e-Journal of Nondestructive Testing - ISSN 1435-4934 - www.ndt.net

EWSHM 2024
11th European Workshop on Structural Health Monitoring

Transfer learning applications for corrosion prediction for electrodissoluted specimens

Lucio PINELLO[1], Vasiliki PANAGIOTOPOULOU[1], Marco GIGLIO[1], Lourdes VASQUEZ GOMEZ[2], Luca MATTAROZZI[2], Claudio SBARUFATTI[1]
[1] Department of Mechanical Engineering, Politecnico di Milano, Milan, 20156, Italy, e-mail:
lucio.pinello@polimi.it vasiliki.panagiotopoulou@polimi.it marco.giglio@polimi.it
claudio.sbarufatti@polimi.it
[2] Istituto di Chimica della Materia Condensata e di Tecnologie per l'Energia - ICMATE, Consiglio Nazionale delle Ricerche; Padova, Italy, e-mail:
lourdes.vazquezgomez@cnr.it luca.mattarozzi@cnr.it

Abstract. Corrosion monitoring is a critical aspect of Structural Health Monitoring (SHM), as it presents inherent challenges in tracking its progression and requires strategically placed sensors for effective surveillance. Moreover, the intricate interplay between environmental factors and the material composition of corroding components further complicates this phenomenon.

In this study, a novel neural network framework is designed to predict corrosion evolution based on environmental data, focusing on three distinct aluminium alloys: Al7475, Al2024, and F357. Dedicated neural network models are developed for Al7475 and F357, trained on electrodissolution data, including process parameters and profilometry measurements. These models predict corrosion progression in terms of volume loss, corroded depth, and area based on charge measurements. To address the challenge of material-environment coupling, transfer learning is employed to adapt the neural network to Al2024.

The oxidation charge passed during the electrodissolution mimics in an artificial way the corrosion process due to environmental conditions in the attack site. Consequently, the implemented neural network establishes a robust connection between corrosion process measurements and profilometry data, enabling accurate corrosion progression prediction based on oxidation charge. This framework allows the prediction of a component's residual useful life based on the assessment of the corrosion state, with the knowledge of volume loss further supporting mechanical assessment by evaluating resistant section or pit dimensions, ultimately aiding in the pit-crack transition analysis.

Keywords: corrosion monitoring, SHM, transfer learning, neural network

https://doi.org/10.58286/29626

Introduction

In the framework of Structural Health Monitoring (SHM), corrosion monitoring is a critical aspect due to its intrinsic probabilistic nature and the intricate interplay between environmental factors and the material under corrosion. Corrosion phenomena are usually monitored by exploiting non-destructive techniques (NDT) such as acoustic emission (AE) [1] [2], electrochemical noise measurements (ENM) [3] [4], eddy current [5], electrochemical impedance spectroscopy (EIS) [6].

However, these methods allow only the assessment of the current state of the corrosion process. The difficulties in performing prognosis are of serious concern since corrosion can seriously affect the integrity of structural components for instance by reducing their thickness and mechanical strength. In addition, if one thinks about localised corrosion, corrosion pits can lead to localised stress concentration and become initiating points for fatigue cracks [7] [8] [9]. Thus, the corrosive phenomena jeopardise the structural integrity of mechanical components [10] [11].

For this reason, machine learning (ML) techniques are being implemented in the SHM framework for predicting the corrosion rate for different kinds of corrosive phenomena [12] [13] [14]. Within the ML framework for corrosion rate prediction, growing interest is being posed in transfer learning (TL) techniques [15] [16], which allow adapting a ML model developed on a database to a new set of data.

Therefore, in this work, accelerated corrosion tests are performed on three aluminium alloys (Al7475, Al2024, and Al F357) through electrodissolution tests in a $NaCl\ 5\% + HCl\ 0.5\ M$ solution simulating a marine environment. The evolution of the corrosion process is monitored by performing profilometries to obtain corrosion penetration depth and corroded transversal area measurements. The oxidation charge passed during the electrodissolution mimics in an artificial way the corrosion process due to environmental conditions in the attack site. Consequently, neural networks are implemented to establish a robust connection between corrosion process measurements and profilometry data, enabling accurate corrosion progression prediction based on oxidation charge.

Eventually, a transfer learning technique named Multi-Fidelity Model (MFM) is used to adapt the neural networks developed for Al 7475 and Al F357 to the Al 2024 alloy. This process mimics a real-world scenario, for which, as new corrosion measurements are available for the monitored structure, predictions based on previous experiences should be made for prognosis, where, in this case study, the previous experience is the knowledge of the corrosion rate for one or more similar materials. In addition, the prediction of the corroded volume, in terms of maximum corroded depth and maximum corroded area, may be used for estimating the residual mechanical strength of the structure and to assess whether a corrosion pit is growing and becoming a crack or not.

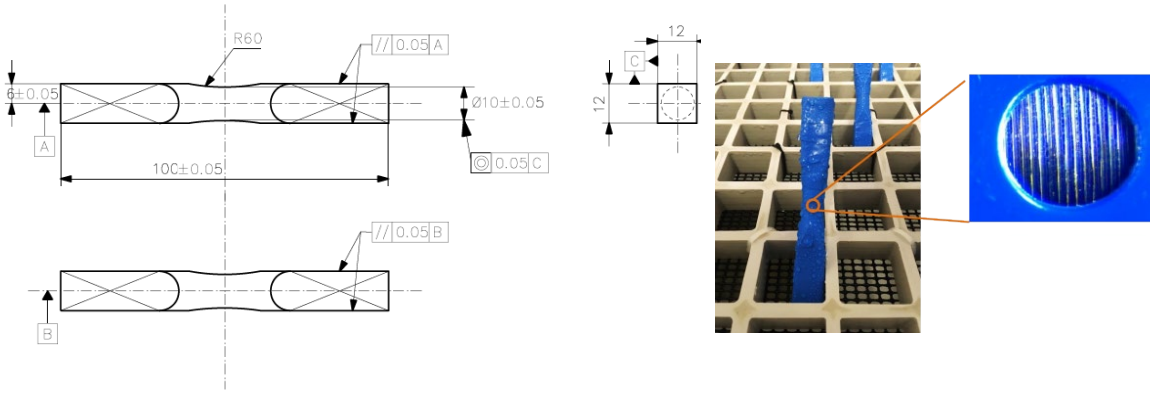
1. Materials and Methods

1.1 Materials

Three aluminium alloys have been selected to be tested as representative of several components used in commercial products: Al 7475, Al 2024, and Al F357. Twelve specimens for each material were available, whose geometry is depicted in **Fig. 1**. Before performing the tests, a vinyl tape was used to realise a mask to cover the surface of the specimens except for a circle with a diameter of $1.2\ mm$ at the centre of the specimens, as reported in **Fig. 1(b)**. The circle location is the intersection between the horizontal and vertical axis of symmetry

in **Fig. 1(a)**. The vinyl tape mask was realised to force the corrosion on the exposed circle, preventing other regions from being corroded.

Eventually, the exposed surface was etched with a $NaOH$ 4.0 M water-based solution for 1 minute to remove the existing alumina on the exposed surface.



(a) **(b)**

Fig. 1. Specimens subjected to electrodissolution tests: (a) geometry of the specimens and (b) specimen covered with the vinyl tape with the exception of a 1.2 mm circular hole in the center.

1.2 Electrodissolution

The electrodissolution test is performed in a $NaCl$ 5% + HCl 0.5 M naturally aerated water-based solution, to mimic the exposure of the specimens to a marine environment. A potentiostatic test is carried out until a charge of 3.0 C is obtained. The Metrohm Autolab M204 potentiostat/galvanostat has been used together with a standard three electrode cell, with two Pt counter electrodes and a saturated calomel (SCE) reference electrode.

Before starting the tests, the open circuit potential (OCP) was measured after 5 minutes of exposure to reach a stable potential and observe the difference in free-corrosion potential for the three alloys in the selected solution. Then, a potentiostatic reduction was carried out for 5 minutes with a voltage V_{red}^{*} to remove possible residual alumina, that may not be removed by the etching procedure, and promote the activation of the entire exposed surface. This procedure aims at improving the reproducibility of the geometry of the corrosion defect, removing the dependencies from the specimen surface inhomogeneity Eventually, the potentiostatic oxidation was performed by holding constant the voltage to a value of V_{ox}^{*} until the measured charge reached values of 1.0 C, 2.0 C, and 3.0 C. At each of those charge values, the tests were stopped to allow profilometry measurements to track corrosion evolution. The values of V_{red}^{*} and V_{ox}^{*} depend on the alloy tested, as reported in **Table 1**. At each stop, the specimens were washed with water and air-dried to allow profilometry measurements.

The reason for the different testing voltages is the different nobility of the alloy used in this experimental campaign.

Table 1. Applied potential V_{red}^{*} and V_{ox}^{*} during the potentiostatic test for each alloy.

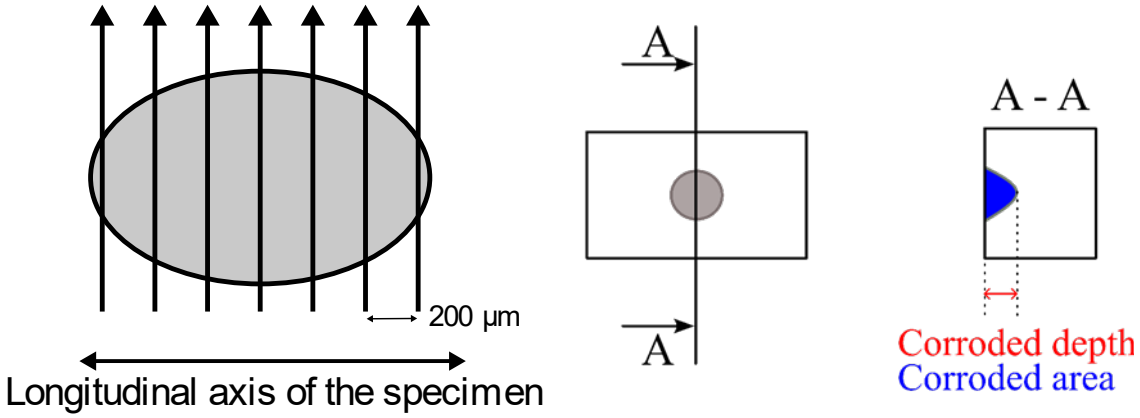
Material	**Potentiostatic Reduction Voltage V_{red}^{*} [V]**	**Potentiostatic Oxidation Voltage V_{ox}^{*} [V]**
Al 7475	$-1.55\ V\ (SCE)$	$-0.55\ V\ (SCE)$
Al 2024	$-1.2\ V\ (SCE)$	$-0.2\ V\ (SCE)$
Al F357	$-1.5\ V\ (SCE)$	$-0.2\ V\ (SCE)$

Being the charge defined as the integral of the current over time, measuring the charge during a corrosion process gives an estimation of the evolution of the corrosion since large charge variations, over a given time period, can be associated with a faster corrosion.

1.3 Profilometry

Profilometries were performed at each stop, e.g., when the charge reached the values of $1.0\ C$, $2.0\ C$, and $3.0\ C$, to measure the corroded profile. The stylus profilometer Bruker Dektak XT has been used and the Sensofar software SensoMap 9 was selected to process the profilometries.

Seven profiles are tracked at each stop with an offset of $200\ \mu m$ between each profile, as depicted in **Fig. 2(a)**, with the 4th profile taken in the centre of the exposed area. The tracking of the profiles allows the estimation, for each profile, of the corroded depth, taken as the lowest point among all the profiles, and of the corroded area, taken transversally as depicted in **Fig. 2(b)**. Eventually, the maximum depth and the maximum area are taken as the parameters describing the corrosion process evolution.



(a) **(b)**

Fig. 2. Description of the profilometry measurements: (a) Direction and offset of the profilometries performed on the electrodissoluted specimens and (b) Schematic of the data extracted from each profilometry: maximum corroded depth and maximum corroded area, evaluated at the cross-section of the specimen.

1.4 Methodology

The acquired database consists of 12 maximum corroded depth and area measurements for each charge value ($1.0\ C$, $2.0\ C$, and $3.0\ C$) and each alloy (Al 7475, Al 2024, and Al F357), as shown in **Fig. 3** (a) and (b), respectively, for corroded depth and area. As it is possible to notice, the corroded area data in **Fig. 3** (b) are almost overlapped while

this is not true for the corroded depth data. More specifically, the Al 2024 shows a higher penetration depth of the corrosion process, as it stands above the other two alloys in **Fig. 3** (a).

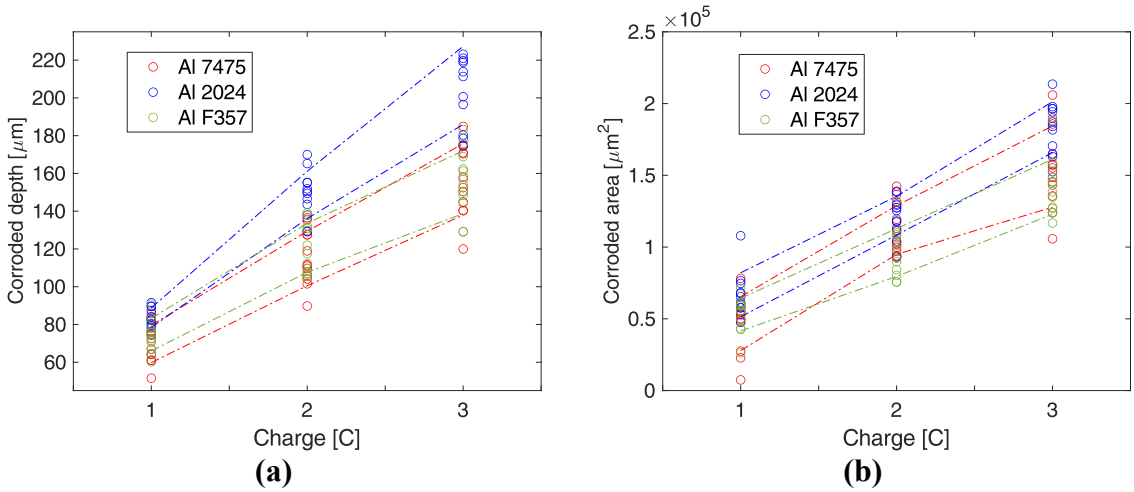


Fig. 3. Databases of (a) maximum corroded depth and (b) maximum corroded area versus the charge for the three alloys. The dashed lines represent the mean value plus and minus the standard deviation.

Neural networks were implemented to correlate the charge values with the acquired measurements, in particular, the NNs predictions can either be (i) the corroded depth or (ii) the corroded transversal area. Thus, two NNs are developed for each alloy. This work proposes a transfer learning framework based on the database of **Fig. 3** by assuming that the data of Al 2024 are available only for 1.0 C and 2.0 C. Instead, data for Al 7475 and Al F357 are available for all the charge values. Therefore, there is the need to estimate the evolution of the corrosion process for Al 2024 at 3.0 C.

For this scope, the Multi-Fidelity Method (MFM) has been implemented since it adapts well to the available database. This approach is based on the availability of two datasets: (i) a large dataset characterised by low-fidelity data, for instance, data that deviates from the real system of interest, and (ii) a small dataset composed of high-fidelity data, which is accurate and reliable data. To build data-based models it is aimed to have a large dataset, but using the low-fidelity data implies the realization of a model that is not very accurate, and, on the other hand, the high-fidelity data may be too few to build a model. Thus, to tackle the challenge of insufficient high-fidelity data, it is possible to use the information from low-fidelity data to improve prediction accuracy. The idea is to build (i) a low-fidelity model based on the low-fidelity data, and (ii) a multi-fidelity model based on the high-fidelity data and the output of the low-fidelity model.

Referring to **Fig. 4**, the low-fidelity model receives as input the low-fidelity data and gives low-fidelity outputs, and the MFM model receives as input the high-fidelity data and the low-fidelity output, giving high-fidelity outputs. Thus, this approach leverages the availability of a huge low-fidelity dataset, with which it is possible to have accurate low-fidelity predictions, with the availability of a small high-fidelity dataset, with which it is not possible to build a reliable model due to the few data available. By combining the two datasets in the method explained before, it is possible to build a Multi-Fidelity model that is based on both datasets.

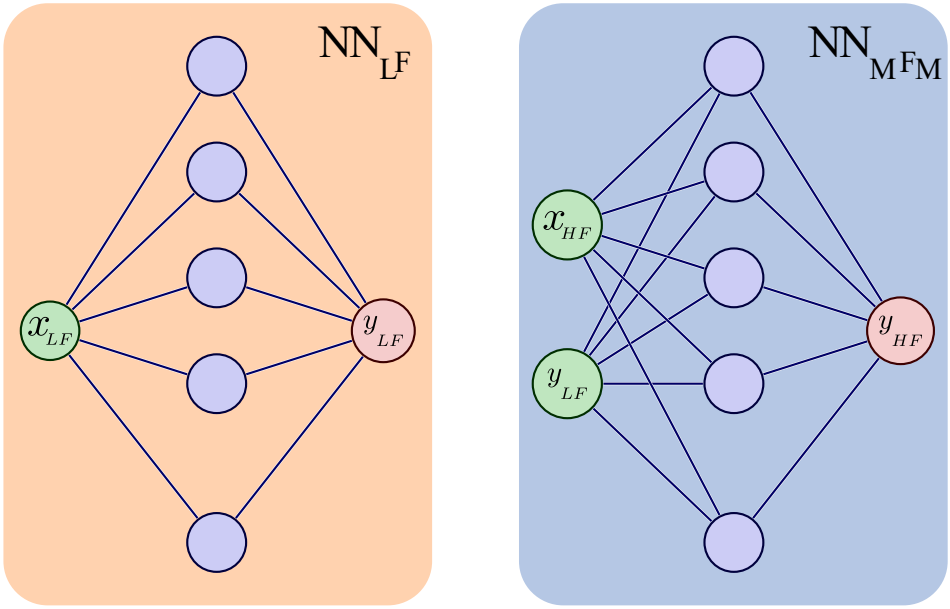


Fig. 4. Schematic of the Multi-Fidelity Method approach.

For the purpose of the work, the low-fidelity data can be seen as corrosion data of the Al 7475 and Al F357 alloys, while the high-fidelity data can be data of the material of Al 2024. The low-fidelity dataset is termed as source domain, while the high-fidelity dataset is called target domain. In this way, it is possible to transfer the knowledge acquired from one material to another one. For instance, data coming from Al 7475 and Al F357 are inconsistent for the corrosion depth with respect to Al 2024 measured depths, therefore they can be considered as low fidelity data, e.g., source domain.

As resumed in **Table 2**, three scenarios are going to be analysed: (1) the source domain consists of Al 7475 data and Al 2024 is the target domain, (2) F357 is the source domain and Al 2024 is the target domain, and (3) the source domain consists of both Al 7475 and F357 data. In particular, given the database shown in **Fig. 3**, the focus is going to be set on the corroded depth since the data regarding the transversal corroded area are close to each other.

Table 2. Scenarios analysed in this study.

Scenario	Source domain	Target domain
1	Al 7475	Al 2024
2	Al F357	
3	Al 7475, Al F357	

The low-fidelity and high-fidelity NNs implemented consist of a single hidden layer of 60 neurons with one input (charge) and one output (depth). The multi-fidelity NN differs from the previous two only for the number of inputs, that is two. The optimisation algorithm is the adam algorithm the maximum number of iterations was set to be 1000. The NNs are implemented within the sklearn.neural_network.MLPRegressor Python function.
Before training the NNs, the available data (source domain and target domain) are normalised to lay in the [0;1] domain where it is worth remembering that the target domain does not include Al 2024 data at 3.0 C. Therefore, when the latter data will be normalised, they will stay outside of the [0;1] domain. The three models are compared based on the MAPE (Mean Absolute Percentage Error).

2. Results and Discussion

This section illustrates the results of the methodology in Section 1.4 Methodology applied to the three scenarios shown in **Table 2**.

Starting with the first scenario, in which Al 7475 data are the source domain and Al 2024 data at 1.0 C and 2. 0 C are the target domain, **Fig. 5** shows the predictions of the three NNs implemented. As it is possible to notice, the Multi-Fidelity Model output (in brown) is closer to the mean value of Al 2024 data at 3.0 C (named as test samples and represented in orange in the figure) with respect to the Low-Fidelity and High-Fidelity models' predictions, respectively in red and orange. This is confirmed also by comparing the MAPE for the three models reported in **Table 3**.

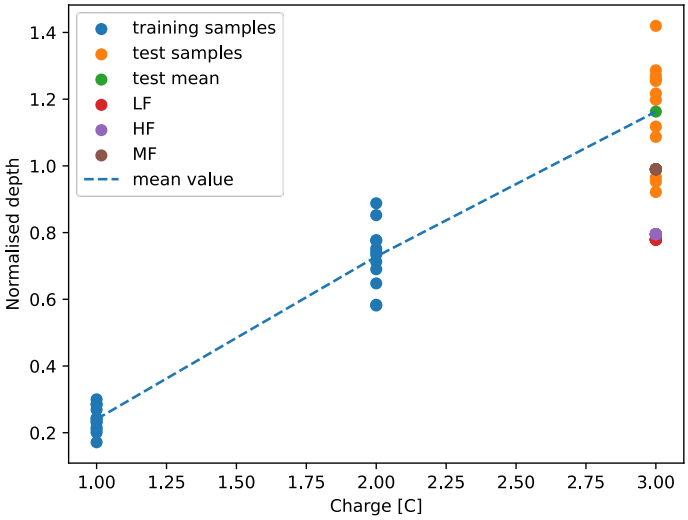


Fig. 5. Comparison of the predictions of the Low-Fidelity Model (in red), High-Fidelity Model (in purple), and Multi-Fidelity Model (in brown) outputs for the 1st scenario (source domain: Al 7475; target domain: Al 2024).

Table 3. MAPE comparison of Low-Fidelity, High-Fidelity, and Multi-Fidelity models for the 1st scenario (source domain: Al 7475; target domain: Al 2024).

Low-Fidelity Model	High-Fidelity Model	Multi-Fidelity Model
32.99%	31.59%	14.86%

Moving on to the 2nd scenario, that is when the source domain consists of Al F357 data and the target domain is composed of Al 2024 data at 1.0 C and 2.0 C, **Fig. 6** and **Table 4** show the comparison of, respectively, the predictions of the models and the MAPE computed on those predictions. As before, the Multi-Fidelity model is the one that shows the best performance among the three methods being closer to the mean value of the test sample (Al 2024 at 3.0 C).

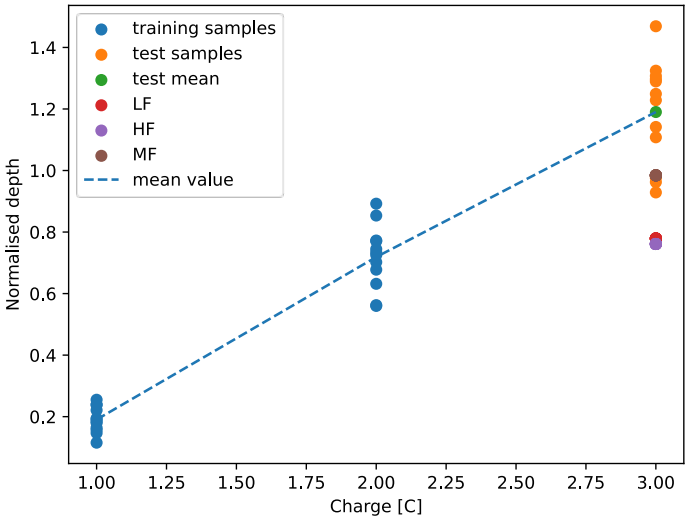


Fig. 6. Comparison of the predictions of the Low-Fidelity Model (in red), High-Fidelity Model (in purple), and Multi-Fidelity Model (in brown) outputs for the 2nd scenario (source domain: Al F357; target domain: Al 2024).

Table 4. MAPE comparison of Low-Fidelity, High-Fidelity, and Multi-Fidelity models for the 2nd scenario (source domain: Al F357; target domain: Al 2024).

Low-Fidelity Model	High-Fidelity Model	Multi-Fidelity Model
34.53%	36.04%	17.29%

Finally, the third scenario addresses the case in which the source domain consists of Al 7475 and Al F357 data, while the target domain is made of Al 2024 data at 1.0 C and 2.0 C. By looking at **Fig. 7**, it is possible to observe that, as in the previous cases, the Multi-Fidelity Model shows superior performances with respect to the other two models, being closer to the mean value of Al 2024 data at 3.0 C. **Table 5** confirms the previous observation since the Multi-Fidelity Model is the one with the lowest MAPE compared to the Low and High-Fidelity models.

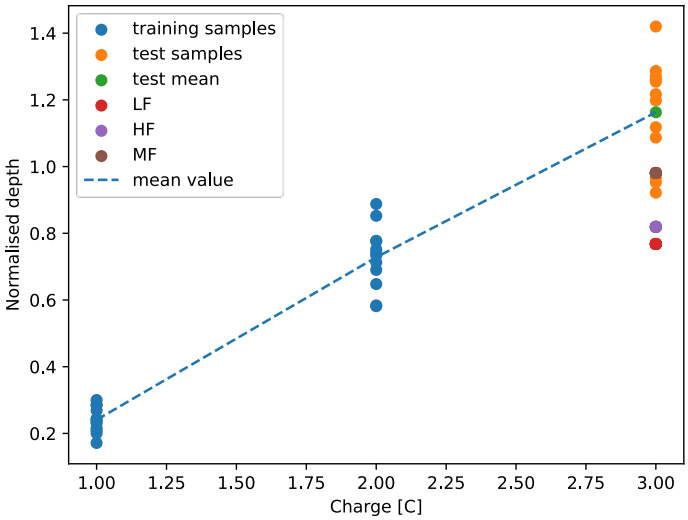


Fig. 7. Comparison of the predictions of the Low-Fidelity Model (in red), High-Fidelity Model (in purple), and Multi-Fidelity Model (in brown) outputs for the 3rd scenario (source domain: Al F357; target domain: Al 2024).

Table 5. MAPE comparison of Low-Fidelity, High-Fidelity, and Multi-Fidelity models for the 3rd scenario (source domain: Al F357; target domain: Al 2024).

Low-Fidelity Model	High-Fidelity Model	Multi-Fidelity Model
33.95%	29.55%	15.65%

The results show that the proposed framework, e.g., the MFM, can be used for adapting the knowledge acquired on the corrosion process of one or more materials to a new one, as in the presented work, since the Multi-Fidelity Model is the one showing the lowest MAPE in all the three scenarios analysed. In particular, it can be interesting to notice that the third scenario seems to be a trade-off of the first two scenarios. By comparing **Table 3**, **Table 4**, and **Table 5** it is possible to notice that the MAPE of both Low-Fidelity and Multi-Fidelity models for the 3rd scenario are in between the ones of the first two. Therefore, in cases where there are more discrepancies between the datasets, using two source domains may not be the best choice. However, this allows the comparison of the first two scenarios since the use of the Al 7475 data alone as the source domain showed the lowest MAPE. By looking at **Fig. 3(a)** it is possible to observe that the mean value and the standard deviation of the Al 7475 data at $3.0\ C$ are slightly greater than the ones of the Al F357. More specifically, the difference is of $1.5261\ \mu m$ and $2.2850\ \mu m$ for the mean and the standard deviation, respectively. Therefore, these small differences in the distribution of the data at $3.0\ C$ for Al 7475 and Al F357 can be the reason behind the slightly lower errors shown in the first scenario rather than in the second one.

3. Conclusion

The work presented focuses on a transfer learning framework for corrosion evolution estimation. The transfer learning approach proposed is based on the Multi-Fidelity Method (MFM) and it is applied to data acquired for electrodissoluted specimens. In particular, profilometries are used to collect maximum corroded depth and transversal area for 12 specimens for three aluminium alloys (Al 7475, Al 2024, and Al F357). The electrodissolution process takes place in a $NaCl\ 5\% + HCl\ 0.5\ M$ naturally aerated water-based solution and the charge was monitored during the process. When the charge reached the values of $1.0\ C$, $2.0\ C$, and $3.0\ C$ the electrodissolution was stopped to take profilometry measurements.

The transfer learning framework is developed assuming that Al 2024 data at $3.0\ C$ are missing, therefore the aim is to predict the corroded depth given the knowledge of the data coming from Al 7475 and Al F357. Therefore, the data of the latter alloys are used as source domain while the target domain consists of Al 2024 data acquired at $1.0\ C$ and $2.0\ C$. Three scenarios are studied in which the source domain consists of (1) Al 7475 alone, (2) Al F357 alone, and (3) both Al 7475 and Al F357.

The results showed that the first scenario is the one for which the proposed method gives the best results (lowest MAPE, Mean Absolute Percentage Error) and that the third scenario is a trade-off between the first and the second. This is due to the data distribution of the three alloys studied.

Future works may focus on comparing the proposed method with other approaches, like statistical methods or component analysis-based methods, such as Transfer Component Analysis for instance. In addition, transfer learning between different corrosion phenomena may be investigated.

In conclusion, this work paves the way towards the prediction of the residual useful life estimation of components because it is possible to evaluate the pit-crack transition and the residual mechanical strength given the size of the corrosion damage.

References

[1] N. Morizet, N. Godin, J. Tang, E. Maillet, M. Fregonese and B. Normand, "Classification of acoustic emission signals using waveltes and Random Forests: Application to localised corrosion," Mechanical System and Signal Processing, 2016.

[2] C. Jomdecha, A. Prateepasen and P. Kaewtrakulpong, "Study on source location using an acoustic emission system for various corrosion types," NDT & E International, 2007.

[3] A. Homborg, T. Tinga, X. Zhang, E. van Westing, P. Oonincx, J. de Wit and J. mol, "Time–frequency methods for trend removal in electrochemical noise data," *Electrochimica Acta,* 2012.

[4] S. Hoseinieh, A. Homborg, T. Shahrabi, J. Mol and B. Ramezanzadeh, "A Novel Approach for the Evaluation of Under Deposit Corrosion in Marine Environments Using Combined Analysis by Electrochemical Impedance Spectroscopy and Electrochemical Noise," *Electrochimica Acta,* 2016.

[5] B. Graig, R. Lane and D. Rose, "Corrosion Prevention and Control: A Program Management Guide for Material Selection," Advanced Materials, Manufacturing, and Testing Information, 2006.

[6] S. Harris, M. Hebbron and M. Mishon, "Corrosion Sensors to Reduce Aircraft Maintenance," in *NATO RTO-STOMP-AVT-144*, Vilnius, Lituania, 2006.

[7] D. Horner, B. Connolly, S. Zhou, L. Crocker and A. Turnbull, "Novel images of the evolution of stress corrosion cracks from corrosion pits," *Corrosion Science,* 2011.

[8] A. Turnbull, L. Wright and L. Crocker, "New insight into the pit-to-crack transition from finite element analysis of the stress and strain distribution around a corrosion pit," *Corrosion Science,* 2010.

[9] S.-H. W. Y. Xu, "Estimating the effects of corrosion pits on the fatigue life of steel plate based on the 3D profile," *International Journal of Fatigue,* 2015.

[10] "Relazione di indagine sull'incidente ferroviario del 29 giugno nella stazione di Viareggio," Ministero delle Infrastrutture e dei Trasporti - Direzione Generale per le Investigazioni ferroviarie, 2012.

[11] "Report of the enquiry into the sinking of the Erika off the coasts of Brittany on 12 December 1999," French Marine Accident Investigation Office, 1999.

[12] L. Bertolucci Coelho, D. Zhang, Y. Van Ingelgem, D. Steckelmacher, A. Nowè and H. Terryn, "Reviewing machine learning of corrosion prediction in a data-oriented perspective," *npj Materials Degradation ,* 2022.

[13] T. Sutojo, S. Rustad, M. Akrom, A. Syukur, G. Shidik and H. Dipojono, "A machine learning approach for corrosion small datasets," *npj Materials Degradation,* 2023.

[14] C. Taylor and B. Tossey, "High temperature oxidation of corrosion resistant alloys from machine learning," *npj Materials Degradation,* 2021.

[15] V. Vangrunderbeek, L. Bertolucci Coelho, D. Zhang, Y. Li, Y. Van Ingelgem and H. Terryn, "Exploring the potential of transfer learning in extrapolating accelerated corrosion test data for long-term atmospheric corrosion forecasting," *Corrosion Science,* 2023.

[16] G. Canonaco, M. Roveri, C. Alippi, F. Podenzani, A. Bennardo, M. Conti and N. Mancini, "A transfer-learning approach for corrosion prediction in pipeline infrastructures," *Applied Intelligence,* 2021.