Bridge Maintenance, Safety, Management, Digitalization and Sustainability – Jensen, Frangopol & Schmidt (eds)

Advanced framework for degradation modeling of operating structures

M.E.A. Ben Seghier, A. Jiménez Rios & J. Dai
Oslo Metropolitan University, Oslo, Norway

V. Plevris
Qatar University, Doha, Qatar

ABSTRACT: The safe operation of steel structures, such as bridges, is of paramount importance to mitigate potential issues. Consequently, the continuous and thorough monitoring of their operational conditions is imperative to uphold their safety and reliability. However, the inexorable process of corrosion, catalyzed by environmental conditions, leads to the inevitable deterioration of structural integrity over time. This research endeavors to predict the extent of corrosion in the primary cables of bridges through the application of advanced methodologies based on machine learning techniques. The execution of the proposed model necessitates the utilization of an extensive database encompassing diverse characteristics pertaining to the environmental properties of the surrounding region. The performance of the proposed models is rigorously assessed using a comprehensive suite of statistical and graphical metrics. The findings of this investigation underscore the effectiveness of the recommended solutions, surpassing previously established methodologies in addressing this pressing issue. The demonstrated success of the proposed model augurs favorably for its potential utility in furthering research into the dependability assessment of suspension bridge main cables.

1 INTRODUCTION

Suspension bridges rely on main cables made of high-strength carbon steel with a zinc protection layer (Barton Scott et al., 2000) to carry traffic loads and maintain structural stability. These cables are susceptible to corrosion, which depends on manufacturing and environmental factors (Chavel & Leshko, 2012). Current inspection methods involve removing the external cover to assess internal wire condition, but this provides limited information due to practical constraints and budget limitations (Mayrbaurl & Camo, 2004). Accurate prediction of the annual corrosion rate in these cables is crucial for ensuring safety, determining remaining useful life, and optimizing maintenance and replacement efforts while minimizing costs (Ben Seghier, Corriea, et al., 2021).

Corrosion growth on steel structures poses a formidable challenge, with substantial cost implications (El Amine Ben Seghier et al., 2018; Hussein Farh et al., 2023). Bridge cables, in particular, endure harsh environmental conditions, experiencing fluctuating temperatures and varied weather patterns over their lifespan (Kim et al., 2002). Numerous studies have delved into the critical environmental factors affecting metal corrosion in bridge cables, highlighting temperature (T), pH levels, humidity duration, contaminant and ion concentrations, and extended wetness periods as the key determinants (Saha, 2012). Bettie et al. (Betti & Yanev, 1999) discovered that pH values below 4 can result in water infiltration into internal cables in New York City suspension bridges, while Eiselstein and Caligiuri (Eiselstein & Caligiuri, 1987) emphasized the significance of water contaminant accumulation within the cables. Stahl

DOI: 10.1201/9781003483755-57

(Stahl & Gagnon, 1995) pointed out that issues could also be linked to imperfections in protective coatings or cable manufacturing. Further investigations (Deeble Sloane Matthew et al., 2013; Suzumura & Nakamura, 2004) have as well demonstrated the pivotal role of humidity and temperature fluctuations during the day in corrosion rates acceleration within main cables under ambient conditions. Collectively, these findings underscore the detrimental impact of environmental factors on bridge cable corrosion rates, emphasizing the need of appropriate advanced models for predicting corrosion rates using available environmental data from inspection reports.

The power-law model (Romanoff, 1957), is commonly used to predict corrosion rates in steel structures, including carbon steel wires. It relies on time and constant variables derived from simple fitting methods based on experimental data. However, its efficiency is limited, especially in terms of its applicability and precision to a larger database, resulting in inaccurate corrosion rate estimations in various scenarios (Kamrunnahar & Urquidi-Macdonald, 2010). Recently, the research community has turned its attention to machine learning techniques, recognizing their potential to effectively capture the intricate relationships through the implementation of complex anomaly detection algorithms (Jiménez Rios et al., 2023). Karanci and Betti (Karanci & Betti, 2018a) were pioneers in this area, implementing artificial intelligence (AI) models to model the annual corrosion rate in suspension bridge main cables. They compared linear regression, artificial neural networks (ANN), and support vector regression (SVR) to tackle the problem, but their approaches yielded less-than-ideal results, attributed to the complexity of the database and the limited generalization capabilities of the AI models. Ben Seghier et al. (Ben Seghier, Corriea, et al., 2021) optimized the single machine learning (ML) model control parameters using more advanced AI-models such as hybrid-ML techniques with meta-heuristic algorithms (i.e. marine predators algorithm (MPA)). When compared to single-ML models, their results were more accurate. This highlights the pressing need for more advanced AI models to enhance the predictive accuracy of such complex problems.

The primary objective of this research is to introduce a new predictive model designed to estimate the annual corrosion rate by analyzing environmental conditions from a global dataset. To address this challenge, we assess the performance of different ensemble learning models, namely, decision tree (DT), random forest (RF), adaptive boosting (ADB), and extreme gradient boosting (XGB). Furthermore, these models' performance is compared against four conventional regression techniques, i.e. multiple linear regression (MLR), ridge regression (RR), lasso regression (LR), and elastic net regression (ER). Various evaluation criteria are applied to assess the performance of ensemble and regression models in term of efficiency and accuracy.

2 METHODOLOGY

2.1 *Database description*

In this paper, we delve into an in-depth analysis of the intricate patterns of corrosion rate in suspension bridge main cables, which exhibit nonlinear behaviors contingent upon the environmental factors at play. To accomplish this, we draw upon a comprehensive database meticulously compiled by Karanci and Betti (Karanci & Betti, 2018a, 2018b). The database encompasses over 250 testing sites distributed across 33 countries (e.g. USA, Germany, Norway, Mexico, . . ., etc). Within this extensive collection, an uncover 309 distinct measurements are used, which shed light on the annual corrosion rate of carbon steel specimens, each subjected to diverse atmospheric conditions over the course of a year. The tests consisted on exposing carbon steel specimens to different atmospheric circumstances.

This study considers the impact of six crucial environmental factors, which serve as input variables. Specifically, these factors encompass temperature (T, °C), relative humidity (RH,%), the duration of moisture on the metal's surface (TOW, %), annual precipitation (P, mm), rainwater pH (pH), and chloride ion concentration (Cl^-, mg/L). These variables collectively represent the surrounding environmental conditions. The primary focus of this investigation centers on the annual corrosion rate (mm/year), which is regarded as the principal output parameter and is observed in the main cable wires. Table 1 presents a statistical summary of the database, including the mean (X_{Mean}), standard deviation (X_{STD}), the minimum value (X_{Min}) and the maximum value (X_{Max}). Figure 1 shows the Pearson correlation matrix of the studied parameters. It can be seen that there is moderate to low correlation between the input variables and the target output, indicating that all the utilized variables are suitable to be included in the models' development.

Table 1. Dataset description of the input and output variables.

Type	Variables	X_{Mean}	X_{STD}	X_{Min}	X_{Max}
Inputs	Temperature, °C	15.30	8.75	-3.10	29.82
	RH, %	0.46	0.18	0.00	0.98
	TOW, %	68.70	14.10	33.30	91.10
	Precipitation, mm	882.43	486.83	13.00	3677.00
	pH	4.99	0.87	3.45	7.37
	Cl^-, mg/L	12.35	27.17	0.01	192.73
Outputs	Corrosion rate, µm	38.53	43.76	3.30	376.70

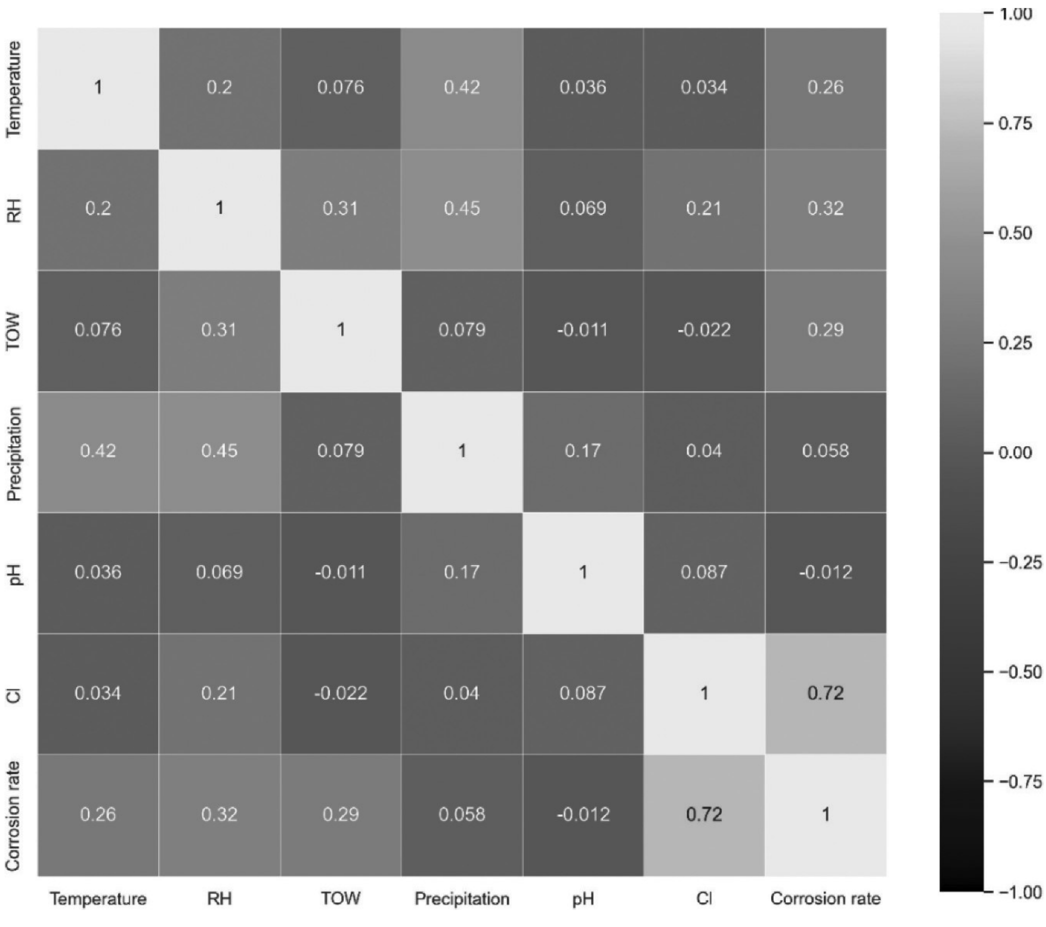

Figure 1. Pearson correlation matrix of the input and output variables.

2.2 *Regression models*

Four conventional regression techniques were implemented in this study as follows:

- Multiple linear regression (MLR): a statistical technique used in data analysis and modeling to examine the relationship between a dependent variable (or response variable) and two or more independent variables (or predictor variables).

- Ridge regression (RR): also known as L2 regularization, it is a linear regression technique used in statistics and machine learning to address the problem of multicollinearity and overfitting in MLR.
- Lasso regression (LR): an abbreviation for "Least Absolute Shrinkage and Selection Operator" regression, which is a linear regression technique used in statistics and machine learning for variable selection and regularization.
- Elastic net regression (ER): it is a linear regression technique that combines both Lasso (L1 regularization) and Ridge (L2 regularization) regression methods to address the limitations of each individual approach.

2.3 *Ensemble learning models*

Ensemble Learning (EL) models are machine learning (ML) techniques that combine the predictions of multiple individual models (base models or learners) to improve overall predictive performance, generalization, and robustness. The idea behind ensemble learning is that by aggregating the predictions of multiple models, the ensemble can often produce more accurate and reliable results than any single model on its own. One base model and three ensemble learning models were employed in this work, as follows:

- Decision tree (DT) –Base model-: used for both classification and regression tasks. It is a versatile and interpretable model that can be represented graphically as a tree structure.
- Random forest (RF) –EL-model-: a technique that combines the predictive strength of multiple DT to produce more accurate and robust results.
- Adaptive boosting (ADB) -EL-model-: a technique that designed to improve the performance of weak learners (classifiers that perform slightly better than random guessing) by combining them into a strong ensemble model.
- Extreme gradient boosting (XGB) -EL-model-: an extension of the gradient boosting machine (GBM) algorithm, designed to be highly optimized, scalable, and capable of delivering excellent predictive performance.

Besides, a k-fold cross-validation was used to prevent over-fitting of all EL-models and regression approaches. The data were divided into 70% for building the predictive models, while 30% were attributed for the validation of the developed predictive models. The performance of all models was measured based on comparative statistical indicators as described in the following (Ben Seghier, Gao, et al., 2021; Ben Seghier, Kechtegar, et al., 2021):

1. Root mean square error (RMSE):

$$RMSE = \frac{1}{n}\sum_{i=1}^{n}\left(C_{rate,i}^{Act} - C_{rate,i}^{Pre}\right)^2 \tag{1}$$

2. Mean absolute error (MAE):

$$MAE = \frac{1}{n}\sum_{i=1}^{n}\left|C_{rate,i}^{Act} - C_{rate,i}^{Pre}\right| \tag{2}$$

3. Coefficient of determination (R^2):

$$R^2 = 1 - \frac{\sum_{i=1}^{n}\left(C_{rate,i}^{Act} - C_{rate,i}^{Pre}\right)^2}{\sum_{i=1}^{n}\left(C_{rate}^{Avg} - C_{rate,i}^{Pre}\right)^2} \tag{3}$$

4. Confidence interval (CI):

$$CI = WI \times NSE \tag{4}$$

$$WI = 1 - \frac{\sum_{i=1}^{n}\left(C_{rate,i}^{Act} - C_{rate,i}^{Pre}\right)^2}{\sum_{i=1}^{n}\left(\left|C_{rate,i}^{Pre} - C_{rate}^{Avg}\right| + \left|C_{rate,i}^{Act} - C_{rate}^{Avg}\right|\right)^2} \tag{5}$$

$$NSE = 1 - \frac{\sum_{i=1}^{n}\left(C_{rate,i}^{Act} - C_{rate,i}^{Pre}\right)^2}{\sum_{i=1}^{n}\left(C_{rate,i}^{Act} - C_{rate}^{Avg}\right)^2} \tag{6}$$

5. U95:

$$U95 = 1.96 \times \sqrt{Standard\ deviation^2 - RMSE^2} \tag{7}$$

Where $C_{rate,i}^{Act}$, $C_{rate,i}^{Pre}$ and C_{rate}^{Avg} are the actual, predicted, and average values of the *i*-th annual corrosion rate from the test result measurements. The lower the RMSE, MAE, and U95 values, the higher the projected values by the utilized model, and the higher the CI and R^2 values, the higher the model accuracy.

3 RESULTS AND DISCUSSION

3.1 *Predictive regression models*

The performance evaluation of the four regression models implemented in this study is presented in Table 2. It can be seen from the obtained results that the best performing regression model among the utilized ones is the ER. This latter yielded the lowest RMSE, MAE and U95 values flowed by LR, while the MLR yielded the lowest performance during both phases, training, and testing. In addition, the highest obtained R^2 value is 0.825 and 0.742 by the EL during the training and testing sets respectively. The overall results show low adaptation to the complexity of the used database, indicating the drawbacks of the used regression models.

Table 2. Performance evaluation of the four regression models.

Phase	Models	RMSE	MAE	R^2	CI	U95
Training	MLR	22.991	16.331	0.758	0.530	46.134
	RR	23.528	15.707	0.780	0.479	47.107
	LR	21.910	15.236	0.803	0.516	43.899
	ER	**19.859**	**14.328**	**0.825**	**0.576**	**39.673**
Testing	MLR	29.421	16.391	0.543	0.239	58.225
	RR	20.442	14.480	0.679	0.457	40.545
	LR	18.834	13.427	0.727	0.481	37.743
	ER	**18.245**	**13.310**	**0.742**	**0.501**	**36.565**

* Bold numbers represent the best results.

Furthermore, Figure 2 presents the scatter plots of the different employed regression models. Besides from the relatively low prediction capacity of the regression models explored in this paper, another important drawback worth highlighting is the fact that such models could result into negative values of annual corrosion rate (see dots plotted below the horizontal axis, close to the origin), which does not hold any physical meaning. Thus, the applicability of such models can result on misleading values of the annual corrosion rate.

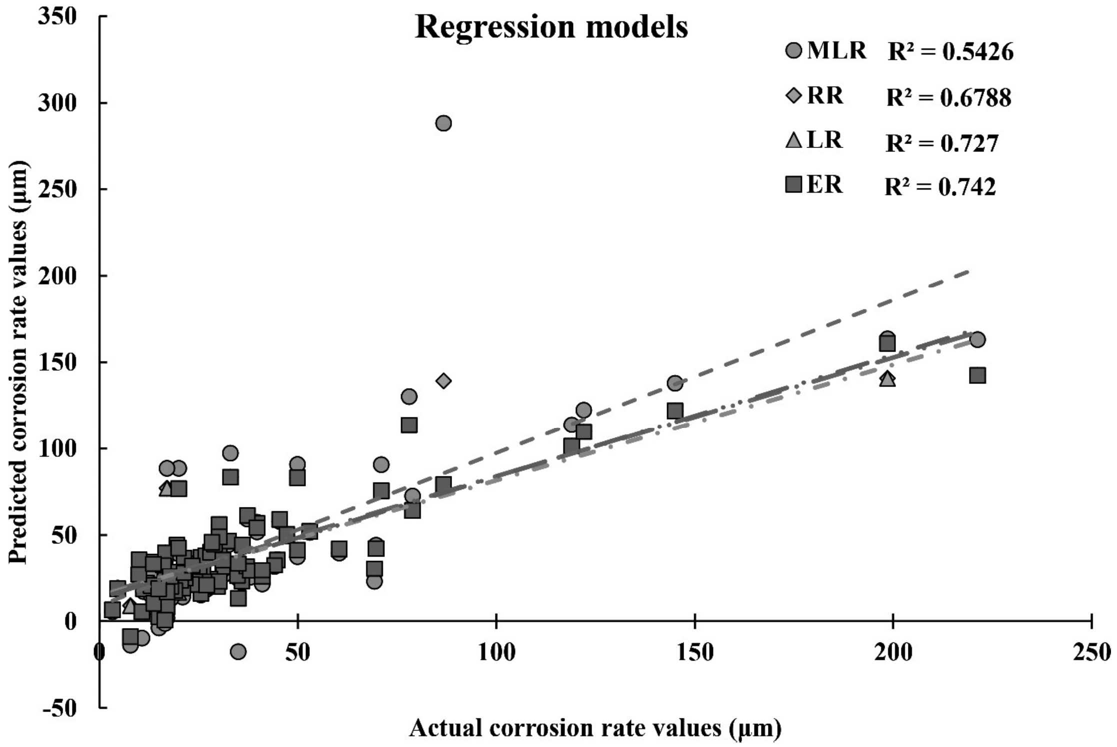

Figure 2. Scatter plots of the regression models.

3.2 *Predictive ensemble learning models*

The performance evaluation of the single based models (i.e. DT) and the three ensemble learning models (i.e. RF, ADB and XGB) implemented in this study are presented in Table 3. According to the results, the DT model showed perfect results during the training phase, whereas the model yielded the lowest results during the testing phase, indicating an overfitting and inaccurate results obtained by this technique a single base model. The DT results are obtained as follow: RMSE = 37.999 µm/year, MAE=16.997 µm/year and CI=-0.117, which is the lowest among all other ensemble learning models. The best performing ensemble learning algorithm is the XGB, where this observation is based on all the utilized statistical indicators, including the obtained R^2 value (0.999 and 0.941 for the training and testing sets respectively). These results highlight the suitability of ensemble learning models to be implemented in annual corrosion prediction tasks.

Table 3. Performance evaluation of the three ensemble learning models.

Phase	Models	RMSE	MAE	R^2	CI	U95
Training	DT**	0.000	0.000	1.000	0.750	0.000
	RF	9.579	5.245	0.970	0.684	19.770
	XGB	**1.604**	**1.145**	**0.999**	**0.746**	**4.010**
	ADB	11.030	8.492	0.954	0.694	20.897
Testing	DT	37.999	16.997	0.380	-0.117	74.978
	RF	14.876	10.966	0.863	0.549	29.784
	XGB	**8.858**	**6.733**	**0.941**	**0.678**	**18.315**
	ADB	14.562	12.173	0.887	0.582	25.936

* Bold numbers represent the best results.
** DT shows perfect results due to the overfitting problem.

Similar to the previous section, Figure 3 presents the scatter plots of the base single model (i.e. DT) and the different ensemble learning models using the testing phase results. Besides from the higher performance already commented, it can be observed that under no circumstances, ensemble learning models provide negative values of annual corrosion rate, thus respecting the physical meaning of the task at hand. In addition, the performance of the models can be ranked as follows: XGB>ADB>RF>DT.

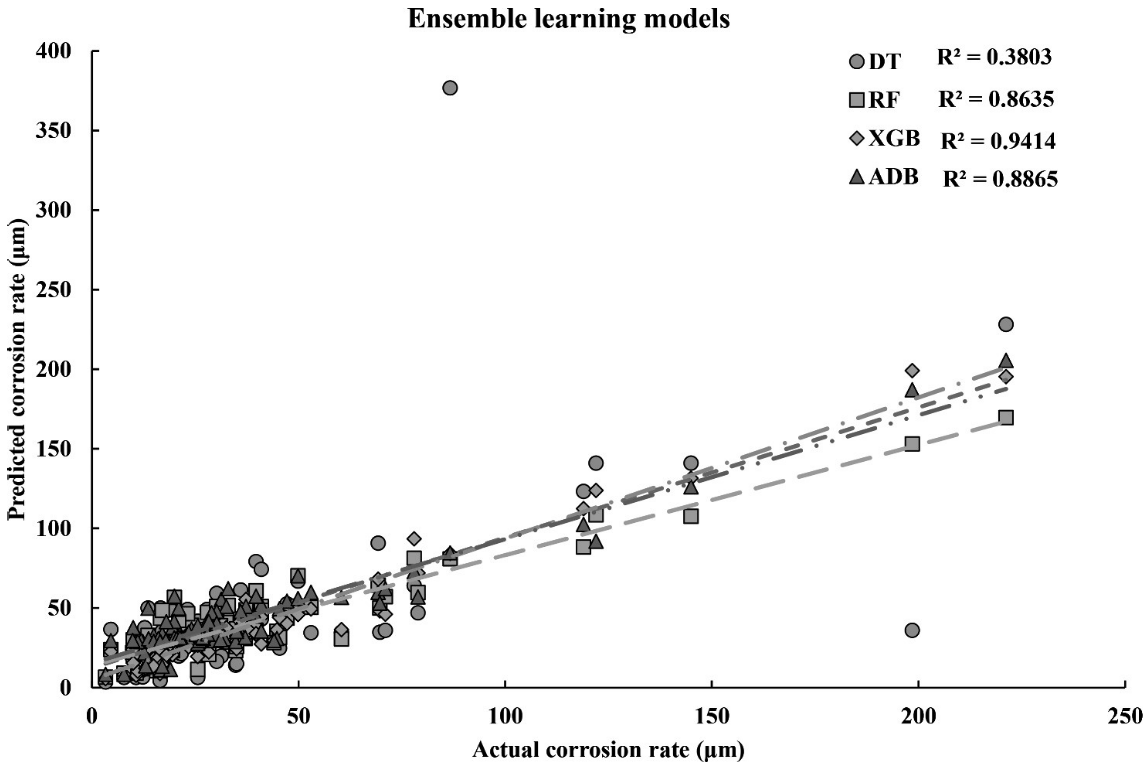

Figure 3. Scatter plots of the ensemble learning models.

4 CONCLUSION

The prediction of the annual corrosion rate in suspension bridge main cables presents a complex and nonlinear issue characterized by chaotic patterns and stochastic tendencies. In this paper, we have assessed the performance of different ensemble learning models, namely, decision tree (DT), random forest (RF), adaptive boosting (ADB), and extreme gradient boosting (XGB), to predict annual corrosion rates using a comprehensive database composed of 309 measurements of annual corrosion rate of carbon steel specimens as a function of diverse atmospheric conditions over the course of a year, distributed across 33 countries. Furthermore, these models' performance was compared against four conventional regression techniques, i.e. multiple linear regression (MLR), ridge regression (RR), lasso regression (LR), and elastic net regression (ER).

In general, except for the DT model, the ensemble learning models presented a better performance than the conventional regression models. XGB was the best performing ensemble learning algorithm with R^2 values reaching 0.999 and 0.941 for the training and testing sets, respectively. Moreover, all ensemble learning models respected the physical meaning of the predicted outcome parameter and none of them resulted in negative estimated values.

REFERENCES

Barton Scott, C., Vermaas Garry, W., Duby Paul, F., West Alan, C., & Betti, R. (2000). Accelerated Corrosion and Embrittlement of High-Strength Bridge Wire. *Journal of Materials in Civil Engineering*, *12*(1), 33–38. https://doi.org/10.1061/(ASCE)0899-1561(2000)12:1(33)

Ben Seghier, M. E. A., Corriea, J. A. F. O., Jafari-Asl, J., Malekjafarian, A., Plevris, V., & Trung, N.-T. (2021). On the modeling of the annual corrosion rate in main cables of suspension bridges using combined soft computing model and a novel nature-inspired algorithm. *Neural Computing and Applications*, *33*(23), 15969–15985. https://doi.org/10.1007/s00521-021-06199-w

Ben Seghier, M. E. A., Gao, X.-Z., Jafari-Asl, J., Thai, D.-K., Ohadi, S., & Trung, N.-T. (2021). Modeling the nonlinear behavior of ACC for SCFST columns using experimental-data and a novel evolutionary-algorithm. *Structures*, *30*, 692–709. https://doi.org/10.1016/j.istruc.2021.01.036

Ben Seghier, M. E. A., Kechtegar, B., Nait Amar, M., A.F.O. Correia, J., & Trung, N.-T. (2021). Simulation of the ultimate conditions of fibre-reinforced polymer confined concrete using hybrid intelligence models. *Engineering Failure Analysis*, *128*, 105605. https://doi.org/10.1016/j.engfailanal.2021.105605

Betti, R., & Yanev, B. (1999). Conditions of Suspension Bridge Cables: New York City Case Study. *Transportation Research Record*, *1654*(1), 105–112. https://doi.org/10.3141/1654-12

Chavel, B. W., & Leshko, B. J. (2012). *Primer for the Inspection and Strength Evaluation of Suspension Bridge Cables* (FHWA-IF-11-045). U.S. Department of Transportation Federal Highway Administration. https://www.fhwa.dot.gov/bridge/pubs/if11045.pdf

Deeble Sloane Matthew, J., Betti, R., Marconi, G., Hong Ah, L., & Khazem, D. (2013). Experimental Analysis of a Nondestructive Corrosion Monitoring System for Main Cables of Suspension Bridges. *Journal of Bridge Engineering*, *18*(7), 653–662. https://doi.org/10.1061/(ASCE)BE.1943-5592.0000399

Eiselstein, L. E., & Caligiuri, R. D. (1987). *Atmospheric corrosion of the suspension cables on the Williamsburg Bridge*. ASTM International. https://www.astm.org/stp25841s.html

El Amine Ben Seghier, M., Keshtegar, B., & Elahmoune, B. (2018). Reliability analysis of low, mid and high-grade strength corroded pipes based on plastic flow theory using adaptive nonlinear conjugate map. *Engineering Failure Analysis*, *90*, 245–261. https://doi.org/10.1016/j.engfailanal.2018.03.029

Hussein Farh, H. M., Ben Seghier, M. E. A., Taiwo, R., & Zayed, T. (2023). Analysis and ranking of corrosion causes for water pipelines: a critical review. *npj Clean Water*, *6*(1), 65. https://doi.org/10.1038/s41545-023-00275-5

Jiménez Rios, A., Plevris, V., & Nogal, M. (2023). Bridge Management through Digital Twin-based Anomaly Detection Systems: A Systematic Review. *Frontiers in Built Environment*. https://doi.org/10.3389/fbuil.2023.1176621

Kamrunnahar, M., & Urquidi-Macdonald, M. (2010). Prediction of corrosion behavior using neural network as a data mining tool. *Corrosion Science*, *52*(3), 669–677. https://doi.org/10.1016/j.corsci.2009.10.024

Karanci, E., & Betti, R. (2018a). Modeling Corrosion in Suspension Bridge Main Cables. I: Annual Corrosion Rate. *Journal of Bridge Engineering*, *23*(6), 04018025. https://doi.org/10.1061/(ASCE)BE.1943-5592.0001233

Karanci, E., & Betti, R. (2018b). Modeling Corrosion in Suspension Bridge Main Cables. II: Long-Term Corrosion and Remaining Strength. *Journal of Bridge Engineering*, *23*(6), 04018026. https://doi.org/10.1061/(ASCE)BE.1943-5592.0001234

Kim, H.-K., Lee, M.-J., & Chang, S.-P. (2002). Non-linear shape-finding analysis of a self-anchored suspension bridge. *Engineering Structures*, *24*(12), 1547–1559. https://doi.org/10.1016/S0141-0296(02)00097-4

Mayrbaurl, R. M., & Camo, S. (2004). *Guidelines for Inspection and Strength Evaluation of Suspension Bridge Parallel Wire Cables*. The National Academies Press. https://doi.org/10.17226/23338

Romanoff, M. (1957). *Underground corrosion* US Government Printing office. https://nvlpubs.nist.gov/nistpubs/Legacy/circ/nbscircular579.pdf

Saha, J. K. (2012). *Corrosion of constructional steels in marine and industrial environment: frontier work in atmospheric corrosion*. Springer Science & Business Media. https://link.springer.com/book/10.1007/978-81-322-0720-7

Stahl, F. L., & Gagnon, C. P. (1995). *Cable corrosion in bridges and other structures: Causes and solutions*. American Society of Civil Engineers (ASCE).

Suzumura, K., & Nakamura, S.-i. (2004). Environmental Factors Affecting Corrosion of Galvanized Steel Wires. *Journal of Materials in Civil Engineering*, *16*(1), 1–7. https://doi.org/10.1061/(ASCE)0899-1561(2004)16:1(1)