} \frac{(1-\xi_{S_i}) - \xi_{T_i}}{2} sin\{a_{TM}[T_1 - M_S]\} \qquad \text{A.38}$$

$$\Xi_{TT}(T_1,\sigma) = -(1-\chi) a_{TM} \frac{(1-\xi_{S_i}) - \xi_{T_i}}{2} sin\{a_{TM}[T_1 - M_S]\} \qquad \text{A.39}$$

$$\frac{\partial \Xi_{ST}}{\partial T_1} = -\chi {a_{TM}}^2 \frac{(1-\xi_{S_i}) - \xi_{T_i}}{2} cos\{a_{TM}[T_1 - M_S]\} \qquad \text{A.40}$$

$$\frac{\partial \Xi_{TT}}{\partial T_1} = -(1-\chi) {a_{TM}}^2 \frac{(1-\xi_{S_i}) - \xi_{T_i}}{2} cos\{a_{TM}[T_1 - M_S]\} \qquad \text{A.41}$$

$$\Xi_{SNT}(T_1,\sigma) = \Xi_{TNT}(T_1,\sigma) = \frac{\partial \Xi_{SNT}}{\partial T_1} = \frac{\partial \Xi_{TNT}}{\partial T_1} = \frac{\partial \Xi_{ST}}{\partial \sigma} = \frac{\partial \Xi_{SNT}}{\partial \sigma} = \frac{\partial \Xi_{TT}}{\partial \sigma} = \frac{\partial \Xi_{TNT}}{\partial \sigma} = 0 \qquad \text{A.42}$$

- Zone d-g:

$$\Xi_{ST}(T_1,\sigma) = \Xi_{SNT}(T_1,\sigma) = \Xi_{TT}(T_1,\sigma) = \Xi_{TNT}(T_1,\sigma) = \frac{\partial \Xi_{ST}}{\partial T_1} = \frac{\partial \Xi_{SNT}}{\partial T_1} = \frac{\partial \Xi_{TT}}{\partial T_1} = \frac{\partial \Xi_{TNT}}{\partial T_1} = \frac{\partial \Xi_{ST}}{\partial \sigma} = \frac{\partial \Xi_{SNT}}{\partial \sigma} = \frac{\partial \Xi_{TT}}{\partial \sigma} = \frac{\partial \Xi_{TNT}}{\partial \sigma} = 0 \qquad \text{A.43}$$

A.3 Covariance vectors

The EHF requires to impose the value of the process error covariance Q, the initial state error covariance Q_0 and the measurement error covariance R, here are the values which gave us the best filter performance.

$$Q = \begin{bmatrix} 10^{-20} & 0 & 0 & 0 & 0 & 0 & 0 & 0 \\ 0 & 10^{-4} & 0 & 0 & 0 & 0 & 0 & 0 \\ 0 & 0 & 10^{-10} & 0 & 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 10^{-10} & 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 & 10^{-1} & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 & 0 & 10^{-2} & 0 & 0 \\ 0 & 0 & 0 & 0 & 0 & 0 & 10^{-3} & 0 \\ 0 & 0 & 0 & 0 & 0 & 0 & 0 & 10^{-2} \end{bmatrix} \qquad \text{A.44}$$

$$Q_0 = [10^{-10} \quad 10^{-10} \quad 10^{-10} \quad 10^{-10} \quad 10^{-10} \quad 10^{-10} \quad 10^{-2} \quad 10^{-2}] \qquad \text{A.45}$$

$$R = 10^{-4} \qquad \text{A.46}$$

These values were selected by evaluating for each state variable how confident we were in its estimated value and also which variables to modify first if the filter detected an error.

Index of symbols

SMA	Shape Memory Alloy
PE	Pseudo-Elasticity
SME	Shape Memory Effect or One-Way Shape Memory Effect
TWSME	Two-way Shape Memory Effect
DSC	Differential Calorimetry Spectroscopy
ξ	Martensitic fraction
ξ_T	Twinned martensitic fraction
ξ_S	Detwinned martensitic fraction
σ	Stress applied to the wire
ε	Wire strain
T_1	Wire temperature
T_2	Silicone temperature
e	Strain error
h	Convection coefficient
A_s	Austenite transformation temperature
A_f	Austenite transformation temperature
M_s	Martensite transformation temperature
M_f	Martensite transformation temperature
C_A	Austenite Clausius-Clapeyron coefficient
C_M	Martensite Clausius-Clapeyron coefficient
σ_F^{cr}	Upper critical stress
σ_S^{cr}	Lower critical stress
ε_L	Maximum residual strain
D	Wire elastic modulus
D_A	Austenite elastic modulus
D_M	Martensite elastic modulus
Ω	Stress/transformation coefficient
Θ	Thermoelasticity modulus
χ	TWSME retransformation fraction
d_{NiTi}	NiTi density
d_{Si}	Silicone density
m_1	Wire mass
m_2	Silicone mass
a_1	Wire radius
a_2	Silicone radius
A_0	Wire initial area
c_1	SMA specific heat capacity
c_2	Silicone specific heat capacity
ϱ_A	Austenite electrical resistivity
ϱ_M	Martensite electrical resistivity
r_{Si}	Silicone rubber thermal resistivity
L_h	NiTi latent heat
ν_{NiTi}	NiTi Poisson ratio
L_0	Initial wire length
k_{Si}	Silicone rubber thermal conductivity

Index of tables

Index of figures

Acknowledgements

I want to thank Fabio and Simone for their great help, patience and for the opportunity to work at CNR ICMATE.

I want to thank prof. Prandini for her cordiality and her availability despite her many other commitments.

I want to thank my parents, my girlfriend, my family and my friends for all the support and love throughout this period.