A comparative study of predictive models for Nafion-117 IPMC soft actuators

Burawudi, Kenny K. ; D'Anniballe, Riccardo; Langius, Ruben G.; Carloni, Raffaella

Published in:
IEEE/ASME International Conference on Advanced Intelligent Mechatronics

DOI:
10.1109/AIM46487.2021.9517466

Publication date:
2021

2021 IEEE/ASME International Conference on
Advanced Intelligent Mechatronics (AIM)

A Comparative Study of Predictive Models for Nafion-117 IPMC Soft Actuators

Kenny K. Burawudi, Riccardo D'Anniballe, Ruben G. Langius, and Raffaella Carloni

2021 IEEE/ASME International Conference on Advanced Intelligent Mechatronics (AIM) | 978-1-6654-4139-1/21/$31.00 ©2021 IEEE | DOI: 10.1109/AIM46487.2021.9517466

***Abstract*—Ionic polymer-metal composites are electro-active soft actuators that, when stimulated by an electric field, convert electrical energy into mechanical energy. This study focuses on an ionic polymer-metal composite soft actuator that has been realized with Nafion-117 and platinum electrodes. Three black-box models, i.e., curve fitting, multi-layer perceptron, and long short-term memory recurrent neural network, are designed based on the forces exerted by the soft actuator at fixed displacements when stimulated by different voltages. The capability of the three black-box models to predict unseen forces is evaluated and compared. This study shows that the multi-layer perceptron has the best predictive capability in capturing the dynamics of unseen force data, with a root mean square error of 0.109 mN and computation time of 13 μs.**

I. Introduction

Ionic polymer metal composites (IPMCs) are a class of electro-active soft actuators, for which the conversion from electrical to mechanical energy is based on mass transfer, i.e., an applied electric field causes the migration of the cations carrying a solvent in the polymer. IPMC are promising soft actuators for bio-medical and bio-robotic applications that require low actuation voltages, slow actuation, large displacements, and the need to work in wet conditions [1].

IPMC soft actuators are complex multi-domain (i.e., mechanical, electrical, thermal, and chemical) dynamical systems. As a consequence, the development of competent models, which can capture the highly nonlinear behaviour they exhibit, results in a challenging task [2]. Current research focuses on modeling the bending response of IPMC soft actuators when stimulated by different voltages. Specifically, the tip displacement has been described either by *white-box models*, i.e., a set of mathematical relations based on physical principles [3], [4], [5], or by *black-box models*, i.e., identification methods that fit experimental input-output data by means of, e.g., nonlinear autoregressions [6], [7] or neural networks [8], [9]. To the author's knowledge, research effort has not been put yet in designing or black-box models for the prediction of the dynamic behaviour of the forces exerted by the tip of IPMC soft actuators, when stimulated by different voltages, even if it constitutes a fundamental performance measure in soft robotic applications [10].

This study focuses on designing and comparing three different black-box models for an IPMC soft actuator realized with Nafion-117 and platinum electrodes. The models should be able to predict the dynamics of the tip forces exerted by the IPMC soft actuator at fixed displacements, when stimulated by different voltages. Specifically, as observed during the experiments, the models should capture the following behaviour: when a voltage is applied, the soft actuator is in its region of actuation, in which the force increases until a peak actuation force. Afterwards, the soft actuator is in its region of dehydration, in which the force decreases due to electrolysis. In the Nafion-117 IPMC soft actuator under investigation, the region of actuation lasts $\sim$20 s, while the region of dehydration $\sim$60 s. Three black-box models are designed and compared in this study, i.e., a curve fitting model (CFM) optimized by the Levenberg-Marquardt Algorithm (LMA) [11], a feed-forward Multi-Layer Perceptron (MLP), and a Long Short-Term Memory (LSTM) recurrent neural network. The models are build on experimental force data that have been collected from one IPMC soft actuator, when placed at different fixed displacements and stimulated by different voltages. The prediction capability of the models is evaluated on unseen force data (at different displacements and voltages) and are compared by using the root-mean square error, sanity check, and computation time for the prediction. The study concludes that the MLP has the best predictive capability with a final root mean square error of 0.109 mN (i.e., a final scaled root mean square error of 0.0570), and computation time of 13 μs, which guarantees its usability in real-time control.

To summarize, the three main contributions of this study are: (i) To design three black-box models for the prediction of the tip forces (which are varying in time) exerted by an IPMC soft actuator. (ii) To compare three different black-box models. (iii) To use artificial neural networks for the prediction of the tip force of an IPMC soft actuator.

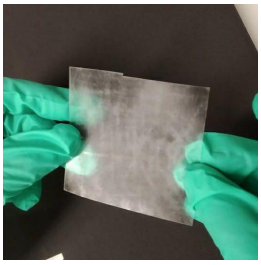
The remainder of the paper is organized as follows. Section II describes the fabrication of the Nafion-117 IPMC soft actuator, measurement set-up, and data collection. Section III presents the data pre-processing and three black-box models. Section IV reports the results, which are discussed in Section V. Finally, concluding remarks are drawn in Section VI.

II. Materials

This Section describes the process for the fabrication of the Nafion-117 IPMC soft actuator, the measurement set-up, and the data collection, which consists in measuring the forces exerted by the tip of the soft actuator at fixed displacement when stimulated by different voltages.

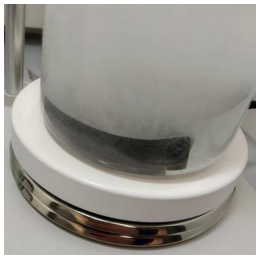
This work was funded by the European Commission's Horizon 2020 Programme as part of the project MAGNIFY under grant no. 801378.

The authors are with the Faculty of Science and Engineering, Bernoulli Institute for Mathematics, Computer Science and Artificial Intelligence, University of Groningen, The Netherlands. {k.k.burawudi,r.g.langius}@student.rug.nl, {r.danniballe,r.carloni}@rug.nl

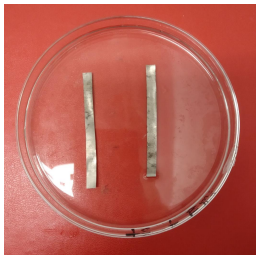


(a) A roughened and washed Nafion-117 sheet.

(b) Nafion-117 submerged in a salt solution.



(c) Metal ion reduction to form the primary plating.



(d) Secondary plating and the two identical samples of the Nafion-117 IPMC soft actuator.

Fig. 1: Fabrication steps of the Nafion-117 IPMC soft actuator.

A. Fabrication of the Nafion-117 IPMC Soft Actuator

The IPMC soft actuator is realized with a composite sandwich structure made of Nafion-117 and platinum electrodes. The actuator is fabricated by a number of chemical processes following the methodology presented in [12]. As summarized in Figure 1, the fabrication steps are: (a) *Surface Treatment:* The Nafion-117 is roughened with sandpaper. The roughening increases the surface area of the polymer that is in contact with the platinum electrodes. Then, the Nafion-117 is washed by immersing it in a boiling acidic solution to ensure that the polymer is completely saturated with protons. (b) *Ion Exchange:* The Nafion-117 is submerged in a salt solution containing platinum ions so that they are absorbed into the polymer. (c) *Primary Plating:* The absorbed platinum ions are reduced to the platinum metallic state, which form the primary electrode. (d) *Secondary Plating:* The primary plating is developed by additional coating of platinum to form the secondary plating, which reduces the electrode's resistance by increasing the mass transfer capability. The produced composite sandwich structure is cut to obtain samples of the Nafion-117 IPMC soft actuator, which have length of 65 mm and width of 5 mm.

B. Measurement Set-up

The measurement set-up consists of the Instron^{TM} ElectroPuls E1000 (www.instron.us, USA) test instrument, equipped with an optical encoder for the linear actuator, and with the Instron^{TM} 2530-5N static load cell (capacity of ± 5 N, sensitivity of 1.6 mV/V to 2.4 mV/V at static rating).

The Nafion-117 IPMC soft actuator sample is placed inside the Instron testing instrument by means of a clamp, internally covered with copper tape. The clamp is rigidly attached to the linear actuator of the Instron test instrument with a 3D-printed custom mount, and it holds the sample horizontally on the load cell [13]. A Basetech BT-305 0-30V DC power supply applies the voltage to the soft actuator through the copper tape connectors, in contact with the platinum electrodes. The forces, exerted by the tip of the soft actuator on the load cell, are recorded by the Instron^{TM} Wavematrix software at a sampling frequency of 10 Hz. The measurement set-up is shown in Figure 2.

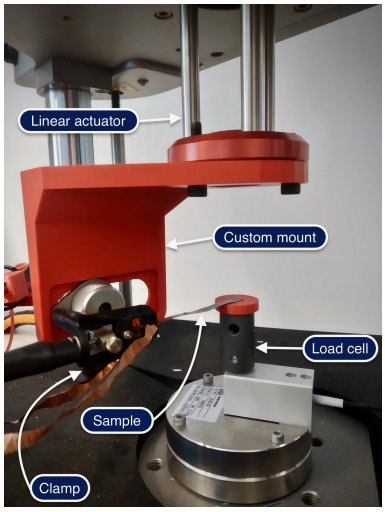


Fig. 2: The soft actuator sample is placed horizontally inside the testing instrument on the load cell by means of a clamp (rigidly attached to the instrument linear actuator with a 3D-printed mount). The DC power supply applies the voltage to the sample through copper tape inside the clamp.

C. Preparation of the Nafion-117 IPMC Soft Actuator

Before starting each round of measurements of the tip force exerted at a fixed displacement for different voltages, the soft actuator is prepared as follows: (i) It is immersed in water for 60 s to fully hydrate it, with the same hydration level for each measurement. The hydration allows the ion migration in the thickness direction of the actuator and, therefore, its bending when a voltage is applied. (ii) It is blotted in filter paper to eliminate the excess of water. (iii) It is placed in the test instrument. (iv) The load cell is calibrated to read 0 N with the weight of the hydrated sample.

D. Data Collection

The linear actuator of the test instrument is controlled to incrementally change its position, which causes also the base of the soft actuator to move. For each fixed displacement of the base of the soft actuator, the force (exerted by its tip) is measured by the load cell when different voltages are applied. The displacement varies from 0 to 20 mm, with increments of 2 mm (so for a total of 11 displacements), while the applied voltages are 2, 3, and 4 V, making it 33 force measurements (one for each voltage-displacement pair). As each measurement is performed three times, $33 \times 3 = 99$ force measurements have been collected in total. Each measurement

is limited to 100 s, i.e., a sufficient amount of time to observe the soft actuator's force response in time. Specifically, when the voltage is applied, the actuator is in its region of actuation ($0-\sim$20 s), in which the force increases until a peak actuation force. Afterwards, the actuator is in its region of dehydration ($\sim$20$-\sim$80 s), in which the force decreases due to electrolysis.

The force data collected from one IPMC soft actuator sample are used to train, validate, and test the thee black-box models developed in this study. Specifically, the data-set of the total force measurements has been divided as follows: 80% for training and 20% for testing. Within the training set, 10% has been used for validation. The testing data-set is unknown to the three black-box models during training and validation, as it is used to evaluate their prediction capability.

III. Methods

This Section presents the data pre-processing and the three black-box models, i.e., the curve fitting model, multi-layer perceptron, and long short-term memory recurrent neural network, for the Nafion-117 IPMC soft actuator.

A. Data Pre-processing

The collected force data have been pre-processed before presenting them to the black-box models. The pre-processing consists of: (i) Removing the calibration offset due to the load cell so that each measurement begins from 0 N. (ii) Extracting data from the first 80 s of the measurements, since the last 20 s of the measurements do not contain new information to be learned by the models. Since data are collected at 10 Hz, there are 800 data points for each measurement and, therefore, there are $800 \times 33 \times 3 = 79.200$ data points collected in the three rounds of measurements for the different displacement and voltages combinations. It should be noted that, since the load cell is calibrated to read 0 N with the weight of the hydrated sample, the forces becomes negative as result of the sample losing mass due to dehydration. (iii) Adding 0 N forces at the beginning of the measurement since the load cell is not able to record the initial low forces. (iv) Running a rolling median sampling window over the data to smooth the force measurements. To account for the increasing noise due to dehydration, the sampling window increases in size with respect to time as $\lfloor 2.5t \rfloor + 5$, where t is the time (logged at 10 Hz), and $\lfloor \cdot \rfloor$ is the floor function. Figure 3 shows an example of pre-processed data overlaid on the measured data.

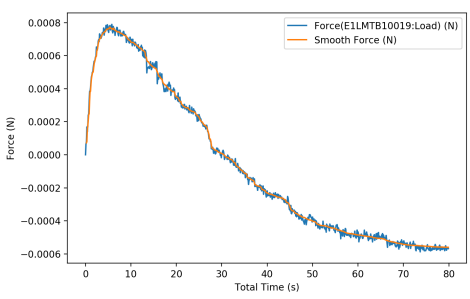


Fig. 3: Example of pre-processed force data overlaid on the measured data.

B. Training Data-set

The original training data-set consists of the 80% of the complete measurement data-set, i.e., 63.360 data points. In this study, two different training data-set have been considered for the training of the three black-box models. Specifically, the two training data-sets are: the *unaugmented training data-set* that contains only the pre-processed data, as described in the previous subsection, i.e., 63.360 data points, and the *augmented training data-set* in which each one of the 63.360 data points has been augmented by a factor 4 for a total of $63.360 \times 4 = 253.440$ data points.

C. Curve Fitting Model

The CFM has been chosen because of the characteristic shape of the data observed in the experiments, in which the soft actuator achieves a peak force and, afterwards, it loses the ability to exert a force due to dehydration.

The CFM assumes that the forces observed at time t are described by $F(t) = g(\vec{x}, t) + \epsilon$, where ϵ is an error (assumed to be drawn from a normal distribution), g is a predefined function whose parameters $\vec{x}$ need to be fitted. Specifically, the forces of the soft actuator can be described by a lognormal probability density function in time for a given voltage-displacement combination, i.e., the function g is thus defined as $g(\vec{x}, t) = \text{lognorm}(t, a) + b$, where a is a distribution scaling parameter, b is a translation parameter, and $\text{lognorm}(t, a) = a^2/(s\ t\sqrt{2\pi}) exp\left(-ln^2\left(\frac{t}{a}\right)/(2s^2)\right)$ is the standardized probability density function, where s is the shape parameter for the lognormal function. The optimal parameters have been found for individual voltage-displacement combinations on the training data-set.

The CFM has to fit three parameters, i.e., a that scales the range of the lognormal function to the range of the exerted force, b that accounts for some of the experimental errors that could not be removed in the pre-processing stage, and s that determines the shape of the lognormal function. To fit the parameters, the LMA is used as damped least-squares optimizer. This optimization process is iterative and the parameters $\vec{x} = [a \quad b \quad s]^T$ are adjusted at each step until the LMA finds a local minimum of the gradient function. In this study, the model parameters are initialized to 1 and the iterations are limited to a maximum of 1000.

D. Multi-Layer Perceptron

The second black-box model is the feed-forward MLP, which has been chosen because it is a simple and powerful method for function approximation [14], [15].

In this study, the MLP has three inputs, i.e., a fixed displacement of the soft actuator, a fixed applied voltage, and the time, and one output, i.e., the exerted force in time. The chosen architecture has three hidden layers, with 250, 10, 250 neurons, respectively, with a 20% dropout [16]. The hidden and output neurons have a sigmoid activation function. The MLP is trained for 20 epochs, the loss function is the mean square error, and the optimizer is Adam [17]. The inputs and output data are scaled to [0,1] by the min-max scaler.

E. Long Short-Term Memory

The third black-box model is a recurrent neural network, and, specifically, a LSTM network [18]. It has been chosen for its ability to model long-term dependencies in data by exploiting a memory cell at each neuron, and to represent differential equations.

In this study, the LSTM network has three inputs, i.e., a fixed displacement of the soft actuator, a fixed applied voltage, and the time, and one output, i.e., the exerted force in time. The chosen architecture has three hidden layers, with 100, 10, 100 LSTM units, respectively, with a 20% dropout. The LSTM network is trained for 7 epochs, the loss function is the mean square error, and the optimizer is Adam. The inputs and output are scaled to [0,1] by the min-max scaler.

F. Models' Evaluation

The root mean square error (RMSE) is one of the performance metrics used to evaluate and compare the models as it represents the average prediction error.

1) RMSE in Training, Validation, and Testing: The predictive capability of each model to predict unseen data is evaluated by comparing the RMSE on the testing data-set with the RMSE on the training and validation data-sets. If similar, it means that the models can predict unseen data.

2) RMSE in Testing: The comparison between the models is made by comparing their RMSEs on the testing data-set.

3) RSME in the Region of Actuation and in the Region of Dehydration: From the data collection, it was noted that the soft actuator has a region of actuation in the first $\sim$20 s, after which a region of dehydration starts. The region of actuation is very important and the models must capture this dynamic response. Therefore, an additional condition to the models' evaluation is a second RMSE on the testing data-set for the prediction made in the first 20 s of each measurement. A lower RMSE in the region of actuation than for the whole measurement (i.e., 80 s after pre-processing) indicates that the model learns to predict the sample's dynamics in the region of actuation. A higher RMSE than for the whole measurement indicates that the model does not learn the sample's dynamics in the region of actuation and, instead, it minimizes the error when the sample is in the region of dehydration (mostly static behaviour).

4) Sanity Check: The RMSE can be misleading because the models can learn a noisy relation, which brings to a low RMSE. A sanity check is, therefore, necessary which consists in overlaying the models' predictions on the measurement data. Only if the predictions show both the region of actuation and of dehydration, the RMSE can be used as performance metric.

5) Computation Time: The computation time for each model is also a metric to evaluate models. A lower prediction time indicate an overall lower model complexity.

IV. Results

This Section reports the training and validation results for the three black-box model and the comparison between them.

A. Curve Fitting Model

The final RMSE of the CFM are reported in Table I. In the case of augmented training data-set, the lower RMSE for the testing data-set compared to the training and validation data-sets indicates that the CFM can predict the unseen testing data. On the contrary, in the case of unaugmented training data-set, the RMSE of the testing data-set is higher and, therefore, the CFM cannot predict the unseen testing data.

TABLE I: Final RMSE for the CFM (in mN).

	augmented data-set	unaugmented data-set
training	0.1208	0.1434
validation	0.1327	0.1334
testing	0.115	0.1537

B. Multi-Layer Perceptron

Figure 4 shows the RMSE losses of the MLP with respect to the epochs for both the augmented and the unaugmented training data-sets. It can be noted that the MLP learns the relationship between the inputs (i.e., displacement, voltage, time) and output (i.e., the forces), and that it performs better in the case of the augmented data-set.

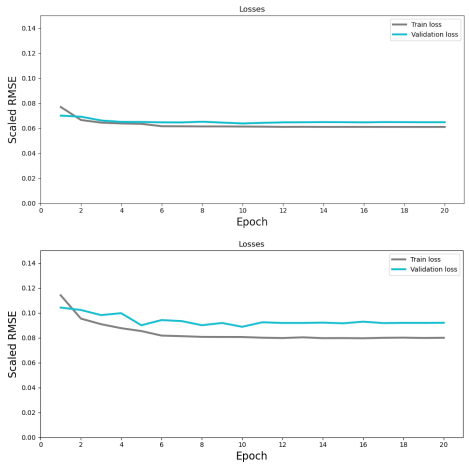


Fig. 4: The scaled RMSE loss of the MLP with respect to the epochs for the augmented training data-set (top) and for the unaugmented training data-set (bottom).

The final scaled RMSE (and the final RMSE) of the MLP are reported in Table II. The lower RMSE for the testing data-set compared to the training and validation data-set indicates that the MLP model can predict the unseen testing data.

TABLE II: Final scaled RMSE (and the final RMSE) for the MLP.

	augmented data-set	unaugmented data-set
training	0.0610 (i.e., 0.1175 mN)	0.0800 (i.e., 0.128 mN)
validation	0.0648 (i.e., 0.1245 mN)	0.0920 (i.e., 0.1484 mN)
testing	**0.0570 (i.e., 0.1095 mN)**	0.1013 (i.e., 0.1633 mN)

C. Long Short-Term Memory

Figure 5 shows the RMSE losses of the LSTM with respect to the epochs for both the augmented and the unaugmented

training data-set. It can be noted that the LSTM learns the relationship between the inputs (i.e., displacement, voltage, time) and the output (i.e., forces), and that it performs better in the case of the augmented data-set.

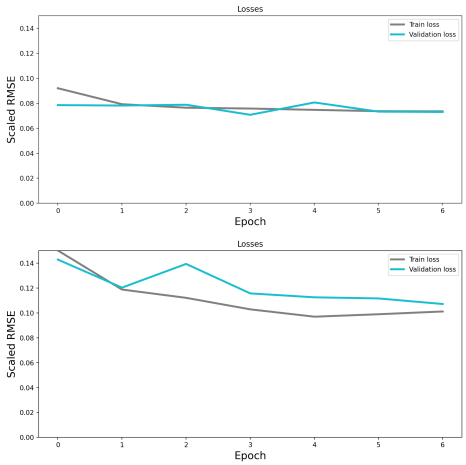


Fig. 5: The scaled RMSE loss of the LSTM with respect to the epochs for the augmented training data-set (top) and for the unaugmented training data-set (bottom).

The final scaled RMSE (and the final RMSE) of the LSTM are reported in Table III. The lower RMSE for the testing data-set compared to the training and validation data-sets indicates that the LSTM can predict the unseen testing data.

TABLE III: Final scaled RMSE (and the final RMSE) for the LSTM.

	augmented data-set	unaugmented data-set
training	0.0669 (i.e., 0.1286 mN)	0.0924 (i.e., 0.1490 mN)
validation	0.0730 (i.e., 0.1402 mN)	0.1071 (i.e., 0.1727 mN)
testing	0.0595 (i.e., 0.1143 mN)	0.0937 (i.e., 0.1511 mN)

D. Comparison

For a complete evaluation and comparison of the prediction capabilities of the three models on the testing data-set, this study calculates the RMSE in the region of actuation, performs a sanity check, and computes the prediction time.

1) RMSE in the Region of Actuation: Table IV shows the RMSE for each model in both the region or actuation (first $\sim$20 s) and in the whole duration of the measurements ($\sim$80 s) for both the augmented and unaugmented training data-set. It can be noted that, in the case of augmented training data-set, only the MLP predicts the dynamics of the forces in the region of actuation better than for the entire duration of the measurements. In the case of the unaugmented training data-set, all the three models predict the dynamics of the forces in the region of actuation better than for the entire duration of the measurements. Moreover, in the case of augmented training data-set, the MLP returns the lowest RMSE in both the region of actuation and over the entire duration of the measurements. In the case of unaugmented training data-set, the MLP returns the lowest RMSE in the region of actuation, while the LSTM returns the lowest RMSE over the entire duration of the measurements. In general, the CFM performs worse than the MLP and LSTM, except for the MLP in the prediction of the forces in the entire duration of the measurements (in the case of the unaugmented training data-set).

TABLE IV: RMSE on the testing data-set for both the augmented and unaugmented training data-set. Units are in mN.

		CFM	MLP	LSTM
<20 s	augmented data-set	0.138	**0.104**	0.122
	unaugmented data-set	0.180	**0.140**	0.145
0-80 s	augmented data-set	0.115	**0.109**	0.114
	unaugmented data-set	0.154	0.163	**0.151**

Figure 6 compares the RMSEs of the MLP and LSTM in the region of actuation for the augmented training data-set. It can be noted that the MLP is more reliable than the LSTM when presented with the unseen data.

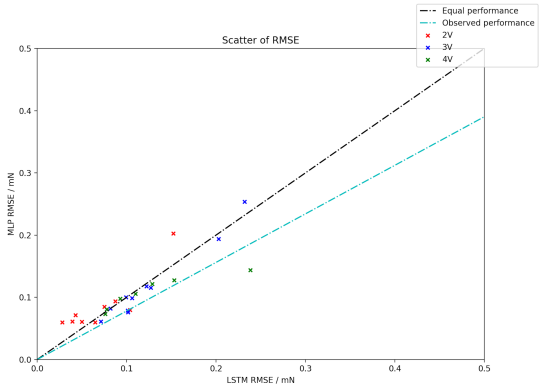


Fig. 6: The RMSEs of the MLP against the LSTM in the region of actuation, for the augmented training data-set.

2) Sanity Check: Figure 7 shows the sanity check for the MLP and, namely, the predicted dynamics of the tip forces for different displacement/voltage combinations (in red) overlaid on the testing data-set (in blue). The missing data are the ones that were not in the testing data-set and, therefore, are not reported for the sanity check.

3) Computation time: The computation time for the prediction of a single input vector is 1.3 μs for the CFM, 13 μs for the MLP, and 225 μs for the LSTM, making the CFM the fastest model. However, all models are fast enough to be used to predict the actuation forces in real-time.

V. Discussion

This Section discusses the results and the limitations of the models and suggests possible improvements.

The CFM has the lowest performance among the three models, with a higher RMSE on the testing data-set when predicting the dynamics of the forces in both the region of actuation and over the entire process, except in the case with the unaugmented data-set for the entire duration of the experiments (see Table IV) in which, nevertheless, it cannot predict the unseen testing data (see Table I). The CFM is the simplest of the three investigated black-box models (with only 99 parameters), making the model worth investigating. The

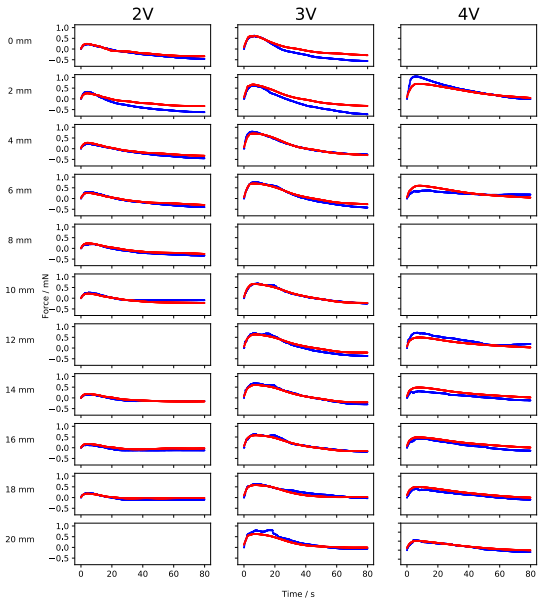


Fig. 7: MLP predicted forces in time for different displacement/voltage combinations (in red) overlaid on the testing data-set (in blue).

simplicity of the CFM also makes it the most computationally efficient in the prediction, which would be best suited for real-time predictions. The MLP is a simplified adaptation of the neural network designed in [9] for the prediction of the tip displacement. It is the most consistent model with generally good performance across all metrics of evaluation, and it returns more reliable predictions. The model's strongest point is its relative simplicity as a universal function approximator [14]. The LSTM can learn dependencies with the training data-set (similar RMSEs in the training and validation data-sets, see Table III). It is difficult, however, to reconcile the added complexity of the LSTM with a minor improvement in the prediction capability and an higher computation time than in the MLP.

A. Limitations and future Outlook

A first limitation in this study lies in the data collection. Due to dehydration, the data collection results in a difficult process, which makes the creation of a large experimental data-set challenging for black-box models. Secondly, the sample's loss of mass over the course of a measurement, due to dehydration in the experiments of this study, results in unpredictable errors between repeated measurements. Modeling the weight loss of the sample may help to reduce these errors as they could be considered in the data pre-processing. Finally, the predictive capabilities of the MLP and LSTM could be improved by including the physical knowledge on the underlying dynamics of the soft actuator.

VI. Conclusion

This study investigates three black-box models for the prediction of the dynamics of the tip forces of a Nafion-117 IPMC soft actuator at fixed displacements and when stimulated by different voltages. The models are the CFM (optimized by the LMA), the MLP, and the LSTM. The study also proposed a framework to compare the performances of these models in relation to each other, contributing to a black-box model selection paradigm for IPMC soft actuators.

The study showed that the assumption made by the CFM does not hold for unseen testing data, and that the MLP and the LSTM are better models that perform similarly with regard to their prediction capabilities. However, the MLP is best suited to model the actuation forces of a Nafion-117 IPMC soft actuator. This observation was made on balance of three factors. The first is the RMSE on the predictions of unseen testing data of 0.104 mN in the region of actuation and a RMSE of 0.109 mN (i.e., a final scaled RMSE of 0.0570) over the entire sample response (when an augmented training data-set is used). The second is the consistency of predictions across all voltages conditions. The third is the prediction time of 13 μs.

References

[1] M. Shahinpoor and K. J. Kim, "Ionic polymer-metal composites: IV. Industrial and medical applications," *Smart Materials and Structures*, vol. 14, no. 1, pp. 197–214, 2005.

[2] M. Shahinpoor and K. J. Kim", "Ionic polymer-metal composites: I. Fundamentals," *Smart Mat. and Structures*, vol. 10, pp. 819–833, 2001.

[3] E. T. Enikov and G. S. Seo, "Numerical analysis of muscle-like ionic polymer actuators," *Biotechnology Progress*, vol. 22, no. 1, pp. 96 – 105, 2008.

[4] A. Punning *et al.*, "A distributed model of ionomeric polymer metal composite," *Journal of Intelligent Material Systems and Structures*, vol. 20, no. 14, pp. 1711 – 1724, 2009.

[5] C. Jo *et al.*, "Recent advances in ionic polymer?metal composite actuators and their modeling and applications," *Progress in Polymer Science*, vol. 38, pp. 1037 – 1066, 2013.

[6] M. Annabestani and N. Naghavi, "Nonlinear identification of IPMC actuators based on ANFIS-NARX paradigm," *Sensor and Actuators A: Physical*, vol. 209, pp. 140 – 148, 2014.

[7] R. D'Anniballe, G. Paoletta, and R. Carloni, "A polyurethane-based electrospun nanofiber bundle soft actuator: Fabrication, modeling, and control," in *IEEE Int. Conference on Soft Robotics*, 2021.

[8] D. Q. Truong and K. K. Ahn, "Design and verification of a non-linear black-box model for ionic polymer metal composite actuators," *J. of Intel. Mat. Sys. and Structures*, vol. 22, no. 3, pp. 253 – 269, 2011.

[9] ——, "Modeling of an ionic polymer metal composite actuator based on an extended Kalman filter trained neural network," *Smart Mat. and Structures*, vol. 23, no. 7, pp. 140 – 148, 2014.

[10] L. Yang *et al.*, "Prediction of the actuation property of Cu ionic polymer-metal composites based on backpropagation neural networks," *American Chemical Society*, vol. 5, pp. 4067–4074, 2020.

[11] J. J. Moré, *The Levenberg-Marquardt algorithm: Implementation and theory*. Springer Berlin Heidelberg, 1978.

[12] V. De Luca *et al.*, "Ionic electroactive polymer metal composites: Fabricating, modeling, and applications of postsilicon smart devices," *J. of Polymer Science Part B*, vol. 51, pp. 699 – 734, 2013.

[13] R. Carloni *et al.*, "A variable stiffness joint with electrospun P(VDF-TrFE-CTFE) variable stiffness springs," *IEEE Robotics and Automation Letters*, vol. 3, no. 2, pp. 973–978, 2018.

[14] K. Hornik, "Approximation capabilities of multilayer feedforward networks," *Neural Networks*, vol. 4, no. 2, pp. 251 – 257, 1991.

[15] G. Cybenko, "Approximation by superpositions of a sigmoidal function," *Math. of Control, Signals and Syst.*, vol. 5, no. 4, p. 455, 1992.

[16] N. Srivastava *et al.*, "Dropout: A simple way to prevent neural networks from overfitting," *J. of Machine Learning Research*, vol. 15, no. 2, pp. 1929 – 1958, 2014.

[17] Z. Zhang, "Improved Adam optimizer for deep neural networks," in *Proceedings IEEE/ACM Int. Symposium on Quality of Service*, 2018.

[18] A. Sherstinsky, "Fundamentals of Recurrent Neural Network (RNN) and Long Short-Term Memory (LSTM) network," *Physica D: Nonlinear Phenomena*, vol. 404, 2020.