Time series analysis for database completion and forecast of sensors measurements: application to concrete structures.

Luis Rincon[1,2], Bassel Habeeb[3], Emilio Bastidas-Arteaga[2], Elsa Eustáquio[4], Ameur Hamami[2], João Marcelino[4], Luís Oliveira Santos[4], Jose C. Matos[1], Yina F. Muñoz[1]

[1] *Universidade do Minho, ISISE, ARISE, Department of Civil Engineering, Guimarães, Portugal, id9635@alunos.uminho.pt*

[2] *Laboratory of Engineering Sciences for Environment UMR CNRS 7356, La Rochelle University, France, luis.rincon_prada@univ-lr.f*

[3] *Institute for Research in Civil and Mechanical Engineering UMR CNRS 6183, University of Nantes*

[4] *Laboratório Nacional de Engenharia Civil, Lisboa, Portugal*

ABSTRACT Despite the high durability level associated with reinforced concrete bridges, they are nonetheless susceptible to natural hazards and extreme events can impair their performance and serviceability throughout their lifespan. For that reason, maintenance, rehabilitation, and repair actions on existing structures are projected to rise and currently account for about 50% of the construction sector spending in most developed nations. To establish long-term maintenance schedules, it is vital to know the state of a structure and its degradation over time. Therefore, the monitoring of structures has become a necessary task to guarantee their use throughout their lifespan. Maintenance and inspection schemes depend on these systems that periodically or continuously collect information using chemical, optical, sound sensors, among others. However, the reliability of these sensors depends on environmental factors, durability, and even power outages. When any of these factors affect the sensors, their acquisition of information can be interrupted temporarily or permanently. This paper focuses on the competition of this missing data. The study uses one year of data from sensors monitoring a reinforced concrete structure that suffered interruptions in the acquisition processes. To reduce possible uncertainties that affect the analysis of the degradation of the materials and the reliability of the structures, the database of concrete electrical resistivity and concrete temperature of the sensors were analyzed, and time-series analysis method, artificial neural network models and generalized linear and non-linear models were used specifically to fill in the missing database values and perform predictions. Finally, the results are discussed, and recommendations are established for the application of this methodology for the analysis of the sensors used.

Keywords Autoregressive integrated moving average models, Structural health monitoring, Sensors, Artificial neural network models, Generalized linear models.

I. INTRODUCTION

The long-term durability and safety of a structure depend on the conditions to which it is exposed, such as extreme events, unpredicted load and natural hazards. According to global statistics on bridge collapses, natural hazards are the primary reason for failure. The French government recently disclosed that out of the 12,000 maintained bridges, 840 are at risk of collapse following the incident of the Genoa motorway bridge collapse, a problem that is widespread throughout Europe as well (Willsher, Tondo and Henley, 2018). These conditions must be considered during the design, construction, and maintenance phases. Corrosion is one of the primary issues faced by concrete structures in coastal or estuary areas, where the presence of chloride ions, as well as carbon dioxide, can penetrate concrete and compromise the structural integrity of the bridge reducing its service life. For the contamination with chlorides, the corrosion process occurs primarily through the reaction described in Equation 1. Corrosion begins when the concentration of chlorides in the corrosion cell reaches a threshold value, destroying the protective passive film, and causing the ferrous ions to react with hydroxyl ions present in concrete. This reaction results in the production of a white precipitate of ferrous hydroxide which oxidizes to the more familiar forms of brown-toned oxides (Bastidas-Arteaga, 2009).

$$Fe^{2+} + 2OH^{-} \rightarrow Fe(OH)_2 \tag{1}$$

In natural exposure, the corrosion of reinforcing steel can be highly variable due to uncertainties in concrete properties, environmental conditions, and other factors (Marsh and Frangopol, 2008).

In recent years, there has been a growing interest in the use of sensors in concrete structures to monitor their condition and detect early corrosion (Llorens, Serrrano and Valcuende, 2019; Shevtsov *et al.*, 2022). Sensors provide real-time information about the condition of the structure, which can be crucial for making informed decisions about maintenance and repair. Concrete resistivity sensors proved to be durable for long-term monitoring and are particularly important in coastal areas where exposure to saltwater can accelerate the corrosion process, making regular monitoring of concrete resistivity a crucial part of structural maintenance (Figueira, 2017). Concrete resistivity is a critical parameter in the durability of reinforced concrete structures, as it measures the electrical resistance of concrete, which is a key indicator of the presence of chloride ions that can lead to corrosion initiation of the steel reinforcement within the structure (Azarsa and Gupta, 2017). Early signs of corrosion can be detected, and appropriate measures can be taken to prevent further damage, which can greatly extend the life of the structure.

Several intrinsic and external factors may affect the electrical resistivity of the concrete (Azarsa and Gupta, 2017). Temperature is one of the key influence factors influencing concrete resistivity, as the temperature of concrete increases, its electrical resistivity decreases (Presuel and Liu, 2012). Mainly, the increase in temperature causes a growth in the mobility of ions within the pore solution of the concrete, which in turn leads to an increase in the electrical conductivity of concrete, generally expressed as an inverse linear relation to electrical resistivity (Pereira *et al.*, 2009; Presuel and Liu, 2012). Therefore, when using sensors to measure the electrical resistivity of concrete, it is common to install temperature sensors to account for these variations in the data analysis (Azarsa and Gupta, 2017).

One of the challenges associated with long-term structural health monitoring using electrical resistivity sensors is the possibility of discontinuous measurements and missing data. This can occur when the sensor is unable to make continuous measurements due to various factors, such as power outages, sensor malfunctions, or data transmission issues. For the data analysis, it is crucial to carefully consider the data collection process and ensure that the sensor is functioning properly to minimize the occurrence of discontinuous measurements and missing data. However, even when the sensor is functioning properly, there may be instances where certain data points are missing due to signal noise or other factors. The presence of discontinuous measurements and missing data can significantly affect the accuracy of the data collected, making it difficult to obtain a complete picture of the phenomena. To address this issue, researchers have developed various statistical forecasting methods that can be used to estimate missing data points based on the available information (Habeeb *et al.*, 2021). Despite the accuracy of these, data imputation methods are highly dependent on the quality and quantity of the available data (Habeeb *et al.*, 2021).

This article presents a study on the use of statistical forecasting methods in monitoring systems for concrete structures near coastal areas. Data from temperature and concrete electrical resistivity sensors from a concrete structure in Portugal are analyzed and statistical forecasting methods are implemented in the recovery of missing information. The results suggest that these methodologies can be useful tools to improve the quality of sensor data and increase the effectiveness of monitoring systems in the early detection of corrosion and other structural problems. The paper is structured as follows: Section 2 presents the methodology implemented. Section 3 presents the application's information. Section 4 describes the results and discussion of the procedure and the forecast. Finally, the conclusions of the research are presented.

II. METHODOLOGY

A. Times series forecasting model.

Time series forecasting models make future predictions based on the statistical information of historical data considering pattern recognition. This method uses several techniques to extract important statistics and characteristics from time series data, like the trend, seasonality, and irregular components. By quantifying the main features of data and random variation, time-series analysis has become a widely applicable approach (Habeeb *et al.*, 2021).

Autoregressive Integrated Moving Average (ARIMA) models are among the most widely used statistical models for short-term time series analysis (Ho and Xie, 1998). These models have three fundamental characteristics: the autoregressive (AR) model, the moving average (MA) model and the integrated differencing (I) model. The autoregressive part of the model uses past values of the time series to predict future values and can be described as

$$X_t = c + \sum_{i=1}^{p} \varphi_i X_{t-i} + \varepsilon_t \; ; \; \varepsilon_t \sim WN(0, \sigma_\varepsilon^2) \tag{2}$$

where X_t is the state, c is a constant, φ_i is a parameter of the model, ε_t is a random white noise WN and σ_ε^2 is the variance of the random white noise.

The moving average (MA) component of the model uses the error of the previous prediction to improve the accuracy of the current prediction and can be described as

$$X_t = \omega + \sum_{i=1}^{q} \theta_i \varepsilon_{t-i} + \varepsilon_t \tag{4}$$

where θ is a parameter of the model and ω, often equals to zero, is the expectation of X_t.

The ARIMA model is used to fit the historical dataset of the resistivity sensors, in which the parameters fitted by the model are used to represent the Box & Jenkins forecasting 95 % confidence intervals. In addition to capturing the historical patterns of the dataset (Box and Jenkins, 1990).

B. Artificial Neural Network models (ANN)

ANN is based on the understanding and imitation of the human brain, where each neuron processes the information separately and simultaneously (Ukrainczyk, Banjad and Ukrainczyk, 2004). ANN models have existed since 1943 (Landahl, McCulloch and Pitts, 1943), however, have become popular, as they overcome the deficiencies of mechanistic and statistical models in prediction and only be applied to certain research areas in civil engineering such as material simulation (Hegazy, Tully and Marzouk, 1998; O. Akande *et al.*, 2014), material deterioration (Huang, 2010; Chou, Ngo and Chong, 2017; Kim, Kim and Lee, 2020; Rincon *et al.*, 2022), among others.

C. Generalized Linear and Non-Linear Models

Generalized linear models (GLM) are a class of statistical models that allow modeling the relationship between a response variable and one or more predictor variables, using a link function and probability distribution appropriated for the type of data being modeled (Dunn, 2023). GLM has been widely used in the estimation of nonlinear models in statistics (Yang, Yu and Zhong, 2023). On the other hand, generalized nonlinear models (GNLM) are an extension of GLMs that allow modeling nonlinear relationships between response and predictor variables through nonlinear link functions, a useful factor when the relationship between variables is suspected to be nonlinear, or when the linear link function is not adequate to model the relationship between variables (De Marco *et al.*, 2013). The present research used GLM and GNLM to forecast the missing data of concrete temperature considering a normal distribution of data given by the relationship between the concrete temperature and the concrete resistivity sensor.

III. APPLICATION DESCRIPTION

A. The test bed

Concrete electrical resistivity and temperature data were collected from a reinforced concrete bridge located in Portugal. The data were measured between July 2015 and August 2016, with a daily periodicity.

B. Sensors

Concrete electrical resistivity was measured using a two-graphite electrode resistivity sensor that consisted of measuring the electrical resistance of concrete through the insertion of two graphite electrodes (height=10mm and diameter=8 mm) with a set spacing (50 mm). Its installation was done by removing a concrete core and installing the electrodes in two 81 mm holes followed by sealing the core. Electric contacts from the electrodes were protected using a two components epoxy resin.

The temperature was measured using RTD platinum temperature sensors installed in the same core of the two graphite electrode resistivity sensors (Bruno *et al.*, 2021). Data acquisition was performed with a Datataker DT80 Universal Input Data Logger.

C. Sensor's output

The two graphite electrode resistivity sensors took a daily measure of the concrete electrical resistivity of the bridge, measured in ohms as shown in Figure 1(left). Some data are missing due to problems with the data acquisition system or the power supply unit of the data acquisition system. The concrete temperature was measured in degrees Celsius using the same daily frequency.

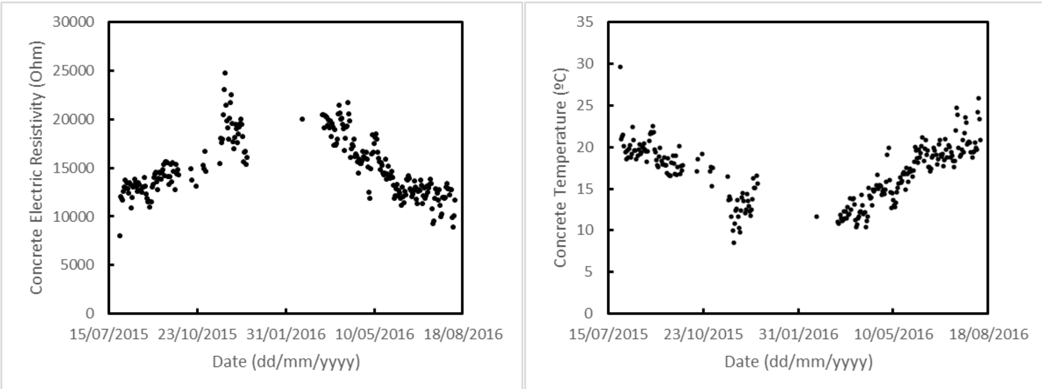


FIGURE 1. (left) Concrete electric resistivity and (right) Concrete temperature of the bridge analyzed.

IV. RESULTS AND DISCUSSION

A. Data filling for concrete resistivity

Figure 2 shows the period analyzed from September 2015 to March 2016, which contemplates the loss of data at seven points in time, the size days of these missing data vary between 1 to 61 days as seen in Table 1. In these seven gaps, no daily temperature or resistivity values were recorded. An additional gap is created to compare the forecast with measured data (Table 2). Figure 2 (Left) shows the localization in time of the missing data in orange and the artificial gap created in blue.

TABLE 1. Information on the missing data

Initial date	Ending date	Gap size
27/09/2015	28/09/2015	1
03/10/2015	16/10/2015	13
18/10/2015	22/10/2015	4
23/10/2015	30/10/2015	7
03/11/2015	18/11/2015	15
20/12/2015	19/02/2016	61
20/02/2016	13/03/2016	22

TABLE 2. Information on the artificial gap created.

Initial date	Ending date	Gap size
30/06/2016	22/07/2016	22

For the data-filling process, we first used the ARIMA method to establish a fitting model that will be used by the ANN model to make the forecasting. Figure 2(Right) shows in detail the prediction results for the second gap contemplated between October 3 and 16, 2015. It also shows the lower and upper limits of the Box Jenkins method that establishes with a reliability of 95% the area where the missing data may be located. With this information and the ARIMA model, the gap data are trained and predicted using ANN.

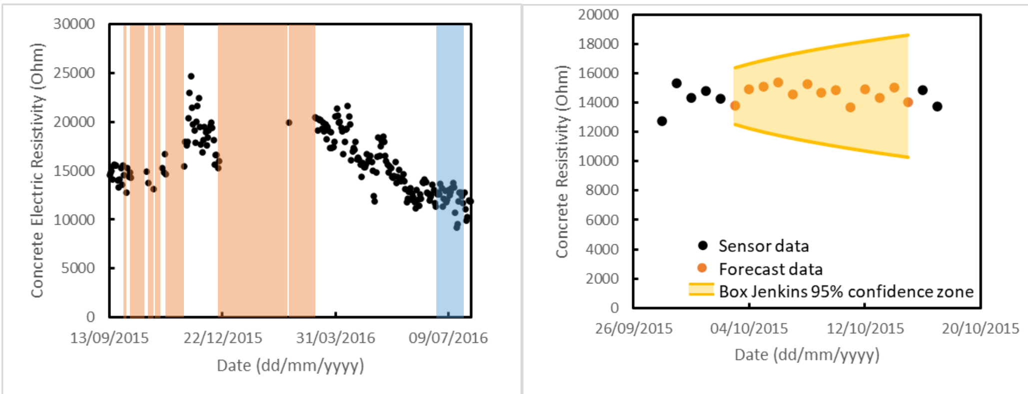


FIGURE 2. **(Left) Localization of the missing data (orange) and an artificial gap for analysis purposes (blue), and (Right) concrete resistivity forecasting from 3 to 16 October 2015.**

Table 3 presents the main error metrics of the predicted values for the different gaps using the above methodology. The results suggest that in general, the model may be providing fairly accurate predictions according to the fitting ARIMA model except for the first gap which predicts the concrete resistivity of a single day and the RMSE-MAE value suggest that the prediction may not be accurate.

TABLE 3. Errors indicators for the forecast data of concrete resistivity.

Model	ME	RMSE-MAE	MAPE (%)
Gap 1	0.21	5.01	1.29
Gap 2	-0.02	0.09	0.11
Gap 3	0.00	0.45	0.15
Gap 4	-0.04	0.05	0.08
Gap 5	-0.01	0.07	0.08
Gap 6	0.00	0.13	0.05
Gap 7	0.01	0.07	0.06
Artificial gap	0.03	0.17	0.06

B. Data filling for temperature

For the forecast of the temperature values, the correlation between temperature and resistivity was analyzed, as shown in Figure 3, where an inversely proportional relationship between the analyzed variables can be observed.

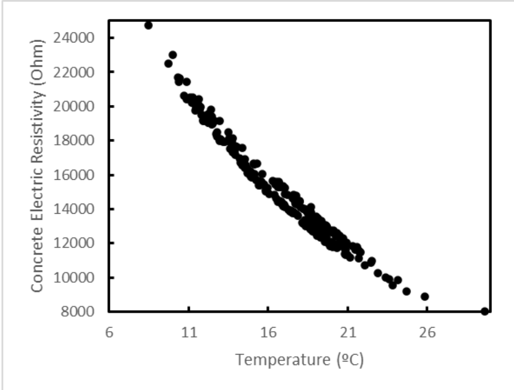


FIGURE 3. **Relation between concrete resistivity and temperature sensors.**

From this relationship, generalized linear and nonlinear models were used to predict temperature values. Normal distributions were considered and a potential model for GNLM was established. Figure 4 presents the results for the GLM and GNLM models, where a good fit with the existing data is observed, but the change from descending to ascending trend found in the winter solstice zone presents a small trend of temperature increase that could be corrected if the predictive models had enough data to consider the seasonality.

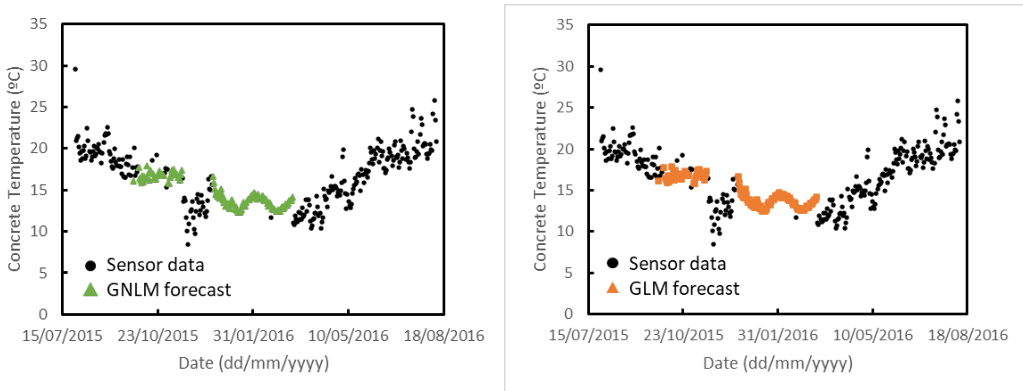


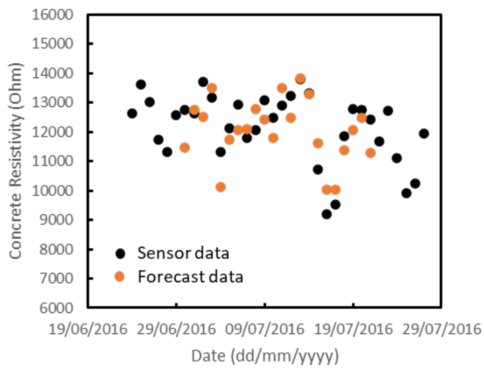
FIGURE 4. **Temperature forecast of GLM and GNLM**

C. Accuracy of predictive models

An analysis was performed to verify the correlation of the data with the predictions. For this, an artificial gap of 23 days was established between June 30 and July 22, 2016. Figure 5 shows the resistivity measured by the sensor and that forecasted using ANN, and the temperature predictions measured by the sensor and the forecast using the GLM and GNLM models. Both figures show an adequate approximation. The main error indicators were calculated, such as mean error (ME), square root of the average of the square errors (RMSE), mean absolute error (MAE), mean absolute percentage error (MAPE) and R^2 for the virtual gap analyzed (See Table 4). The results indicate that the two models used have a small deviation from the true values according to MAPE and present decent estimates according to R^2. However, the combination of high RMSE-MAE and relatively low MAPE suggests that the model may need further improvement.

TABLE 4. Errors indicators for the forecast data of temperature.

Model	ME	RMSE-MAE	MAPE (%)	R^2 (%)
GLM	0.78	0.28	5.48	67.5
GNLM	0.74	0.27	5.36	67.5



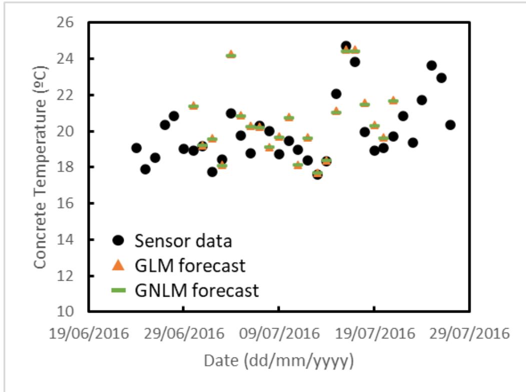


FIGURE 5. **(Left) Resistivity and (right) temperature forecast for the artificial gap.**

CONCLUSIONS

The time series prediction using the ARIMA method and the ANN proposes a suitable solution to the problem of missing data from structural health monitoring systems. The ARIMA model provides R^2value of 65.9% which gives an acceptable value for the forecasting performed.

The proposed GNLM provides R^2value of 67.5%, accuracy in forecasting, similar to GLM with an $R^2$of 67.5%. The amount of data that the models can forecast with an acceptable error was not analyzed in this article. These models fit adequately due to the clear relationship between temperature and resistivity, which allows for filling the missing temperature gaps.

However, it is important to note that the main function of the resistivity sensor is to indicate the presence of chloride ingress in the concrete. During this period the relationship between resistivity and temperature remained constant. However, this relationship can vary with the chloride ions concentration, making the forecast of temperature from concrete resistivity a less accurate procedure if the reduction in electrical resistance with increasing chlorides is not considered. Further work in this area will focus on considering this non-linear relationship as well as integrating information from other sensors or environmental measurements to improve the completion algorithms.

Much of the missing data for the problem analyzed is located around the winter solstice, which can be one of the days with the lowest temperatures, but this is not reflected in the prediction. Therefore, a model that considers seasonality could give better results in the maximum and minimum temperature events.

ACKNOWLEDGES

This work is financed by national funds through FCT - Foundation for Science and Technology, under grant agreement 2021.05862.BD attributed to the first author.

This work was partly financed by FCT / MCTES through national funds (PIDDAC) under the R&D Unit Institute for Sustainability and Innovation in Structural Engineering (ISISE), under reference UIDB / 04029/2020, and under the Associate Laboratory Advanced Production and Intelligent Systems ARISE under reference LA/P/0112/2020.

REFERENCES

Azarsa, P. and Gupta, R. (2017) 'Electrical Resistivity of Concrete for Durability Evaluation: A Review', *Advances in Materials Science and Engineering*, 2017, pp. 1–30. Available at: https://doi.org/10.1155/2017/8453095.

Bastidas-Arteaga, E. (2009) *Probabilistic service life modeling of RC structures subjected to the combined effect of chloride-induced corrosion and cyclic loading*. Universidad de los Andes. Available at: https://tel.archives-ouvertes.fr/tel-01065817.

Box, G.E.P. and Jenkins, G. (1990) *Time Series Analysis, Forecasting and Control*. USA: Holden-Day, Inc.

Bruno, S. *et al.* (2021) 'Technical Proposal for Monitoring Thermal and Mechanical Stresses of a Runway Pavement', *Sensors*, 21(20), p. 6797. Available at: https://doi.org/10.3390/s21206797.

Chou, J.-S., Ngo, N.-T. and Chong, W.K. (2017) 'The use of artificial intelligence combiners for modeling steel pitting risk and corrosion rate', *Engineering Applications of Artificial Intelligence*, 65(August), pp. 471–483. Available at: https://doi.org/10.1016/j.engappai.2016.09.008.

Dunn, P.K. (2023) 'Generalized linear models', in *International Encyclopedia of Education(Fourth Edition)*. Elsevier, pp. 583–589. Available at: https://doi.org/10.1016/B978-0-12-818630-5.10077-6.

Figueira, R. (2017) 'Electrochemical Sensors for Monitoring the Corrosion Conditions of Reinforced Concrete Structures: A Review', *Applied Sciences*, 7(11), p. 1157. Available at: https://doi.org/10.3390/app7111157.

Habeeb, B. *et al.* (2021) 'Stochastic Carbon Dioxide Forecasting Model for Concrete Durability Applications', in *18th International Probabilistic Workshop, Lecture Notes in Civil Engineering 153*, pp. 753–765. Available at: https://doi.org/10.1007/978-3-030-73616-3_58.

Hegazy, T., Tully, S. and Marzouk, H. (1998) 'A neural network approach for predicting the structural behavior of concrete slabs', *Canadian Journal of Civil Engineering*, 25(4), pp. 668–677. Available at: https://doi.org/10.1139/cjce-25-4-668.

Ho, S.L. and Xie, M. (1998) 'The use of ARIMA models for reliability forecasting and analysis', *Computers & industrial engineering*, 35(1–2), pp. 213–216.

Huang, Y.-H. (2010) 'Artificial Neural Network Model of Bridge Deterioration', *Journal of Performance of Constructed Facilities*, 24(6), pp. 597–602. Available at: https://doi.org/10.1061/(ASCE)CF.1943-5509.0000124.

Kim, J.J., Kim, A.R. and Lee, S.W. (2020) 'Artificial neural network-based automated crack detection and analysis for the inspection of concrete structures', *Applied Sciences (Switzerland)*, 10(22), pp. 1–13. Available at: https://doi.org/10.3390/app10228105.

Landahl, H.D., McCulloch, W.S. and Pitts, W. (1943) 'A statistical consequence of the logical calculus of nervous nets', *The Bulletin of Mathematical Biophysics*, 5(4), pp. 135–137. Available at: https://doi.org/10.1007/BF02478260.

Llorens, M., Serrrano, Á. and Valcuende, M. (2019) 'Sensores para la Determinación de la Durabilidad de Construcciones de Hormigón Armado.', *Revista ingeniería de construcción*, 34(1), pp. 81–98. Available at: https://doi.org/10.4067/S0718-50732019000100081.

De Marco, A. *et al.* (2013) 'Assessing ozone and nitrogen impact on net primary productivity with a Generalised non-Linear Model', *Environmental Pollution*, 172, pp. 250–263. Available at: https://doi.org/10.1016/j.envpol.2012.08.015.

Marsh, P.S. and Frangopol, D.M. (2008) 'Reinforced concrete bridge deck reliability model incorporating temporal and spatial variations of probabilistic corrosion rate sensor data', *Reliability Engineering & System Safety*, 93(3), pp. 394–409. Available at: https://doi.org/10.1016/j.ress.2006.12.011.

O. Akande, K. *et al.* (2014) 'Performance Comparison of SVM and ANN in Predicting Compressive Strength of Concrete', *IOSR Journal of Computer Engineering*, 16(5), pp. 88–94. Available at: https://doi.org/10.9790/0661-16518894.

Pereira, E. *et al.* (2009) 'A Galvanic Sensor for Monitoring the Corrosion Condition of the Concrete Reinforcing Steel: Relationship Between the Galvanic and the Corrosion Currents', *Sensors*, 9(11), pp. 8391–8398. Available at: https://doi.org/10.3390/s91108391.

Presuel, F. and Liu, Y. (2012) 'Temperature Effect on Electrical Resistivity Measurement on Mature Saturated Concrete', in *NACE - International Corrosion Conference Series*.

Rincon, L.F. *et al.* (2022) 'Novel trends on the assessment and management of maritime infrastructures: Outcomes from GIIP project', in *Eighth World Conference on Structural Control and Monitoring (8WCSCM)*. Orlando, USA.

Shevtsov, D. *et al.* (2022) 'Progress in Sensors for Monitoring Reinforcement Corrosion in Reinforced Concrete Structures—A Review', *Sensors*, 22(9), p. 3421. Available at: https://doi.org/10.3390/s22093421.

Ukrainczyk, N., Banjad, I. and Ukrainczyk, V. (2004) 'Application of neural network in predicting damage of concrete structures caused by chlorides', in *International Symposium: Durability and Maintenance of Concrete Structures*, pp. 187–194.

Willsher, K., Tondo, L. and Henley, J. (2018) 'Bridges across Europe are in a dangerous state, warn experts'. Available at: https://www.theguardian.com/world/2018/aug/16/bridges-across-europe-are-in-a-dangerous-state-warn-experts.

Yang, Y., Yu, C. and Zhong, R.Y. (2023) 'Generalized linear model-based data analytic approach for construction equipment management', *Advanced Engineering Informatics*, 55, p. 101884. Available at: https://doi.org/10.1016/j.aei.2023.101884.