buildings

Article

Condition Prediction for Existing Educational Facilities Using Artificial Neural Networks and Regression Analysis

Ahmed M. Hassan [1,*], Kareem Adel [2], Ahmed Elhakeem [2] and Mohamed I. S. Elmasry [3]

[1] Construction and Building Engineering Department, College of Engineering and Technology, Arab Academy for Science Technology and Maritime Transport (AASTMT), Port Said 42511, Egypt
[2] Construction and Building Engineering Department, College of Engineering and Technology, Arab Academy for Science Technology and Maritime Transport (AASTMT), Cairo 11511, Egypt
[3] Construction and Building Engineering Department, College of Engineering and Technology, Arab Academy for Science Technology and Maritime Transport (AASTMT), Alexandria 21523, Egypt
* Correspondence: a.hassan@aast.edu; Tel.: +20-120-079-6900

Abstract: Infrastructural assets such as roads, bridges, and buildings make a considerable contribution to national economies. These assets deteriorate due to aging, environmental conditions, and other external factors. Maintaining the performance of an asset in line with rational repair strategies represents a considerable challenge for decision-makers, who may not pay attention to developing adequate maintenance plans or leave the assets unmaintained. Worldwide, organizations are under pressure to ensure the sustainability of their assets. Such organizations may burden their treasury with random maintenance operations, especially with a limited budget. This research aims to develop a generalized condition assessment approach to monitor and evaluate existing facility elements. The proposed approach represents a methodology to determine the element condition index (CI). The methodology is reinforced with an artificial neural network (ANN) model to predict the element deterioration. The performance of this model was evaluated by comparing the obtained predicted CIs with ordinary least squares (OLS) regression model results to choose the most accurate prediction technique. A case study was applied to a group of wooden doors. The ANN model showed reliable results with R^2 values of 0.99, 0.98, and 0.99 for training, cross-validation, and testing sets, respectively. In contrast, the OLS model R^2 value was 1.00. These results show the high prediction capability of both models with an advantage to the OLS model. Applying this approach to different elements can help decision-makers develop a preventive maintenance schedule and provide the necessary funds.

Keywords: condition prediction; condition assessment; artificial neural networks; asset management; multiple regression analysis

check for updates

Citation: Hassan, A.M.; Adel, K.; Elhakeem, A.; Elmasry, M.I.S. Condition Prediction for Existing Educational Facilities Using Artificial Neural Networks and Regression Analysis. *Buildings* **2022**, *12*, 1520. https://doi.org/10.3390/buildings12101520

Academic Editor: Ahmed Senouci

Received: 18 August 2022
Accepted: 20 September 2022
Published: 23 September 2022

1. Introduction

Assets such as buildings, roads, and bridges represent the infrastructure of countries, which contributes considerably to the national economy. In 2001, the USA federal government owned more than 500,000 buildings [1]. These public assets were valued at more than USD 328 billion. According to the U.S. Census Bureau, the annual value of construction that is put in place for public assets in the United States was between USD 346 and USD 361 billion in 2020 and 2021, respectively [2]. These assets must be preserved and developed to ensure a country's ongoing development. However, they cannot be completely protected from potential deterioration due to aging, the nature of their usage, climatic effects, or geological conditions [3]. For these reasons and when building maintenance is neglected, maintenance work puts a huge burden on a country's budget, especially in developing countries. For instance, in 2017, the Central Agency for Public Mobilization and Statistics in Egypt announced that Egypt possessed almost 14.3 million buildings [4], both residential and non-residential. More than 3.3 million of these assets needed minor or major repair, or complete demolition [5]. This represents approximately 23% of the total assets. A large budget should be established for maintenance operations to ensure these assets function efficiently. According to the last census in 2007, the U.S. Census Bureau reported that the expenditure for maintenance and repairs of residential properties was over USD 54 billion [6]; this represents the consumption of treasury funds if no plans are made to bear these costs in advance.

In terms of decision-making, experts suggest developing an asset management program. This program includes policies, objectives, and strategies before embarking on a specific asset management plan. In this regard, asset management information systems (AMIS) are a crucial topic. An AMIS includes the asset's technical, financial, and historical information. These systems can vary in sophistication from simple spreadsheets that are widely used to advanced systems that use specific models to predict future conditions [7]. Such models provide the most robust analysis and prediction capabilities [8]. However, facility managers confront barriers that impede the implementation of these prediction models. Insufficient information about current asset conditions causes failure in developing these prediction models [9] and the subsequent inability to adopt a generalized approach for assessing building sub-elements [10]. Related studies have proposed approaches to assess infrastructure facilities components and prioritize them for maintenance [9]. However, such efforts neglect how this assessment could be rolled up to the building level and provide more than just a simple view of the building. This highlights the need for a generalized and simplified approach to element condition assessment that can be rolled up to the building level.

Regarding element condition prediction, the past two decades have seen a tremendous growth of interest in machine learning's contribution to asset management [11,12]. The ordinary least squares (OLS) technique has been used as a regressor to solve problems related to prediction as it is constructed on a solid theoretical basis. It is thus used for estimation or as a validation indicator for other estimation techniques [13]. On the other hand, artificial neural networks (ANNs) are among the most widely used supervised learning techniques [14]. They are used as a tool for optimization and prediction in the domain of asset condition assessment. Such a technique saves development time as the ANN can self-learn and identify nonlinear relationships between input and output parameters [15].

Despite many related studies, few have proposed the application of neural networks in the deterioration prediction of buildings [9]. Most researchers have discussed their application to other assets such as bridges, roads [16], and sewer pipes for damage detection [17]. Instead, Markov models have been used to predict building conditions [18,19]. Other studies have used neural networks to automate the visual inspection process and defect classification. For instance, neural networks were used to classify masonry wall cracks [20] or cracks in concrete structures using images [21]. Regarding prediction models, only time-related input parameters were considered as an input for prediction models, thus neglecting other proven indoor and outdoor parameters affecting the service life (e.g., the element's location [22], the nature of the space usage, or element usage rate). These parameters have a significant effect on condition degradation according to ISO 15686–1:2000 [23]. Additionally, very few studies have conducted a comprehensive comparison of ANN with other regression techniques such as OLS.

This study contributes to the field of asset management by developing a generalized approach for predicting an element condition index (CI) since there have been few previous contributions in this domain. The approach can be rolled up to anticipate the building's condition. An ANN and OLS models were used to predict the condition of a specific element. Then, a comprehensive comparison was conducted to determine the most reliable model. The study's objective was to assist decision-makers in developing short-term/long-term proactive maintenance plans and optimally fund the related efforts.

The remainder of this paper is structured as follows: Section 2 presents the theoretical background of asset management concepts and processes, the definition of an ANN, components, inner operations, and contributions in the civil domain as well as an OLS definition and its significance. Section 3 describes the research methodology. Section 4 provides the statistical indicators used to evaluate the ANN and OLS techniques. Section 5 provides detailed results, while Section 6 provides comprehensive conclusions, research limitations, and future directions.

2. Theoretical Background

2.1. Asset Management System (AMS)

The Federal Highway Administration defined asset management as cost-effectively maintaining, upgrading, and operating assets [24]. This covers technical and mathematical analysis along with good business practice and economic theory. It is goal-oriented and includes data aggregation, strategy evaluation, program selection, and feedback components. AMS follows five main processes: (1) condition evaluation; (2) deterioration modeling; (3) repair alternatives and strategies; (4) extent of improvement after repair; and (5) asset prioritization and repair fund allocation [7,25,26]. These processes regularly update the asset database for any upcoming repairs. A database comprising all relevant information must be available to assess the asset condition correctly and in a timely fashion. This information includes, for example, asset type, quantity, age, location, asset hierarchy, design data, and construction data, in addition to maintenance records [8,27].

2.1.1. Condition Evaluation

Condition evaluation is defined as the use of a systematic method designed to produce proportionate, pertinent, and useful information to conduct a technical evaluation of the physical state of an asset [10]. This database is essential for better judgment and the most effective repair choice. It helps decision-makers to develop proactive capital planning and schedule future maintenance interventions [28]. Inspection and data collection are crucial in this process. The inspectors gather all possible information, which helps to evaluate the performance and make the best decisions in different processes. Non-destructive evaluation methods are used when fast and accurate data is needed. The most common tools used for inspection are visual inspection and analyzing images. However, visual inspection has problems related to time and money consumption. Previous efforts have tried to find alternatives by applying neural networks and smart sensors to provide real-time data collection, which is required for analysis [7,29,30].

2.1.2. Deterioration Modeling

Making decisions related to infrastructure maintenance and rehabilitation depends not only on the current measured condition of the assets but also on their expected possible degradation over time. Deterioration models are substantial for any asset management system, as they are used in the future condition forecasting of an asset or its components [7]. There are multitudinous deterioration prediction techniques such as (1) deterministic models, (2) stochastic Markovian models [29], and (3) artificial intelligence techniques such as artificial neural networks, fuzzy logic systems, and genetic algorithms [30]. The condition indices' values are deterioration indicators of either the asset components or the whole asset. For building components, the condition index scale is commonly 0% to 100% [9]. Such a scale comprises from 1 to 4, 5, 6, or 7 conditions according to asset type, with 0% representing a critical (failure) condition and 100% describing a new condition. This numeric scale can be translated into a linguistic representation. Some rating scales and the corresponding linguistic expressions are illustrated in Table 1. Subsequently, when the condition state of an element or structure meets a predefined threshold, the organization activates the maintenance process.

Table 1. Reference of different condition scales.

Year	Ref.	Asset Type	Condition Scale	Linguistic Representation
1997	[31]	Buildings	1–4	Deterioration: (1 = no; 2 = slight; 3 = moderate; and 4 = severe).
2005	[32]	Buildings	0–100	Deterioration: (0–20) = no; (20–40) = slight; (40–60) = moderate; (60–80) = sever; and (80–100) = critical.
1998	[33]	Any Asset	1–7	Condition category: (1 = Failed; 2 = V. Poor; 3 = Poor; 4 = Fair; 5 = Good; 6 = V. Good; and 7= Excellent).
2021	[9]	Buildings	0–100	Condition category: (0–40) = full deterioration; (40–60) = poor quality; (60–75) = Imperial quality; (75–85) = good quality; (85–92) = accepted quality; (92–99) = fine quality; and (99–100) = exemplary quality.
2021	[27]	Buildings	1–6	Condition category: (1 = Very good condition, 2 = Good condition, 3 = Reasonable condition, 4 = Borderline condition, 5 = Bad condition, 6 = Very bad condition)

2.1.3. Repair Alternatives and Strategies

For a decision-making approach, life cycle cost analysis evaluates the total costs accrued over the infrastructure's life, from construction to final replacement or demolition. It is an effective approach for an existing facility to compare the long-term effects of different maintenance strategies and determine the best one. Four types of maintenance strategies have been defined by [34] as follows:

- Corrective maintenance is the simplest strategy. In this strategy, the component is kept operating until failure. So, it is not the strategy that leads to the lowest total cost.
- Time-based maintenance, or preventive maintenance, is the most widely used today. It is most effective in preventing major failures or damage. It is usually appropriate for cases where abrasive, erosive, or corrosive wear occurs and/or material properties are changed due to fatigue. Timely inspections are essential. Such measures are a must so that there is an opportunity to avoid excessive extra costs.
- Condition-based maintenance. This strategy requires additional information about the current component status to determine the device's status. A specific metric describes this current state, and in condition-based maintenance, maintenance activity is triggered when an estimated state reaches a certain threshold. This procedure provides high availability at a reasonable maintenance cost.
- Finally, reliability-centered maintenance not only considers the condition of the system components but also the impact on the system's performance. This strategy is regarded as the most accurate.

Facility managers should always be aware of checking the conditions of buildings to avert any potential damage, which can be costly. Later, after the repair deterioration process has started, some researchers assume that the deterioration trend parallels the pre-correction deterioration trend [7].

2.2. Artificial Neural Networks (ANNs)

ANNs were originally inspired by the human biological nervous system [35]. A neural network is a network of interconnected basic processing elements, units, or nodes whose functionality is based on the human neuron [36]. These neurons are connected by synapses that can transmit signals to each other [37]. It is a vast technological domain where one can implement "human brain decision-making power" into computer programs based on error and approximation [38]. The network's processability depends on the connection weights generated through learning from training patterns [36].

2.2.1. ANN Components

One of the most used neural networks is the multilayer perceptron [39]. It is generally composed of an input layer, an output layer, and a hidden layer(s) that are interconnected successively. Such feed-forward networks utilize backpropagation as a learning algorithm.

The input layer comprises the independent variables that present the basis for a final decision or prediction. It is simply the layer that communicates with the external environment. The hidden layer(s) comprises a group of neurons with an activation function and establishes a link between the input and output layers. The hidden layer's target is to gain and recall vital characteristics and sub-features from input patterns to anticipate the network's result [40]. The output layer contains the dependent variables, or the network results, which are determined based on the values of the input layer nodes. Once the neural network is trained, it can predict the output for unseen input data. Numerous previous studies were used to calculate the number of neurons in the hidden layer. Some have confirmed there is no specific way to determine their number [41]. Others have confirmed that this could be achieved through Equation (1), where n describes the optimum number of neurons in the hidden layer, Ip is the input layer size, and Op is the output layer size [9]. Due to the lack of a general technique, most researchers usually determine the adequate number of processing elements for the hidden layers by conducting trials.

$$n = \frac{2 \times (\mathrm{Ip} + \mathrm{Op})}{3} \tag{1}$$

2.2.2. Inner Operation of ANNs

An artificial neuron consists of inputs, a summation block, an activation block, and a single output, as shown in Figure 1. Neurons from different layers are connected, and each connection has a weight, which changes with the learning process until it reaches the expected target values. The role of the activation function is to introduce the nonlinearity properties into the network. A complex nonlinear model can map the nonlinear relationships between inputs and outputs. The most common activation functions in hidden layers are rectified linear activation [42], logistic (sigmoid), and hyperbolic tangent [43]. Depending on the used function, it usually ranges between (0 to 1) or (−1 to 1). The input of the neurons connected to the next layer will be the output of the activation function. The role of the bias in the trigger function is to provide flexibility for the trigger function to change. It allows for shifting the activation function by adding a constant to the input.

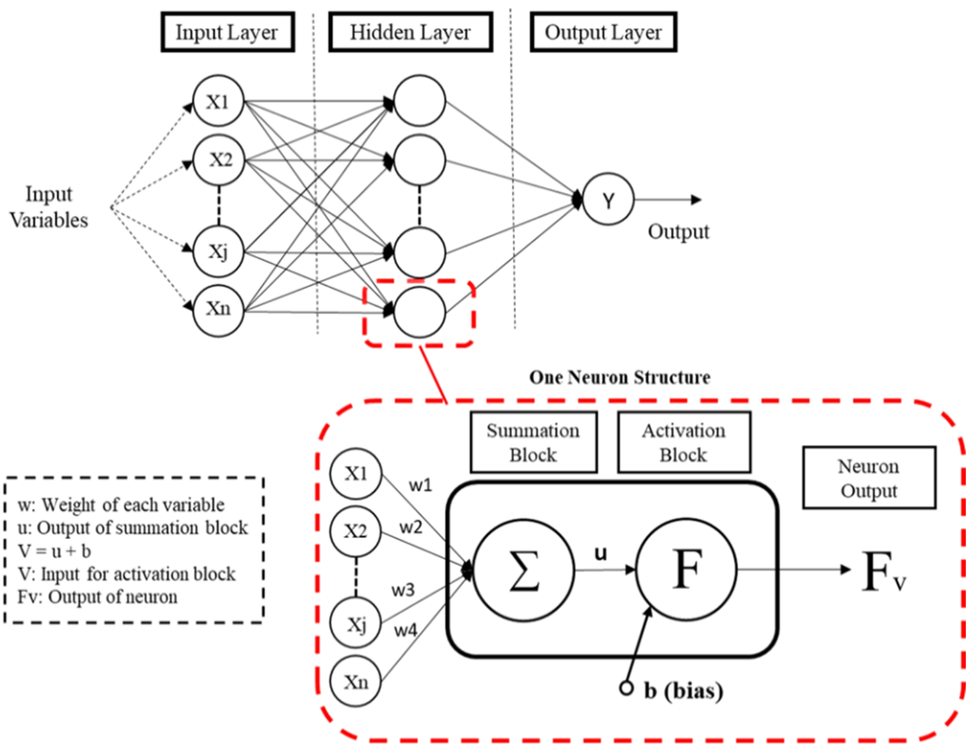


Figure 1. Summary of the artificial neural network structure.

2.2.3. ANNs Application in Construction Domains

ANNs have contributed to much research in civil engineering [44]. In structural engineering, ANNs are applied in pattern recognition and machine learning for structural analysis and design, design automation and optimization, structural system identification, condition assessment, monitoring, and control. In construction engineering, ANNs are applied in construction scheduling and management, construction cost estimation, resource allocation and scheduling, and construction litigation. They are also applied to other engineering fields such as environmental and water resources, traffic, highways, and geotechnical engineering. For instance, a generic object detector trained to identify and classify road damage from still images or real-time video was presented in [45]. Other research has introduced automatic crack detection in concrete structures from images [21]. A bridge management system was developed to evaluate and predict bridge deck deterioration conditions using ANNs [46]. A model capable of estimating the construction cost for road projects was produced using ANNs [41]. Recent efforts have utilized ANNs for estimating road construction costs at different stages using databases from past projects [39]. Additionally, convolutional neural networks were used for detecting key building defects such as mold, deterioration, and stains from images [47] and ANNs were utilized to predict building energy performance [48]. To sum up, prior studies have shown that the ANN has great potential and robust applications in prediction, optimization, classification, decision-making, and specifically, in asset condition assessment.

2.3. Ordinary Least Squares (OLS) Technique

The ordinary least squares technique is one of the simplest and most used regression methods. OLS is a linear regression technique that has been used in much previous research for estimating purposes such as cost estimating [13], housing price forecasting [49], and building condition assessment [12]. Despite its benefits, it is inadvisable to use OLS in cases involving nonlinear relations between the inputs and outputs [13]. Mean squared error (MSE), root mean squared error (RMSE), adjusted root mean squared error, coefficient of multiple determination, and R-squared are used to assess its predictive performance [12,49]. Generally, in the case of describing multiple independent variables that affect one dependent variable, multiple linear regression analysis is used. This can be represented by Equation (2).

$$Y = C + b_1X_1 + b_2X_2 + \ldots + b_nX_n \quad (2)$$

where Y is the total estimated CI, X_1; $X_2, \ldots, X_n$ are measures of independent variables that may help in estimating Y, C is the estimated constant, and b_1; $b_2 \ldots, b_n$ are the coefficients estimated by regression analysis.

3. Methodology

This research used a simplified approach to determine a selected element CI as follow. First, decompose the educational facility into multiple hierarchies such as floors, systems, and zones. Second, create a coding system and develop a database of the selected element using a Microsoft Excel sheet. Third, identify selected element components and possible defects in each component, create a component defect/impact matrix to get the proportional weight of each component in the element, and the data collection process is carried out through visual inspection of the elements of the studied facility. Fourth, data acquisition and analysis are performed to identify the input and output parameters of the neural network and OLS models. Fifth, train and test the neural network model to validate the model's workability and compare the results to OLS output using statistical indicators. Finally, choose the best model to operate. The methodology is illustrated in Figure 2.

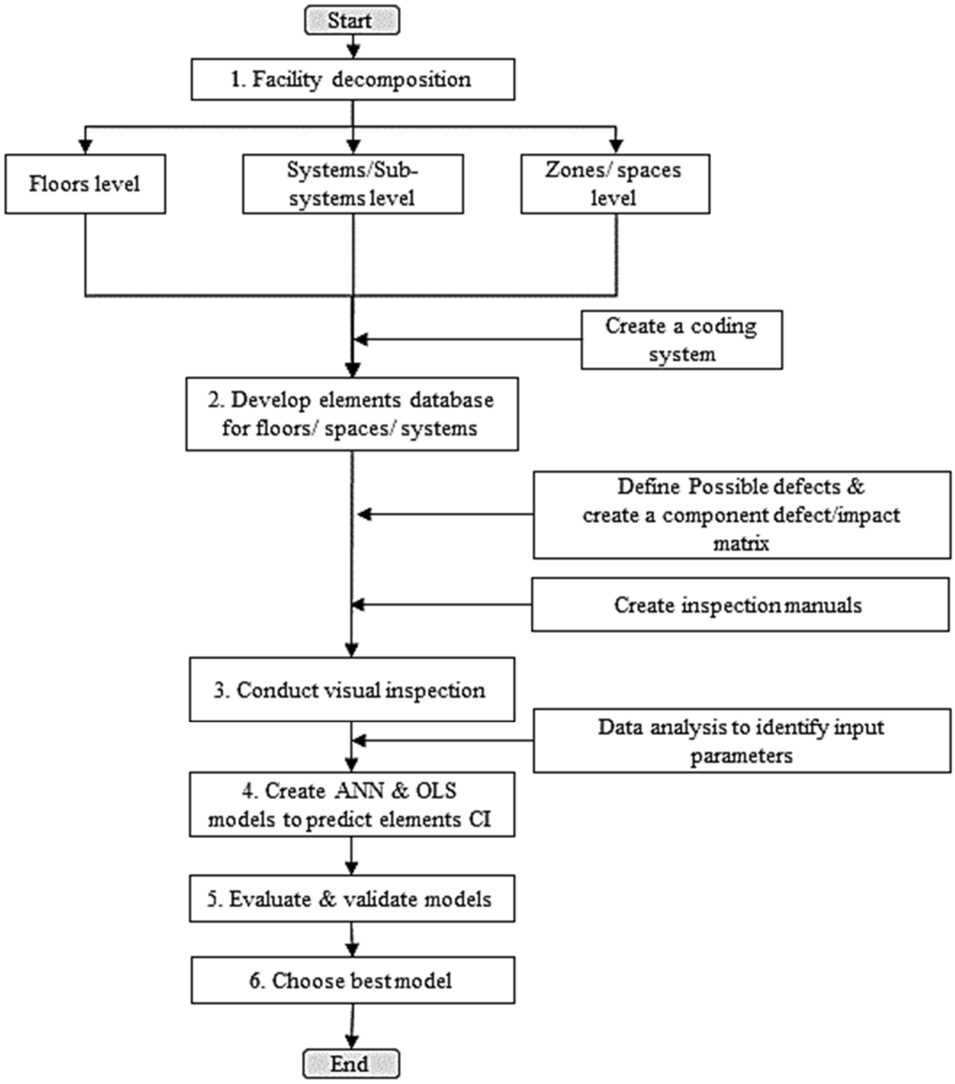


Figure 2. Illustration of the methodology.

3.1. Decomposition of Educational Facility

An educational facility (AASTMT Portsaid branch) was chosen for this study and divided according to the breakdown depicted in Figure 3. This breakdown includes all the existing facility buildings for developing an approach to determine the element condition index (CI), The data was collected from the college and some residential buildings to cover all spectrums and variations that could affect the input parameters.

3.2. Element CI Calculation

The wooden doors were taken as the element for the study; thus, doors were divided upon inspection into four main components that must be checked for defects during visual inspection. These four components included the condition of the door body, the door frame, the hinges and metal fittings, and the paint. Figure 4 represents a part of the guidelines for the items to be checked by the inspector during the inspection process. In this approach, Equation (3) was used to calculate the average condition index of the element [7].

$$\mathrm{CI\,j} = 100 - \frac{\sum_{i=1}^{d} \mathrm{Wi} \times \mathrm{Si}}{100} \quad (3)$$

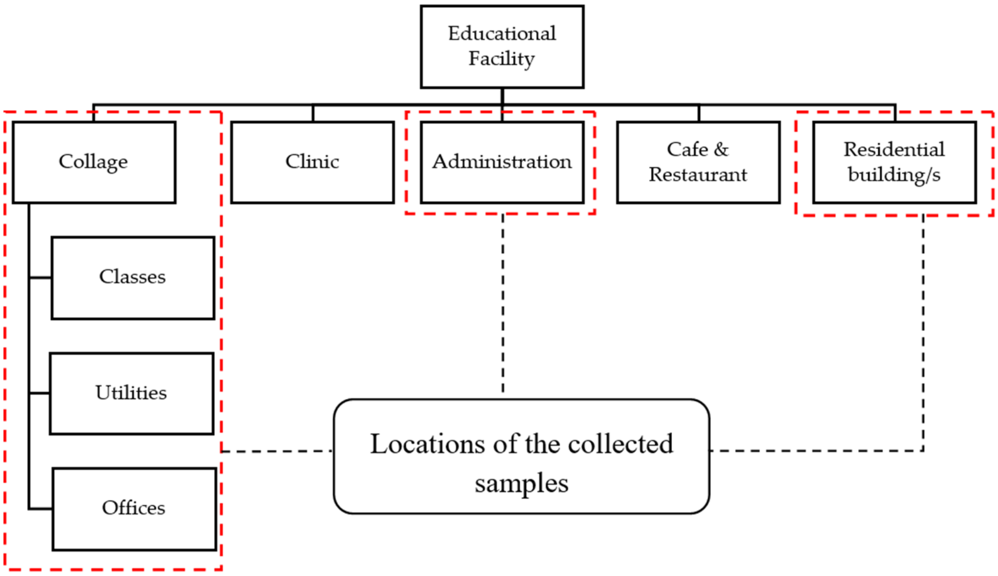
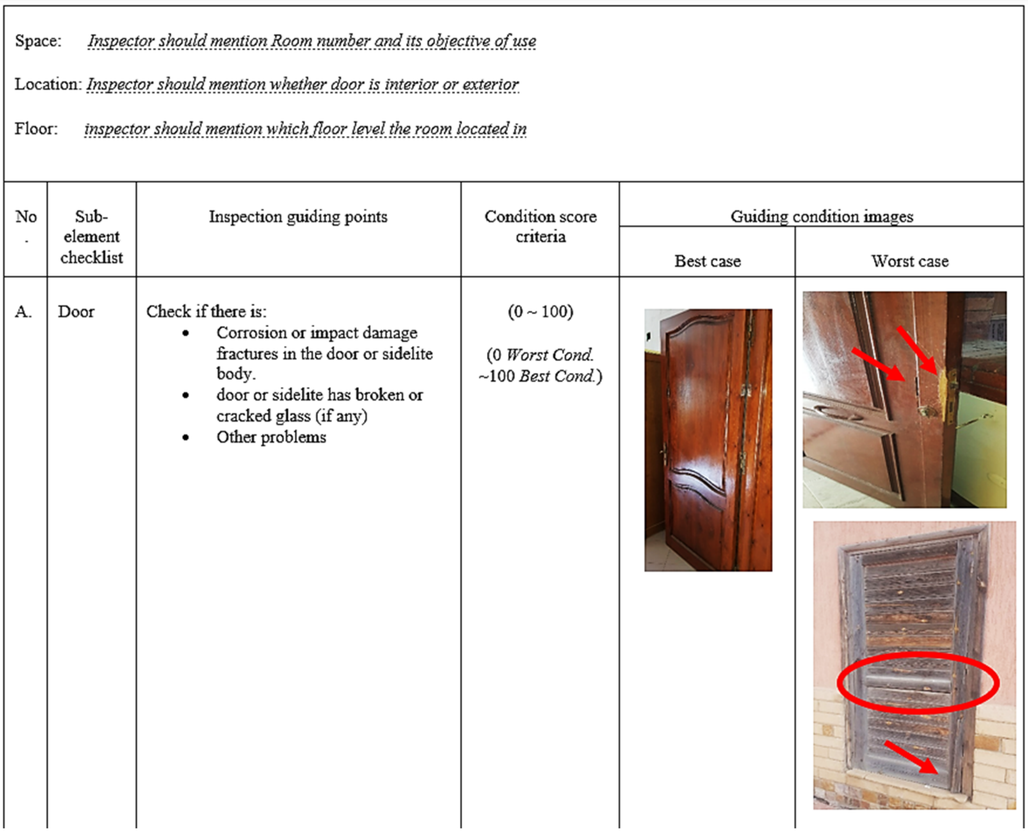


Figure 3. AASTMT Portsaid branch breakdown.



Space: *Inspector should mention Room number and its objective of use*

Location: *Inspector should mention whether door is interior or exterior*

Floor: *inspector should mention which floor level the room located in*

No.	Sub-element checklist	Inspection guiding points	Condition score criteria	Guiding condition images	
				Best case	Worst case
A.	Door	Check if there is: • Corrosion or impact damage fractures in the door or sidelite body. • door or sidelite has broken or cracked glass (if any) • Other problems	(0 ~ 100) (0 *Worst Cond.* ~100 *Best Cond.*)		

Figure 4. Part of the inspection guide.

CI j stands for the condition index for the jth (component or part), where Wi = deficiency weight (i) (from 0 to 1); Si = severity extent for deficiency (i) (from 0 to 100); i = counter for potentially deficient components (j). The CI is a value that ranges from 0 to 100, where from (0 to 10) represents critical condition, from (11 to 24) represents poor condition, from (25 to 49) represents fair condition, and from (50 to 100) represents a good condition.

3.3. Parameters Defect/Impact Matrix

The previous elements were tabulated and given an impact factor according to the corresponding impact area to conclude each deficiency weight, as shown in Table 2, where; (H) high impact (0.71 to 1), (M) moderate impact (0.31 to 0.7), and (L) low impact (0 to 0.3).

Table 2. Defects/impact matrix.

Impact Area \ Defect Type	Corrosion or Fractures in the Door Body	Corrosion or Fractures in the Door Frame	Loss or Malfunction of the Door Hinges or Metal Fittings	Flaking or Cracked Paint
Operational objective of space	H	M	H	L
Safety	H	H	H	L
Architectural objective	H	L	L	H
Weight of deficiency	0.35	0.21	0.26	0.18

3.4. Identification of Input and Output Parameters

To identify the input parameters of the network, factors from previous research [9], in parallel with data gathered during the visual inspection revealed the main input parameters to be used in the network modeling and training. Nine input parameters were selected for the input layer from the analysis of experimental data to evaluate the output. The main deterioration factor is the door's age, which thus represents the main factor. The rest of the factors can be classified as the physical condition of its main components and the accessibility factors. For example, the physical factors of the door components are the body, frame, hinges, and painting conditions. Moreover, the accessibility factors are the location in the space, the floor, and the nature of the space (educational or administrative space). The input and output parameters with descriptions and ranges are shown in Table 3.

Table 3. Neural network input parameters.

Input Parameter	Description	Range
X_1	Door Body Condition	(0 Worst Cond.~100 Best Cond.)
X_2	Door Frame	(0 Worst Cond.~100 Best Cond.)
X_3	Hinges or Metal Fittings	(0 Worst Cond.~100 Best Cond.)
X_4	Painting Condition	(0 Worst Cond.~100 Best Cond.)
X_5	Floor	(0, 1, 2 . . . . N)
X_6	Door's age	(0, 1, 2 . . . . N)
X_7	Door's Usage rate	Low, Moderate, High
X_8	Location of the door in space	Interior, Exterior
X_9	Nature of space use	Educational, Utility, Administrative, Residential
Output Parameter	**Description**	**Range**
Y	Condition Index (CI)	(0 Worst Cond. ~100 Best Cond.)

3.5. Neural Network Topology

The design of the neural network architecture is part of the modeling step. It is a complicated and dynamic process requiring the establishment of the internal structure and rules (i.e., the number of hidden layers and neurons and the type of activation function). The model is created based on the type of data and the response required by the application as shown in Figure 5.

The model was designed to include three successive layers to identify the model's topology: an input layer of nine neurons corresponding to the nine input parameters; an output layer of one neuron as the target output; and one hidden layer of several hidden neurons (NHs) that are determined after multiple trials during the testing phase.

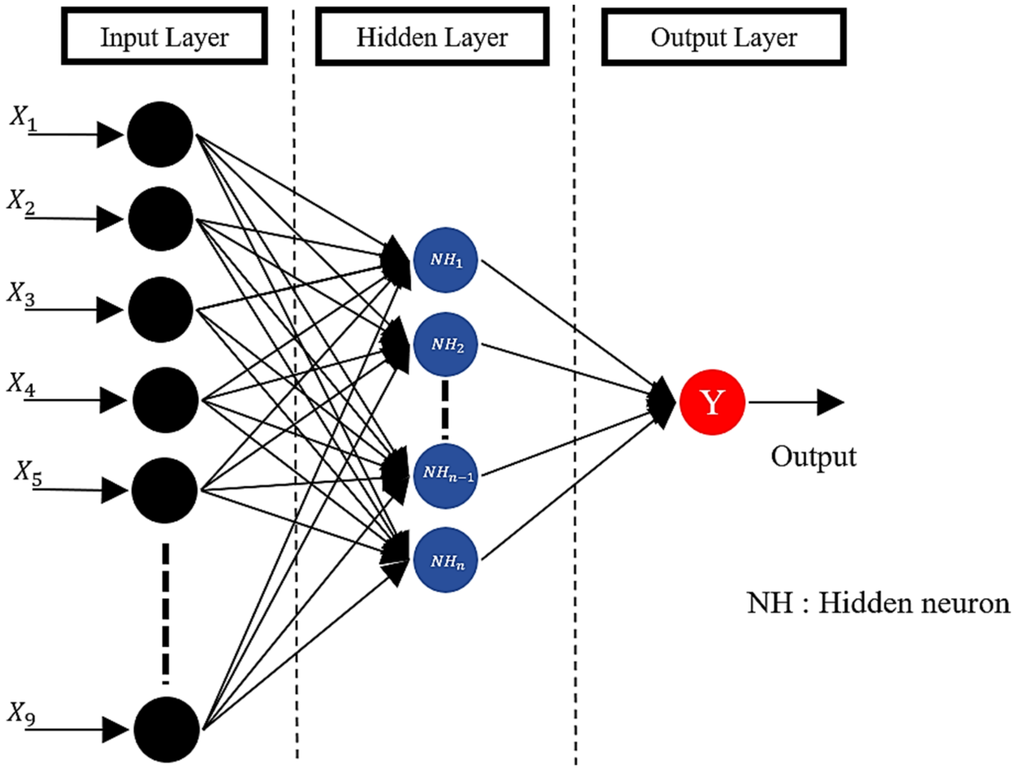


Figure 5. The architecture of the neural network model.

The model was created with NeuroSolutions Software Version 6.0 as a multilayer perceptron artificial neural network. In this study, Equation (1) was applied to calculate the approximate number of NHs. The number of NHs in the hidden layer was almost 7. To ensure the most reliable number, and due to the lack of a general technique as previously described, an error factor of at least ±20% was cited. So, a range of trial models was set between 5 and 15 NHs. The model with the most accurate results was the one corresponding to 5 NHs. The model's back-propagation rule for supervised learning was the Levenberg–Marquardt algorithm, and the nonlinear activation function utilized in the model was the hyperbolic tangent function [41]. The available database was divided into three sets (Training (108 doors), Cross-Validation (13 doors), and Validation (13 doors)). The size of each set does not rely on standard or generalized rules. However, each set should cover all the input parameters' spectrums and variations.

3.6. Processing Phase

The processing of the model included both a training set and a cross-validation set with various features. Training sets were used to learn and record the correlations between the inputs and outputs. On the other hand, the cross-validation set was used to monitor model performance, ensure an optimal level of generalization, and avoid overtraining issues without affecting network weight updates. The validation set was not used during model processing and was reserved for measuring the validity and preparedness of the processed model and handling new cases.

STATA, a commercially available statistical software for data science, was used to develop the OLS model. Such a model determines the same input/output relationship stated in Table 4, and helps evaluate the results obtained by the ANN.

3.7. Simplified Rolling-Up Approach

The building was divided into three main hierarchies: zones that contain multiple spaces that share similar characteristics, physical systems and subsystems (civil systems, architectural systems, etc.), and floors. Accordingly, analysis was done for each element separately based on the inspectors' reports to calculate the condition index (CI) of each element in the zone. Then, elements were sorted according to their corresponding systems to obtain the average CI for each system using a proper proportional weight matrix. To assess the building condition, each subsystem can be weighted according to a set of aspects by following the same methodology used in the defect/impact matrix. Thus, through the desired sorting, the average CI can be obtained per floor, system, zone, or space. Subsequently, the CI of the entire building could be rolled up and calculated. By obtaining the same data for other buildings and through using the proper forecasting input parameters, it is possible to predict the entire building condition. The aforementioned different hierarchies have a key advantage with respect to prioritizing maintenance operations. For example, assuming the classes and studios zone is the most important; then, by using the proper criteria, this zone will have the higher weight. Thus, it comes first in the maintenance schedule. The approach illustrated in Figure 6 was adopted from [50].

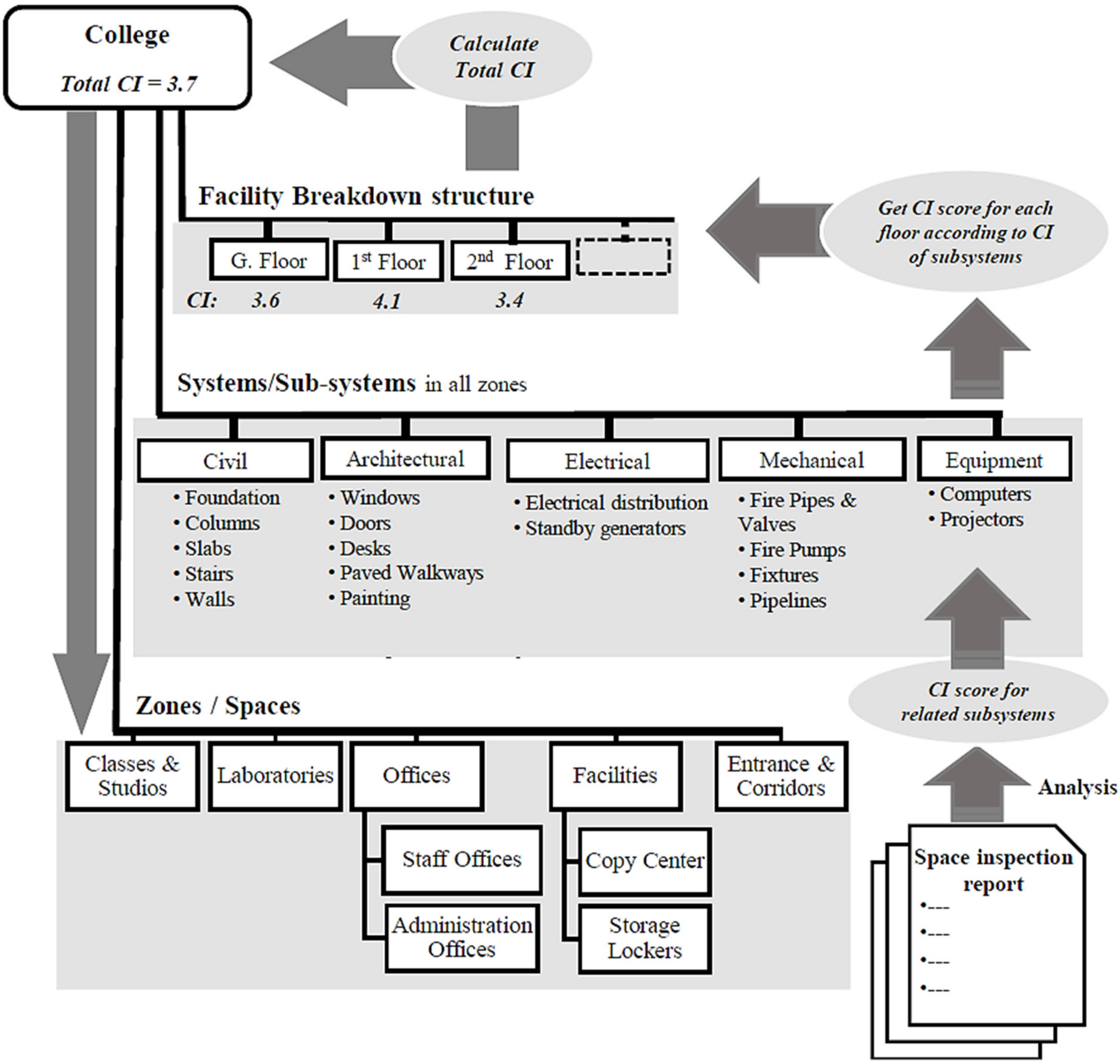


Figure 6. CI rolling up approach.

4. Model Evaluation

The performance of the developed ANN and OLS models was appraised for the interior and exterior doors. Through the evaluation process, the deviation between the output and target data values is measured. The term "output data" refers to the data that the models estimate. The term "target data" represents the actual data obtained from the visual inspection during the inspection process. The statistical indicators that were used to validate the models are the Coefficient of Determination (R^2) (4), Mean Squared Error (MSE) (5), Root Mean Squared Error (RMSE) (6), Mean Absolute Error (MAE) (7), and Mean Absolute Error Percentage (MAEP) (8). These statistical indicators are used to anticipate future performance or test assumptions relying upon other data and attain optimal performance.

$$R^2 = \frac{\sum_{i=1}^{n}\left(\hat{y}i - \tilde{y}i\right)^2}{\sum_{i=1}^{n}\left(yi - \tilde{y}i\right)^2} \tag{4}$$

$$MSE = \frac{\sum_{i=1}^{n}(yi - A.i)^2}{n} \tag{5}$$

$$RMSE = \sqrt{\frac{\sum_{i=1}^{n}(yi - A.i)^2}{n}} \tag{6}$$

$$MAE = \frac{\sum_{i=1}^{n}\frac{|yi - A.i|}{yi}}{n} \tag{7}$$

$$MAEP = \frac{\sum_{i=1}^{n}\frac{|yi - A.i|}{yi}}{n} \times 100 \tag{8}$$

where yi represents the targets, $\hat{y}i$ represents the network outputs, $\tilde{y}i$ depicts the mean of target values, A.i depicts the model predicted values, and n is the sample number.

5. Results and Discussion

5.1. ANN Model Predictive Performance

The artificial neural network (ANN) model results illustrated in Table 4 demonstrate its reliable performance. The MSE, which measures the average squared difference between the output data and target data of the training, cross-validation, and validation sets, was 1.58, 2.63, and 1.59, respectively. The RMSE of the training, cross-validation, and validation sets was 1.26, 1.62, and 1.26, respectively, which showed the ability of the network to fit the data. The MAEP indicates that the percent deviation of the output data from the target data was only about 1.24, 1.56, and 1.24% for the training, cross-validation, and validation sets. The R^2 for the training, cross-validation, and testing sets was 0.99, 0.98, and 0.99, respectively. The value of R^2 represents the proportion of variance for a dependent variable that is explained by the independent variable/s. These values indicate that nearly 99% of the variance in the doors' CIs can be predicted by the model. The 1% error between the output and target data is acceptable compared to the literature [9].

Table 4. Summary ANN model results for the doors.

Set	MSE	RMSE	MAE	MAEP (%)	R^2
Training	1.58	1.26	0.01268	1.27	0.9953
Cross-Validation	2.63	1.62	0.01562	1.56	0.9899
Testing	1.59	1.26	0.01238	1.24	0.99

- Figure 7 shows the validation loss versus the number of epochs over the training and cross-validation sets. The gradual decrease in loss indicates the efficiency of the network for learning useful representations for the inputs and the desired output.

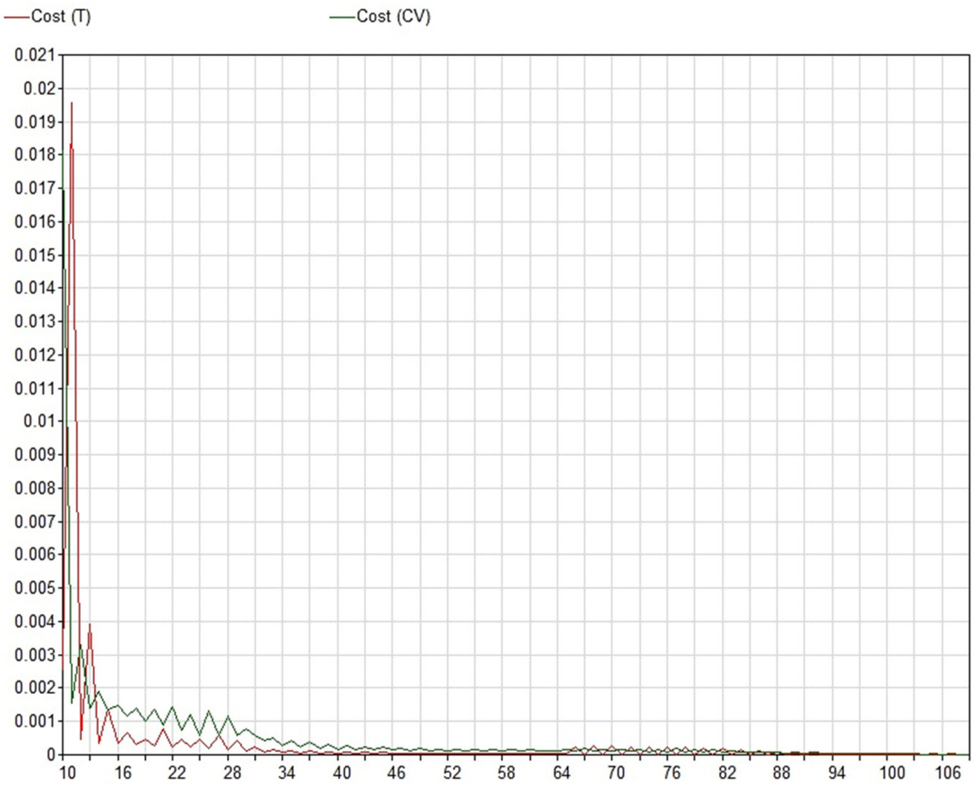


Figure 7. Training and cross-validation loss vs. epochs.

- Figure 8 shows the comparison between the target data and the output data. The percentage error distribution of all samples is shown in Figure 9. Figure 10 shows the coefficient of determination plots for each set. Figure 10a represents the training data set, Figure 10b represents the cross-validation data set, Figure 10c represents the testing data set, and finally, Figure 10d represents the coefficient of determination plot for all samples.

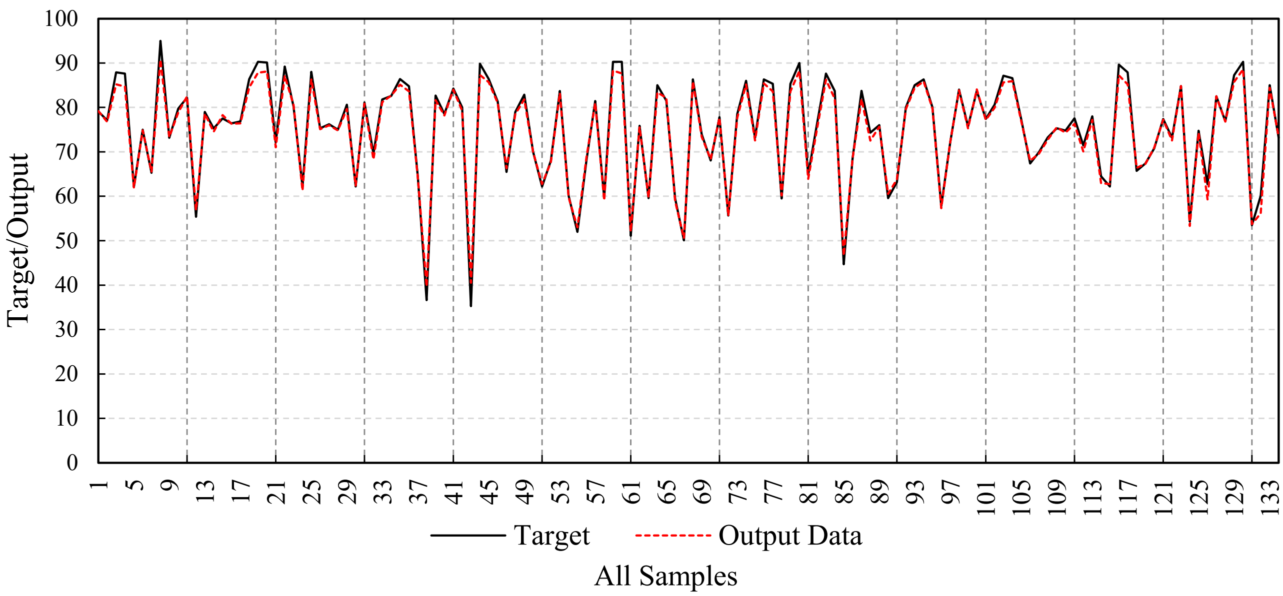


Figure 8. Target and output data for all samples.

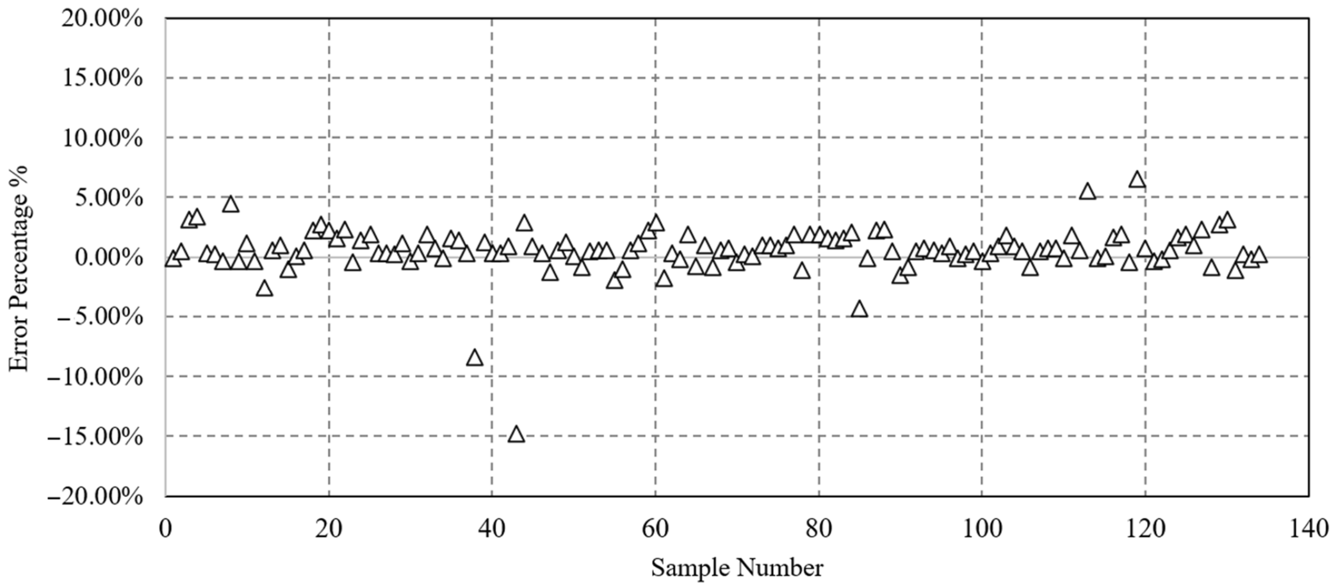


Figure 9. Percentage error distribution of all samples.

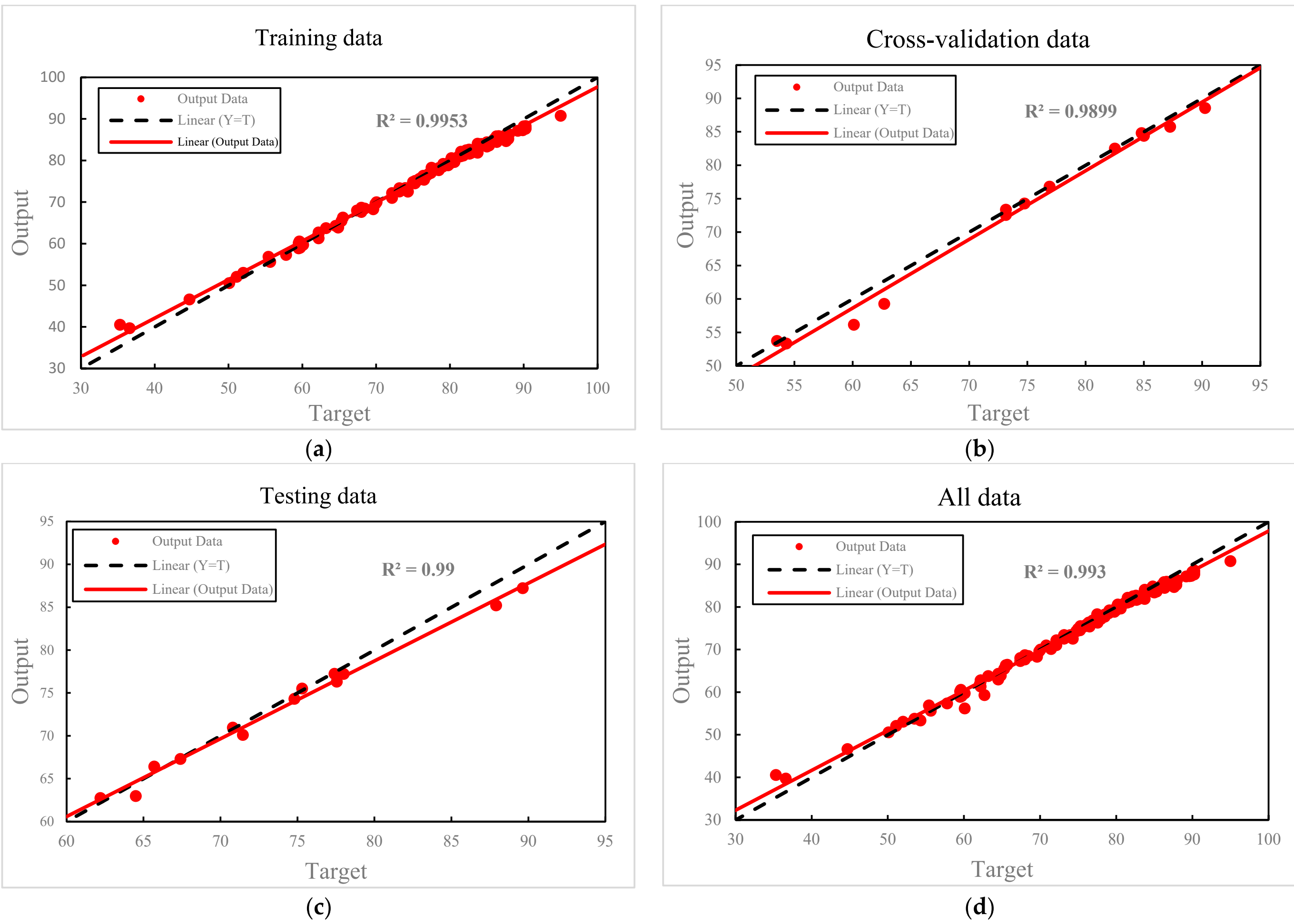


Figure 10. Coefficient of determination plots of ANN for various datasets. (**a**) R^2 value for training dataset; (**b**) R^2 value for cross-validation dataset; (**c**) R^2 value for testing dataset; (**d**) R^2 value for all datasets.

5.2. ANN and OLS Performance Comparison

- Table 5 compares the measurements obtained from both the ANN and OLS techniques in terms of R^2 and RMSE. The values of R^2 and RMSE for ANN are 0.99, and 1.26, respectively, while those for the OLS predictor are 1.00 and 1.08×10^{-6}, respectively. The R^2 of the ANN is lower than that of the OLS, which shows that the OLS model performed better. Table 6 presents a comparison between the two techniques in terms of the MAEP of all samples. The analysis revealed that 95.5% of samples have an MAEP of less than 4% using ANN with no more than 0.5% using OLS.
- According to the obtained results from both models, the most significant input parameters affecting the condition of doors are body condition, frame condition, hinges or metal fittings, painting condition, usage rate, and age, while location, nature of space usage, and floor are statistically insignificant parameters.

Table 5. Comparison of measurements between ANN and OLS.

Statistical Indicator	ANN	OLS
R^2	0.9930	1.00
RMSE	1.26	1.08×10^{-6}

Table 6. Summary of all results.

Error Percentage	ANN		OLS	
	Fre.	**Cum.**	**Fre.**	**Cum.**
0–0.5	43	43	134	134
0.5–1	34	77		
1–2	35	112		
2–3	13	125		
3–4	3	128		
<4	6	134		
MAEP	1.29%		1.1×10^{-6} %	

6. Conclusions

This research is a step forward in the field of asset management, specifically in predicting the future state of the different elements inside buildings, through which it is possible to visualize the building's overall condition. The research aims to develop a generalized approach that identifies a condition index (CI). Such an index could be obtained at the element level and rolled up to the building level to determine the building's overall condition. A case study was applied to a group of interior and exterior wooden doors of an educational facility. The samples were taken from different locations to validate the workability of the approach and help cover the prediction models' input parameters' spectrums and variations. The approach consists of a few steps: (1) decompose the educational facility into multiple hierarchies, such as floors, systems, zones, or spaces; (2) create a coding system and develop a database of doors; (3) determine the wooden door components; (4) create a component defect/impact matrix to evaluate door conditions accurately; (5) conduct a visual inspection of doors; and (6) identify input and output parameters to develop the artificial neural network (ANN) and the ordinary least squares (OLS) models. A back-propagation three-layered ANN was used. The ANN was trained, tested, and compared to the OLS model to validate the model's forecasting accuracy. Then, the best model was chosen. The ANN model showed a slightly lower but reliable result in predicting the door's future according to the inputs that were used and the approach that was followed compared to the OLS technique. The model's R^2 values for the training, cross-validation, and testing sets were 0.99, 0.98, and 0.99, respectively. The results show a strong correlation between ANN output and the actual CI obtained by visual inspection. The average R^2 of ANN was 0.993 compared to the OLS regression technique's R^2, which was 1.00. Both models revealed that some input parameters such as location, nature of space usage, and floor were statistically insignificant. Asset conditions can be predicted, and budgets for maintenance plans can be established by applying the proposed approach to different asset elements. Subsequently, random maintenance can be limited to failures only.

Although this study accomplished its primary objectives, it has some limitations; inadequate maintenance reports can be a barrier to the training and validation of a reliable element prediction model. The different perspectives regarding defect/impact matrix weights can reflect the selected element condition or space condition based on its composing elements and so on. As a result, the overall condition of the building, which may involve structural, architectural, and other elements, requires prediction models that need a large amount of data. Therefore, the need for high computing and storage capabilities is a must, which can cause errors to occur.

This research paper falls within a research agenda that is concerned with developing an AI asset management system for existing buildings. The ANN prediction accuracy could be further improved and escalated by considering other input parameters that may appear in future work. The proposed novel approach can be extended and integrated into the Internet of Things technology to provide real-time condition analysis and prediction. It can be used in conjunction with BIM technology to visually present element condition. It can also be included in an AI decentralized system using blockchain technology, a fast-growing technology. Such technology solves information problems such as immutability and transparency between different stakeholders, which will practically contribute to the industry.

Author Contributions: Conceptualization, A.M.H.; methodology, A.M.H.; software, A.M.H.; validation, M.I.S.E., A.E. and K.A.; data curation, A.M.H.; writing—original draft preparation, A.M.H.; writing—review and editing, M.I.S.E., A.E. and K.A.; visualization, A.M.H.; supervision, M.I.S.E., A.E. All authors have read and agreed to the published version of the manuscript.

Funding: This research received no external funding.

Data Availability Statement: The data presented in this study are available on request from the corresponding author. The data are not publicly available due to data ownership policies by AASTMT.

Conflicts of Interest: The authors declare no conflict of interest.

References

1. Badger, W.W.; Garvin, M.J. Facilities asset management: A new career field for construction management graduates. In Proceedings of the Associated Schools of Construction 43rd Annual International Conference, Flagstaff, AZ, USA, 11–14 April 2006; pp. 12–14.
2. U.S. Census Bureau. Value of Construction Put in Place in the United States. Available online: https://www.census.gov/construction/c30/prpdf.html (accessed on 1 February 2022).
3. Vanier, D.J. Why industry needs asset management tools. *J. Comput. Civ. Eng.* **2001**, *15*, 35–43. [CrossRef]
4. CAPMAS. Distribution of Buildings According to Building Type in 2017. 2017. Available online: https://www.capmas.gov.eg/Pages/StatisticsOracle.aspx?Oracle_id=1962&page_id=5109&YearID=23345 (accessed on 9 April 2022).
5. CAPMAS. Distribution of Regular Buildings According to the Building's Need for Renovation in 2017. 2017. Available online: https://www.capmas.gov.eg/Pages/StatisticsOracle.aspx?Oracle_id=1962&page_id=5109&YearID=23345 (accessed on 9 April 2022).
6. U.S. Census Bureau. Expenditures for Maintenance and Repairs and Improvements. 2008. Available online: https://www.census.gov/construction/c50/c50index.html (accessed on 1 February 2022).
7. Elhakeem, A.A. *An Asset Management Framework for Educational Buildings with Life-Cycle Cost Analysis*; University of Waterloo Waterloo: Waterloo, ON, Canada, 2005.
8. Robjent, L.; Clark, P.; Marti, M.; Freese, R.; Johnson, A.; County, C. *Asset Management Guide for Local Agencies*; Department of Transportation, Office of Research & Innovation: Saint Paul, MN, USA, 2019.
9. Mohamed, A.G.; Marzouk, M. Building condition assessment using artificial neural network and structural equations. *Expert Syst. Appl.* **2021**, *186*, 115743. [CrossRef]
10. Abbott, G.R.; McDuling, J.J.; Parsons, S.A.; Schoeman, J.C. Building condition assessment: A performance evaluation tool towards sustainable asset management. In Proceedings of the CIB World Building Congress, Cape Town, South Africa, 14–18 May 2007.

11. Bartram, S.M.; Branke, J.; Motahari, M. *Artificial Intelligence in Asset Management*; CFA Institute Research Foundation: Charlottesville, VA, USA, 2020.
12. Sun, H.; Burton, H.V.; Huang, H. Machine learning applications for building structural design and performance assessment: State-of-the-art review. *J. Build. Eng.* **2021**, *33*, 101816. [CrossRef]
13. Kim, G.-H.; An, S.-H.; Kang, K.-I. Comparison of construction cost estimating models based on regression analysis, neural networks, and case-based reasoning. *Build. Environ.* **2004**, *39*, 1235–1242. [CrossRef]
14. Fan, C.; Sun, Y.; Zhao, Y.; Song, M.; Wang, J. Deep learning-based feature engineering methods for improved building energy prediction. *Appl. Energy* **2019**, *240*, 35–45. [CrossRef]
15. Matel, E.; Vahdatikhaki, F.; Hosseinyalamdary, S.; Evers, T.; Voordijk, H.J. An artificial neural network approach for cost estimation of engineering services. *Int. J. Constr. Manag.* **2022**, *22*, 1274–1287. [CrossRef]
16. Peraka, N.S.P.; Biligiri, K.P.J. Pavement asset management systems and technologies: A review. *Autom. Constr.* **2020**, *119*, 103336. [CrossRef]
17. Li, D.; Cong, A.; Guo, S.J. Sewer damage detection from imbalanced CCTV inspection data using deep convolutional neural networks with hierarchical classification. *Autom. Constr.* **2019**, *101*, 199–208. [CrossRef]
18. Edirisinghe, R.; Setunge, S.; Zhang, G. Markov model—based building deterioration prediction and ISO factor analysis for building management. *J. Manag. Eng.* **2015**, *31*, 04015009. [CrossRef]
19. Mohseni, H.; Setunge, S.; Zhang, G.; Wakefield, R. Markov process for deterioration modeling and asset management of community buildings. *J. Constr. Eng. Manag.* **2017**, *143*, 04017003. [CrossRef]
20. Dais, D.; Bal, I.E.; Smyrou, E.; Sarhosis, V.J. Automatic crack classification and segmentation on masonry surfaces using convolutional neural networks and transfer learning. *Autom. Constr.* **2021**, *125*, 103606. [CrossRef]
21. Rao, A.S.; Nguyen, T.; Palaniswami, M.; Ngo, T. Vision-based automated crack detection using convolutional neural networks for condition assessment of infrastructure. *Struct. Heal. Monit.* **2021**, *20*, 2124–2142. [CrossRef]
22. Lei, X.; Xia, Y.; Komarizadehasl, S.; Sun, L. Condition level deteriorations modeling of RC beam bridges with U-Net convolutional neural networks. *Structures* **2022**, *42*, 333–342. [CrossRef]
23. Mohseni, H. *Deterioration Prediction of Community Buildings in Australia*; RMIT University: Singapore, 2012.
24. Zimmerman, K.A.; Ram, P.; Bektas, B.A.; Allen, B.; Mugabe, K.; Serulle, N.U. *Using an LCP (Life Cycle Planning) Process to Support Transportation Asset Management: A Handbook on Putting the Federal Guidance into Practice*; Federal Highway Administration: Washington, DC, USA, 2019.
25. Miyamoto, A.; Kawamura, K.; Nakamura, H.J. Bridge management system and maintenance optimization for existing bridges. *Comput.-Aided Civ. Infrastruct. Eng.* **2000**, *15*, 45–55. [CrossRef]
26. Miyamoto, A.; Motoshita, M.J. Development and practical application of a bridge management system (J-BMS) in Japan. *Civ. Eng. Infrastruct. J.* **2015**, *48*, 189–216.
27. Piaia, E.; Maietti, F.; Di Giulio, R.; Schippers-Trifan, O.; Van Delft, A.; Bruinenberg, S.; Olivadese, R.J. BIM-based cultural heritage asset management tool. Innovative solution to orient the preservation and valorization of historic buildings. *Int. J. Archit. Herit.* **2021**, *15*, 897–920. [CrossRef]
28. Piryonesi, S.M.; El-Diraby, T.E. Role of data analytics in infrastructure asset management: Overcoming data size and quality problems. *J. Transp. Eng. Part B Pavements* **2020**, *146*, 04020022. [CrossRef]
29. Phares, B.M.; Washer, G.A.; Rolander, D.D.; Graybeal, B.A.; Moore, M.J. Routine highway bridge inspection condition documentation accuracy and reliability. *J. Bridg. Eng.* **2004**, *9*, 403–413. [CrossRef]
30. Soliman, M.; Frangopol, D.M.; Kim, S.J. Probabilistic optimum inspection planning of steel bridges with multiple fatigue sensitive details. *Eng. Struct.* **2013**, *49*, 996–1006. [CrossRef]
31. Grussing, M.N.; Liu, L.Y.; Uzarski, D.R.; El-Rayes, K.; El-Gohary, N. Discrete Markov approach for building component condition, reliability, and service-life prediction modeling. *J. Perform. Constr. Facil.* **2016**, *30*, 04016015. [CrossRef]
32. Lee, C.K.M.; Cheng, M.N.; Ng, C.K. IoT-based asset management system for healthcare-related industries. *Int. J. Eng. Bus. Manag.* **2015**, *7*, 7–19.
33. Lounis, Z.; Vanier, D.; Lacasse, M. A discrete stochastic model for performance prediction of roofing systems. In Proceedings of the Proceedings CIB World Congress, Gävle, Sweden, 7–12 June 1998; pp. 203–313.
34. Schneider, J.; Gaul, A.J.; Neumann, C.; Hogräfer, J.; Wellßow, W.; Schwan, M.; Schnettler, A.J. Asset management techniques. *Int. J. Electr. Power Energy Syst.* **2006**, *28*, 643–654. [CrossRef]
35. Hagan, M.T.; Demuth, H.B.; Beale, M. *Neural Network Design*; PWS Publishing Co.: Boston, MA, USA, 1997.
36. Gurney, K. *An Introduction to Neural Networks*; CRC Press: Boca Raton, FL, USA, 2018.
37. Aggarwal, C.C. *Neural Networks and Deep Learning*; Springer: Berlin/Heidelberg, Germany, 2018; Volume 10, pp. 973–978.
38. Karsoliya, S.J. Approximating number of hidden layer neurons in multiple hidden layer BPNN architecture. *Int. J. Eng. Trends Technol.* **2012**, *3*, 714–717.
39. Tijanić, K.; Car-Pušić, D.; Šperac, M.J. Cost estimation in road construction using artificial neural network. *Neural Comput. Appl.* **2020**, *32*, 9343–9355. [CrossRef]
40. Günaydın, H.M.; Doğan, S.Z. A neural network approach for early cost estimation of structural systems of buildings. *Int. J. Proj. Manag.* **2004**, *22*, 595–602. [CrossRef]

41. Adel, K.; Elhakeem, A.; Marzouk, M.J. Decentralizing Construction AI Applications Using Blockchain Technology. *Expert Syst. Appl.* **2022**, *194*, 116548. [CrossRef]
42. Chollet, F. *Deep Learning with Python*; Simon and Schuster: New York, NY, USA, 2021.
43. Kecman, V. *Learning and Soft Computing: Support Vector Machines, Neural Networks, and Fuzzy Logic Models*; MIT Press: Cambridge, MA, USA, 2001.
44. Adeli, H.J. Neural networks in civil engineering: 1989–2000. *Comput.-Aided Civ. Infrastruct. Eng.* **2001**, *16*, 126–142. [CrossRef]
45. Angulo, A.; Vega-Fernández, J.A.; Aguilar-Lobo, L.M.; Natraj, S.; Ochoa-Ruiz, G. Road damage detection acquisition system based on deep neural networks for physical asset management. In Proceedings of the Mexican International Conference on Artificial Intelligence, Xalapa, Mexico, 27 October–2 November 2019; pp. 3–14.
46. Assaad, R.; El-adaway, I.H. Bridge infrastructure asset management system: Comparative computational machine learning approach for evaluating and predicting deck deterioration conditions. *J. Infrastruct. Syst.* **2020**, *26*, 04020032. [CrossRef]
47. Perez, H.; Tah, J.H.; Mosavi, A.J. Deep learning for detecting building defects using convolutional neural networks. *Sensors* **2019**, *19*, 3556. [CrossRef]
48. Zhang, H.; Feng, H.; Hewage, K.; Arashpour, M. Artificial Neural Network for Predicting Building Energy Performance: A Surrogate Energy Retrofits Decision Support Framework. *Buildings* **2022**, *12*, 829. [CrossRef]
49. Hong, J.; Choi, H.; Kim, W.-S. A house price valuation based on the random forest approach: The mass appraisal of residential property in South Korea. *Int. J. Strat. Prop. Manag.* **2020**, *24*, 140–152. [CrossRef]
50. Ali, A.; Hegazy, T. Performance assessment framework for healthcare facilities. In Proceedings of the 4th Specialty Conference on Coastal, Estuary and Offshore Engineering, Montreal, QC, Canada, 29 May–1 June 2013.