A Bayesian Network Framework to Predict Compressive Strength of Recycled Aggregate Concrete

Tien-Dung Nguyen, Rachid Cherif, Pierre-Yves Mahieux, Emilio Bastidas-Arteaga

Journal of
Composites Science

MDPI

Article

A Bayesian Network Framework to Predict Compressive Strength of Recycled Aggregate Concrete

Tien-Dung Nguyen [1,2], Rachid Cherif [1], Pierre-Yves Mahieux [1] and Emilio Bastidas-Arteaga [1,*]

[1] Laboratory of Engineering Sciences for the Environment (LaSIE) UMR CNRS 7356, University of La Rochelle, Avenue Michel Crépeau, 17042 La Rochelle Cedex 1, France; ntdung@dut.udn.vn (T.-D.N.); rachid.cherif@univ-lr.fr (R.C.); pierre-yves.mahieux@univ-lr.fr (P.-Y.M.)
[2] Faculty of Road and Bridge Engineering, The University of Danang—University of Science and Technology, 54 Nguyen Luong Bang Street, Lien Chieu District, Danang City 550000, Vietnam
* Correspondence: ebastida@univ-lr.fr; Tel.: +33-(0)5-86-56-22-32

Abstract: In recent years, the use of recycled aggregate concrete (RAC) has become a major concern when promoting sustainable development in construction. However, the design of concrete mixes and the prediction of their compressive strength becomes difficult due to the heterogeneity of recycled aggregates (RA). Artificial-intelligence (AI) approaches for the prediction of RAC compressive strength (f_c) need a sizable database to have the ability to generalize models. Additionally, not all AI methods may update input values in the model to improve the performance of the algorithms or to identify some model parameters. To overcome these challenges, this study proposes a new method based on Bayesian Networks (BNs) to predict the f_c of RAC, as well as to identify some parameters of the RAC formulation to achieve a given f_c target. The BN approach utilizes the available data from three input variables: water-to-cement ratio, aggregate-to-cement ratio, and RA replacement ratio to calculate the prior and posterior probability of f_c. The outcomes demonstrate how BNs may be used to forecast both forward and backward, related to the f_c of RAC, and the parameters of the concrete formulation.

Keywords: Bayesian networks; compressive strength; formulation; recycled aggregate concrete; prediction

Academic Editor: Khandaker M. A. Hossain

Received: 23 November 2024
Revised: 15 January 2025
Accepted: 3 February 2025
Published: 5 February 2025

Citation: Nguyen, T.-D.; Cherif, R.; Mahieux, P.-Y.; Bastidas-Arteaga, E. A Bayesian Network Framework to Predict Compressive Strength of Recycled Aggregate Concrete. *J. Compos. Sci.* **2025**, *9*, 72. https://doi.org/10.3390/jcs9020072

1. Introduction

Construction and demolition waste (CDW) is a significant part of solid waste, around 25% worldwide [1]. Concrete makes up most of CDW, accounting for around 70% of the material [2]. Recycled aggregate (RA) comes from many resources for demolishing traditional concrete for buildings and transportation facilities. Using RA is crucial as a partial substitution in concrete production, as it helps conserve the environment by decreasing the quantity of CDW that must be dumped in landfills for disposal [3].

The hardened performance of RAC is affected by different parameters, including design variables and the properties of RA [4]. In detail, the adhered mortar of RA decreases its quality by increasing its porosity and water absorption (WA), resulting in a decreased hardened performance of recycled aggregate concrete (RAC) [5,6]. Furthermore, the interfacial transition zones (ITZs) between RA and the mortar also affect the properties of concrete [7]. Almost all studies reported that RAC's compressive strength decreased along with increased replacement level [8,9], but some studies have reported the opposite idea, especially at a later age [10,11]. Such different conclusions are based on the remaining

non-hydrated, adhered mortar stuck to the RCA surfaces, thus increasing the late compressive strength. Therefore, determining the link between the proportions of the mix, the properties of RA, and the compressive strength is a challenge.

Artificial-intelligence methods have been developed to predict RAC's compressive strength. Among these methods are: Artificial Neural Networks (ANNs), Radial Basis Function, Neuro-Fuzzy Inference System, Genetic Programming, and Support Vector Machine. Dantas et al. [12] utilized ANNs to predict the strength of RAC at various ages, using 24 input variables. A study from Duan et al. [13] also used 14 input parameters in ANNs to predict the hardened performance of RAC at 28 days. At the same time, Deshpande et al. [14] used ANNs, Model Tree (MT), and Non-linear Regression (NLR) with the same aim by considering the same 14 input parameters as Deshpande et al. [14]. Dabiri et al. [15] implemented NLR and the Random Forest algorithms to estimate RAC's compressive strength and compared them to ANNs and MT models. All of the above studies gave good prediction results with coefficients of determination (R^2) above 0.9 [16]. However, they required a significant amount of data.

Many real-world applications have adopted Bayesian networks because this methodology could be used to estimate specific variables, handle uncertainties, and integrate observations (evidence) to facilitate decision analysis and provide prompt responses to end users [17,18]. The probabilistic inference of BNs requires information about the prior probability distribution, the influence of variables, and the reasoning from cause to effect and from effect to cause [19]. Once the BN is developed, it can be updated anytime, whenever new data become available [20,21]. Furthermore, the BN method works with small datasets and allows for backward prediction, which is useful for the formulation problem when the target compressive strength (f_c) is known. As mentioned above, artificial-intelligence approaches have been widely applied to predict the hardened properties of RAC. However, the potential of BNs, especially Bayesian updating, has not been comprehensively exploited to predict the hardened properties or the formulation of RAC.

The main objective of this paper is to propose a new BNs-based framework to predict the f_c and some formulation parameters of RAC. In Section 2, the architecture of the BN is presented. The BN model is developed using a Python library. Data from many published papers and experimental data are used to assess the conditional probability tables (CPTs). Original experimental data are also used to assess the relative error between the experimental results and the model predictions obtained after Bayesian updating (Section 3). In Section 4, Bayesian inference is used to identify some parameters in the formulation of RAC to reach target f_c values.

2. Bayesian Networks

Bayesian networks were first developed in the late 1970s [22]. BNs are graph models that show the cause–effect relationship between the input and output variables. BNs are based on Bayesian theory, a branch of probability theory. BNs are a popular tool in mathematics and engineering that combines graph and probability principles to effectively address complex problems [23]. An illustration of a BN is a directed acyclic graph (DAG), in which the nodes depict the relationship parameters and the links among them show any causal or informative associations between the variables. Subjective probability in a BN model captures the uncertainties. As depicted in Figure 1, a BN is composed of (a) the variables and directed links, (b) the states for each variable, and (c) the conditional probability for each variable.

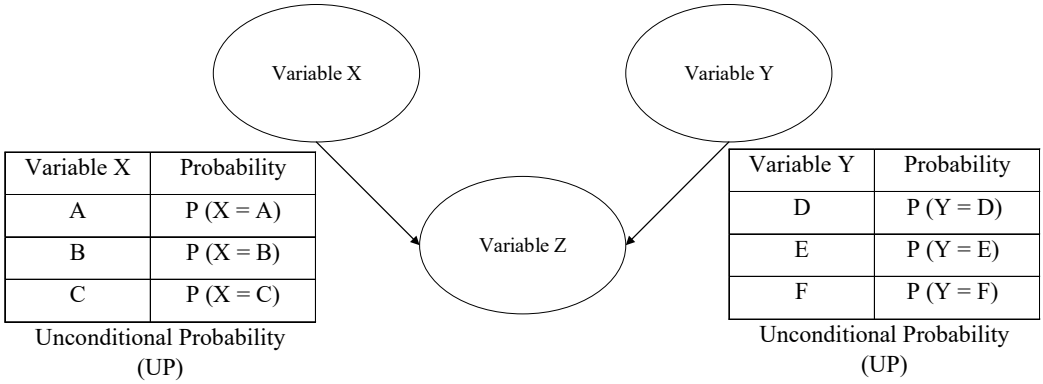
In BNs, the relationships between the variables are described by nodes and links. For example, a variable X is the parent of another variable Z (and vice versa) if a link exists between them. These dependencies are quantified using conditional probability

tables (Figure 1). Based on BNs, the Bayes theorem provides a convincing approach to managing uncertainty by openly displaying the conditional probability connections between parameters [24]. In the BNs approach, for n number of parameters A_i with $i = 1, 2, \ldots, n$, and a given observation B, the probability $P(A_j|B)$ is calculated as follows:

$$P(A_j|B) = \frac{P(B|A_j) \times P(A_j)}{\sum_{i=1}^{n} P(B|A_i) \times P(A_i)} \quad (1)$$

where $P(A_j|B)$ represents the posterior probability of A_j given the condition that B occurs, $P(A_j)$, $P(A_i)$ denote the prior probability of A_j, A_i, and $P(B|A_j)$, $P(B|A_i)$ refer to the conditional probabilities of B given that A_j or A_i occur [25].

The ability of BNs to record both upward inferences, observing the input nodes and deducing potential impacts and backward inferences, observing the influence of child nodes, and deducing potential parent causes, makes them an effective tool for assessing some parameters [17]. Marginalizing the joint probability, it is possible to determine a given parameter probability distribution [19]. Fundamentally, the probability distributions of the nodes composing the BN are updated when evidence is introduced. The conditional probability tables are computed from Equation (1) using information available in the literature [26], derived from data [27], and/or using physics-based models [28].



Variable X	Probability
A	P (X = A)
B	P (X = B)
C	P (X = C)

Unconditional Probability (UP)

Variable Y	Probability
D	P (Y = D)
E	P (Y = E)
F	P (Y = F)

Unconditional Probability (UP)

Variable X	Variable Y	Variable Z		
		Probability		
		I	J	K
A	D	P (Z=I|X=A, Y=D)	P (Z=J|X=A, Y=D)	P (Z=K|X=A, Y=D)
…	…	…	…	…
B	E	P (Z=I|X=B, Y=E)	P (Z=J|X=B, Y=E)	P (Z=K|X=B, Y=E)
C	F	P (Z=I|X=C, Y=F)	P (Z=J|X=C, Y=F)	P (Z=K|X=C, Y=F)

Figure 1. A schematic BN.

3. Developed Bayesian Network

The proposed BN framework was developed following these stages:

1. Data collection and filtering: Experimental studies on RAC are the basis for building the BN. Since testing is costly and requires long analyses and processing times, it is necessary to collect data from previous studies on compressive strength (f_c) of RAC. It is widely known that there are many factors affecting the strength of RAC (e.g., cement content, mineral additives, water-to-cement ratio (w/c), aggregate-to-

cement ratio (a/c), RA replacement ratios (R), etc.). The information on these factors varies for each study in the literature [29–32]. In addition, there are many studies on high-performance concrete with RA where the RAC formulation is quite specific. Considering the availability of data in the studies, we have selected three input parameters that influence f_c: w/c, a/c, and R. The w/c, R are the parameters which have been full investigated in the literature [33,34]. Furthermore, we included a new parameter a/c, which accounts for the skeleton structure which is a key parameter for low strength RAC [35]. In this study, we focused on RAC with an average strength of 20–35 MPa.

2. Model development: This stage requires a deep understanding of the data as well as the relationships between the variables. This stage has the following steps:
   - The architecture of the BN is defined considering causal relationships between the selected parameters and the objective of the model. In this case, the objectives are to estimate f_c (output) as a function of three input parameters: w/c, a/c, and R or to identify the input parameters to achieve a target f_c.
   - The data collected in the first stage was attributed to the input and output nodes in the model. After, the discretization of each node can be determined by considering the data availability and using similar research cases or expert knowledge.
   - By figuring out how frequently the value of the parent node appears when the value of the child node does, we estimated the conditional probability tables of the BN.
3. Updating: Once the architecture of the BN and the CPTs were defined in the BN, prior distributions and evidence could be introduced in the BN for prediction purposes. Two case studies with specific objectives were defined. Figure 2 depicts the considered case studies. The first objective is to predict f_c by updating the parent nodes with the characteristics of a given formulation. The second objective is to determine some parameters of the RAC formulation by updating a target compressive strength.

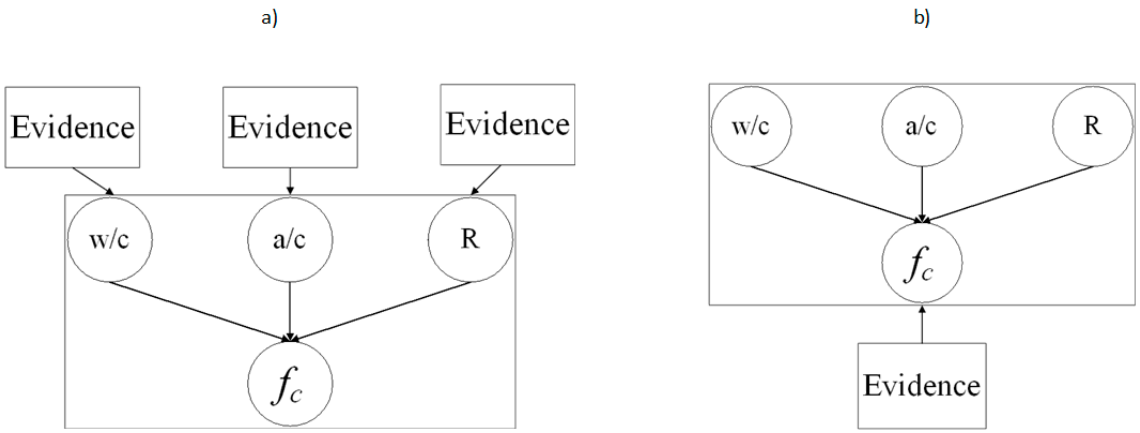


Figure 2. Objectives of the two considered case studies: (**a**) prediction of f_c, (**b**) formulation design for a target f_c.

3.1. BN Architecture and Database Description

The BN architecture proposed in this study is shown in Figure 3. The BN model is developed in Python. The BN has three variables as inputs (w/c, a/c, and R) and f_c as output. This configuration allows estimating the compressive strength as a function of the three parent nodes or determining the parent node values when aiming to reach a given f_c after updating.

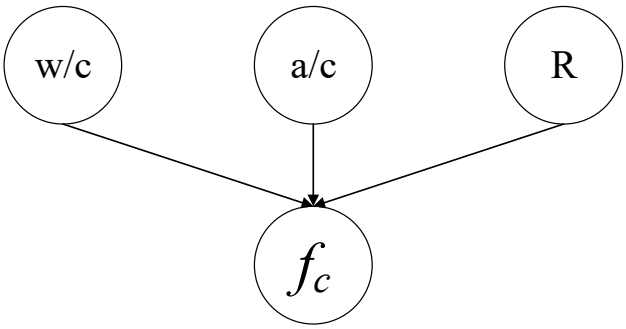


Figure 3. Proposed BNs model.

The f_c of RAC depends on formulation variables such as w/c, RA replacement ratio (R), physical characteristics of RA, etc. [36–39] In this study, the BN considers a database with 81 datasets obtained from various sources in the literature (63 datasets) [29–32], and own experiments (18 datasets). Inconsistent or low-quality data points obtained from other studies may behave as outliers, which could skew the network's learned dependencies and produce inaccurate predictions. To reduce the impact of data quality, we carefully examined and excluded outliers during the setup of the database. However, data quality remains a challenge when using data from different sources for prediction purposes. Appendix A provides a table with the datasets obtained from the literature that were used in this work. For the experiments, f_c was determined for six formulations with w/c = 0.65 to 0.7, a/c = 3.3 to 2.9, and various R values (R = 0%, 20%, 40%, 60%, 80%, 100%). Appendix B summarizes the material and methods of the experiments as well as details the dataset.

Figure 4 displays the scatterplot of the database. About 24% of w/c values are primarily concentrated in the range of 0.6 to 0.65. For the aggregate-to-cement ratio (a/c), 64% of values are concentrated in the range 3 to 3.5. For f_c, 53% of values are found in the range of 25–30 MPa. For R, 33% and 30% of values are far away from the mean and concentrated in the ranges 0–25 and 75–100, respectively. This is mainly due to the fact that in the literature, many studies take as a reference concrete without CRA ($R = 0\%$) or with a complete replacement ratio ($R = 100\%$).

Determining whether a relationship exists among the variables being studied is crucial. Therefore, Figure 5 provides the Pearson correlation coefficients, which quantify these links by determining the correlation between variables. For the parent nodes, as the values range from -0.4 to 0.29, as depicted in Figure 5, it is possible to assume that there are no linear correlations between w/c, a/c, and R. Consequently, we assume that these three parameters are independent. f_c has moderate to fairly strong correlations with the parent nodes, confirming causal relationships between the child and the parent nodes.

3.2. BN Discretization and Prior Probabilities

All nodes are defined on a given domain and partitioned into various states in this stage. The intervals of variables should include the variables' theoretical/physical values and correspond to the values available in the database. These ranges were determined considering the information available in the database and the literature. The information on nodes is depicted in Table 1. In this study, as there is a large variability of values for the nodes w/c, a/c, and f_c, we considered a confidence interval of 95% from the database to define the boundaries of the BN.

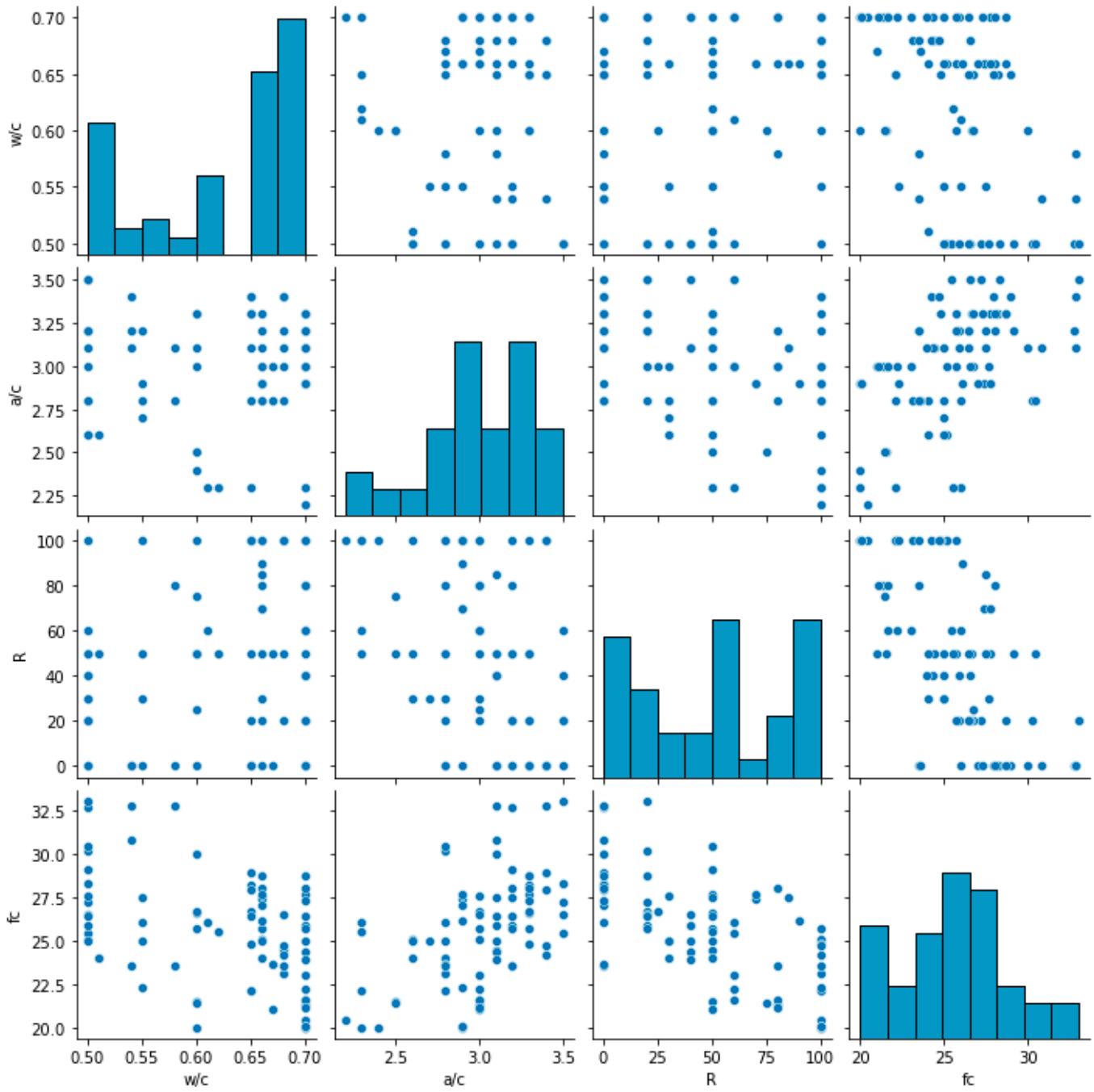


Figure 4. Scatterplot for the database.

Table 1. The information of nodes in the BNs model.

Nodes	Number of States	Boundaries of the BN	Range of States
w/c	4	[0.5, 0.7]	0.5–0.55, 0.55–0.6, 0.6–0.65, 0.65–0.7
a/c	3	[2, 3.5]	2–2.5, 2.5–3, 3–3.5
R (%)	6	[0, 100]	0–10, 10–30, 30–50, 50–70, 70–90, 90–100
f_c (MPa)	5	[20, 35]	20–23, 23–26, 26–29, 29–32, 32–35

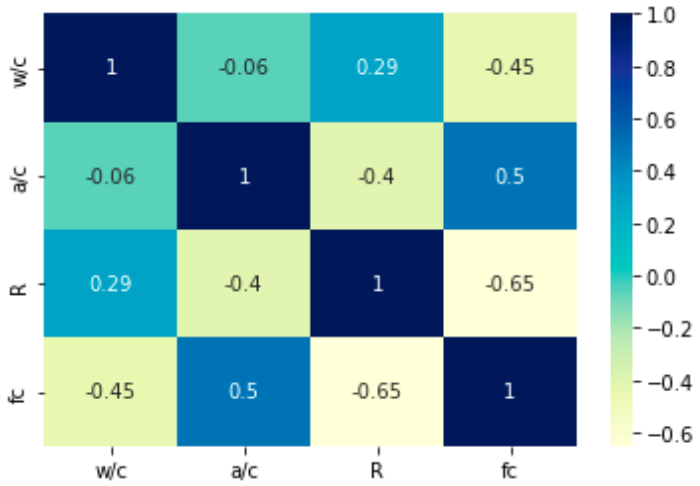


Figure 5. Correlation matrix of input parameters.

The prior probabilities were calculated from the frequencies of the database. The prior distributions for the parent and the child nodes are depicted in Figure 6. In Figure 6, the w/c prior values are focused on the states 0.5–0.55 and 0.65–0.7; these are 22.2% and 59.3%, respectively. Regarding a/c, the values in the prior probability focus on 64.2% at state 3–3.5. The R values in the distribution emphasize the same probability of 22.2% at states 50–70 and 90–100, respectively. The values of f_c concentrate on states 20–23 MPa, 23–26 MPa, and 26–29 MPa, with prior probabilities of 20.9%, 33.3%, and 34.5%, respectively. However, the distribution of R is uniform, indicating that the discretization of the database has approximatively a uniform representation of this parameter that is quite important when we try to predict the performance of RAC. The concentration of data in the first three states for f_c could result in more accurate predictions when targeting a f_c between 20 and 29 MPa.

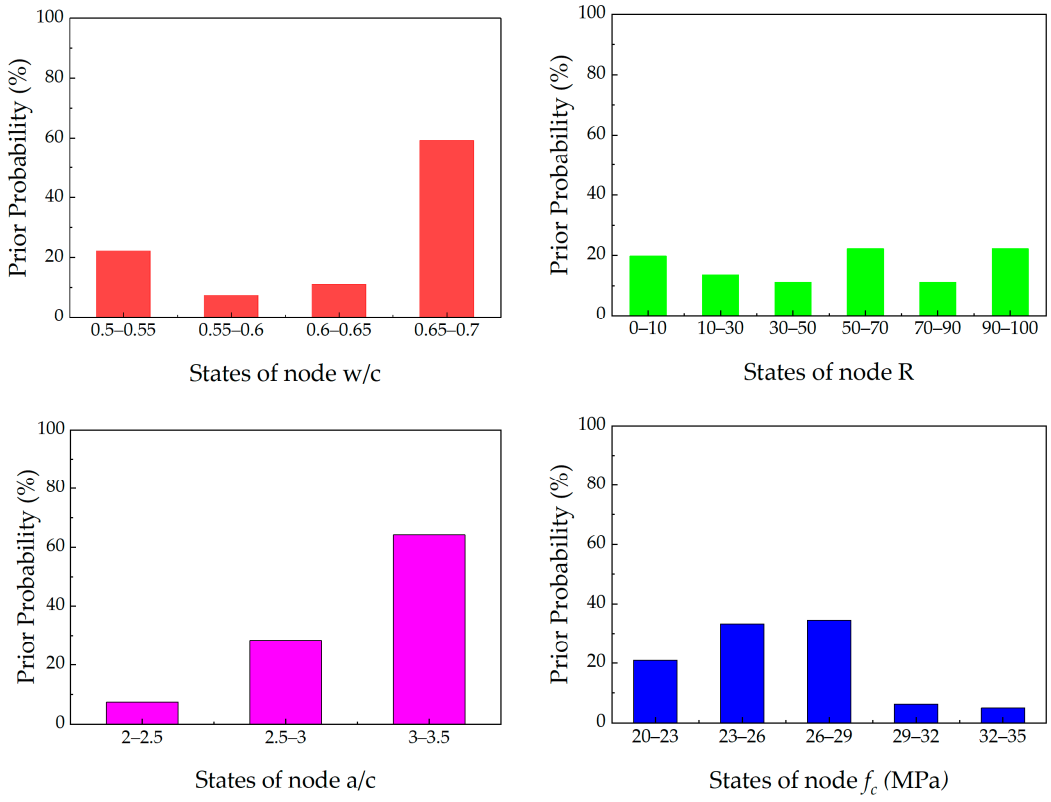


Figure 6. Prior probability of all nodes.

The relationships in a BN are modeled by a set of CPTs. The CPT sizes would be ($4 \times 3 \times 6$) rows and 5 columns (see an extraction in Table 2). These CPTs have been applied for BNs. For example, in Table 2, the CPT values of f_c focus on the state 26–29 when w/c in the state 0.65–0.7, a/c in the state 2–2.5, and R in the state 0–10 (row 67). When increasing R, and keeping the same states for w/c, a/c, the conditional probability values of f_c concentrate in the lower states of compressive strength (rows 68–72). This means that for fixed values of w/c and a/c, the probability of obtaining low compressive strength decreases when the replacement ratio increases.

Table 2. Extract of the CPTs of the BNs model.

No. of Row	Node and States			States of Node f_c and Conditional Probability				
	w/c	***a/c***	***R***	**20–23**	**23–26**	**26–29**	**29–32**	**32–35**
1	0.5–0.55	2–2.5	0–10	0	0	0	1	0
⋮	⋮	⋮	⋮		⋮	⋮	⋮	⋮
67	0.65–0.7	3–3.5	0–10	0	0	1	0	0
68	0.65–0.7	3–3.5	10–30	0	0.33	0.67	0	0
69	0.65–0.7	3–3.5	30–50	0	1	0	0	0
70	0.65–0.7	3–3.5	50–70	0.6	0.2	0.2	0	0
71	0.65–0.7	3–3.5	70–90	0.6	0.4	0	0	0
72	0.65–0.7	3–3.5	90–100	0.5	0.5	0	0	0

3.3. Assessment of Posterior Probabilities

The marginal probability distribution of the output node can be determined by the following equation:

$$P(f_c) = \sum_{w/c,a/c,R} P(f_c|w/c, a/c, R)P(w/c, a/c, R) \quad (2)$$

where $P(w/c, a/c, R) = P(w/c)P(a/c)P(R)$. A Python library was used to implement this BN architecture and determine the posterior probability of f_c by Bayesian inference. Considering the prior probabilities presented in the previous section, the probabilities of the child node are depicted in Figure 7. The results show that in the state 23–26 MPa is the highest probability of f_c, which is equal to 75.1%. The probabilities rise compared to prior values for states 23–26 by 41.8%, whereas the posterior probability of states 20–23 MPa and 26–29 MPa decrease by 12.75% and 24.6%, respectively. For the state 29–32 MPa, the posterior probability is 2.71%, which is low compared to the prior probability of 4.9%. In summary, these results illustrate how the prior information could be used to estimate the distribution of a variable of interest, which is, in this case, f_c. Further applications of the proposed BN framework are provided and discussed in the next section.

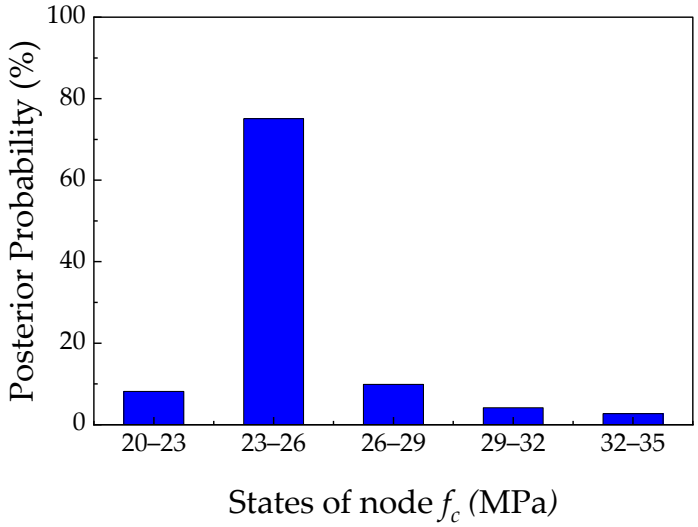


Figure 7. Posterior probability of compressive strength.

4. Applications

4.1. Prediction of Compressive Strength by Bayesian Updating

Among the benefits of the BN-based approach is the ability to update CPTs and node evidence when dealing with new information or data [40,41]. It is then possible to design a model utilizing the best data and information available and incorporate evidence later on by adapting the BN model's performance to new information or data [42,43]. The findings enable the use of already available data as proof of variable correlations or observational data for distribution updates.

In this section, we used our experimental results (Appendix B) to analyze the impact of introducing new knowledge or evidence when reevaluating beliefs or updating probabilities. In the approach, evidence was introduced in the parent nodes to update the model output. Observational data were used to update the output node's posterior probability because it was related to every input node. Experimental data were used to define the prior information of the BN and the evidence of the parent nodes (w/c, a/c and R). Experimental results for f_c were compared with BN predictions to test the accuracy of the BN. The evidence used for the parent nodes is presented in Table 3. This evidence allows the testing of the ability of the BN to predict f_c when changing the replacement ratio.

Table 3. Evidence in BNs updating of the parent nodes (all the evidence is concentrated in one target stage).

Case	*w/c*	*a/c*	*R*
1	[0.65–0.7] = 100%	[3–3.5] = 100%	[0–10] = 100%
2	[0.65–0.7] = 100%	[3–3.5] = 100%	[10–30] = 100%
3	[0.65–0.7] = 100%	[3–3.5] = 100%	[30–50] = 100%
4	[0.65–0.7] = 100%	[3–3.5] = 100%	[50–70] = 100%
5	[0.65–0.7] = 100%	[3–3.5] = 100%	[70–90] = 100%
6	[0.65–0.7] = 100%	[2.5–3] = 100%	[90–100] = 100%

Figure 8 shows the posterior probabilities of the studied cases. With small replacement ratios (R = [0–10%], reference case), the posterior probability values of f_c are 58.5% and 18.2% for the states [26–29 MPa] and [29–32 MPa], respectively. The larger probability for these states indicates that value f_c for this reference concrete is reasonable in comparison

with the results for the other cases because the reference concrete always has the highest strength compared to RAC with the same/close w/c. Furthermore, the reference concrete has no recycled aggregate, which is the main reason for the reduction of f_c because of the large water absorption of CRA, which increases the total water quantity and reduces the compressive strength [44]. When R increases, there is a reduction in the compressive strength, which is consistent with the literature [45] and our experimental observations (Appendix B). When R = 100%, larger probabilities are observed at the states [20–23 MPa] and [23–26 MPa], which correspond to the lower values of f_c. The compressive strength decreases for large R values by two reasons: (i) using large amounts of CRA with high absorption will lead to weak the bond between the RA and the cement matrix, reducing the concrete strength; and (ii) using large amounts of CRA decreases the shear strength of the aggregate matrix because the CRA has a smaller crushing value than the natural aggregate [46].

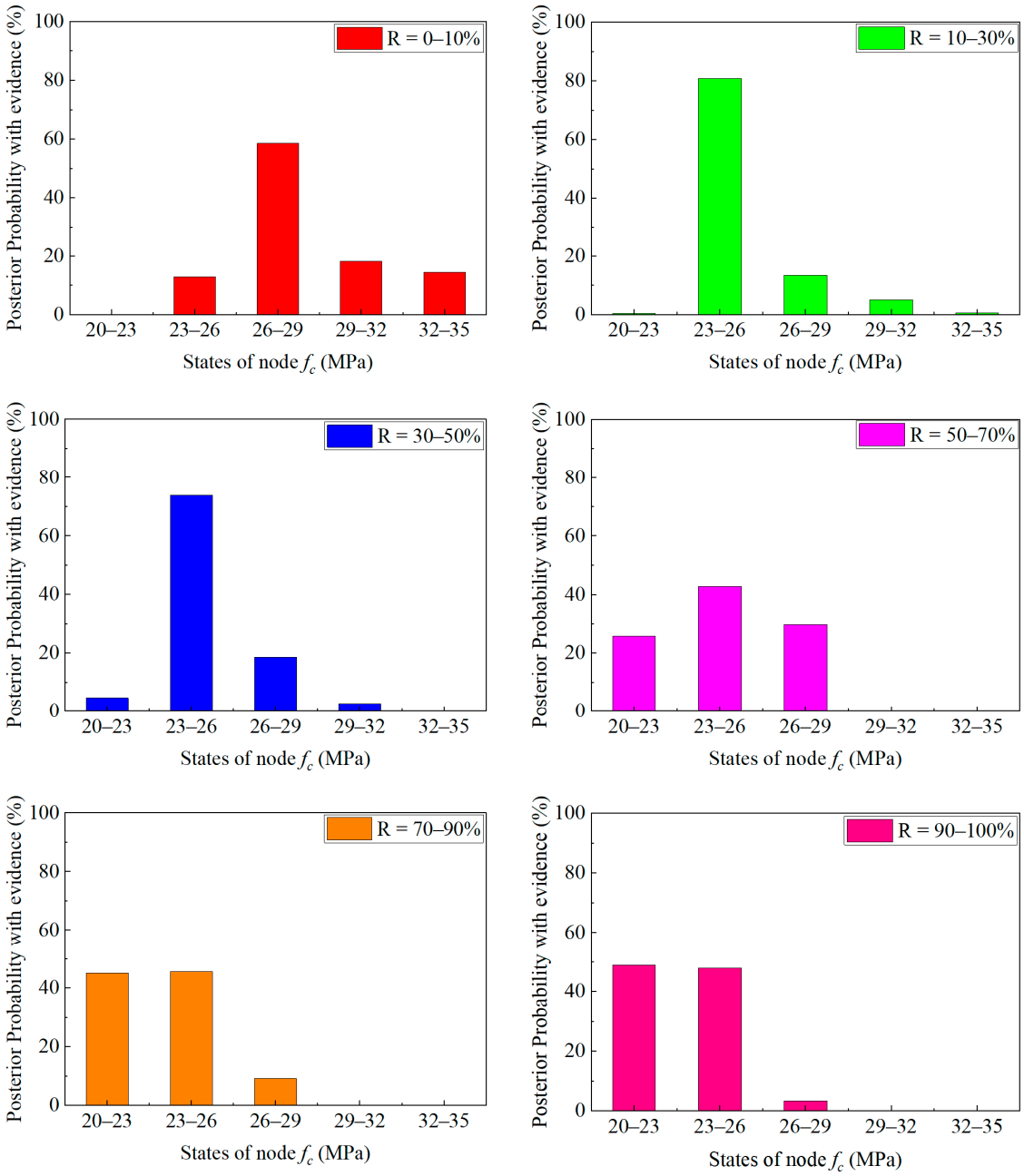


Figure 8. Posterior probabilities of f_c node for several pieces of evidence.

To evaluate the efficiency of the proposed BN configuration, the relative error (E_M) between the mean of f_c obtained using BNs (M_{BNs}) and the mean computed from experimental values (M_E) was used in this study [47]:

$$E_M = \frac{|M_E - M_{BNs}|}{M_{BNs}} \times 100\% \quad (3)$$

The means for f_c estimated from Bayesian inference and experimental data are depicted in Figure 9. M_{BNs} with R = 0% is 29.6 MPa, and the corresponding relative error is 3%. With R = 20%, the value of M_{BNs} decreases to 25.2 MPa, and the relative is 2.7%. These E_M are small, indicating that, for these replacement rates, the predictions have a very good agreement with the experimental data. E_M increases by 3.1%, 8.1%, 8.9% and 10.5% for R = 40%, 60%, 80% and 100%, respectively. The relative errors of values predicted with BNs are low when $R \leq$ 20% and around 10% for high replacement ratios. These small errors demonstrate that the BNs approach outperforms other models, like bagging regressor, decision tree, and gradient boosting, that are employed in the literature with the same dataset [29]. When R =10–30%, the f_c value of the BNs is smaller than the experimental value, which is different from the trend of the previous cases. The reason here may be the lack of data in this case. Furthermore, the error of the BNs increases significantly when considering large replacement ratios. This can be explained by the lack of experimental data or may be due to the nature of the CRA used in our experiments (included recycled concrete, recycled brick, etc.) being different from the CRA used in the studies from the literature (only recycle concrete). These results could be enhanced by filtering the information used to create the BN configuration and optimizing the discretization of the nodes in the BN [48]. Furthermore, the model can include other parent nodes, such as the cement content, the composition of CRA, etc., to reduce the relative errors. This requires more complete datasets.

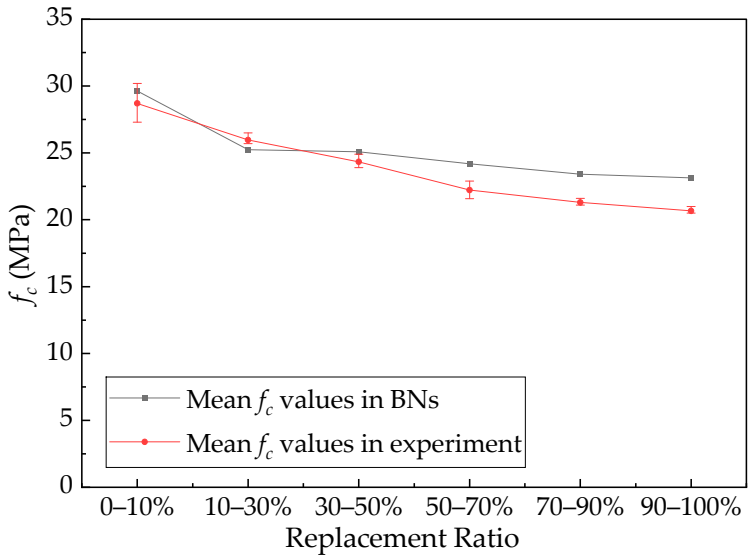


Figure 9. Means for f_c obtained with BNs and experimental data.

4.2. Posterior Probabilities of Parent Nodes for a Target Compressive Strength

In this section, we will fix a target f_c and use the distribution of this target to estimate, by Bayesian updating, the required distributions of the parent nodes. More specifically, the new information will update the BN with a histogram where the target is to obtain a given resistance. Figure 10 shows the configuration of the BN, the prior distributions of the parent nodes, and the evidence of the child node. In this case, we assume that the prior

probabilities of parent nodes are: w/c: 80% for [0.5–0.55] and 20% for [0.55–0.6]; and for a/c and R the prior probabilities have the same values for all the states. This means that we are looking for a formulation with low w/c and we do not make any assumption about the values for a/c and R. We suppose that after that, the target probability value (evidence) for the compressive strength has a probability of 100% belonging to the state [23–26 MPa]. This corresponds to a mean f_c value of 21.5 MPa.

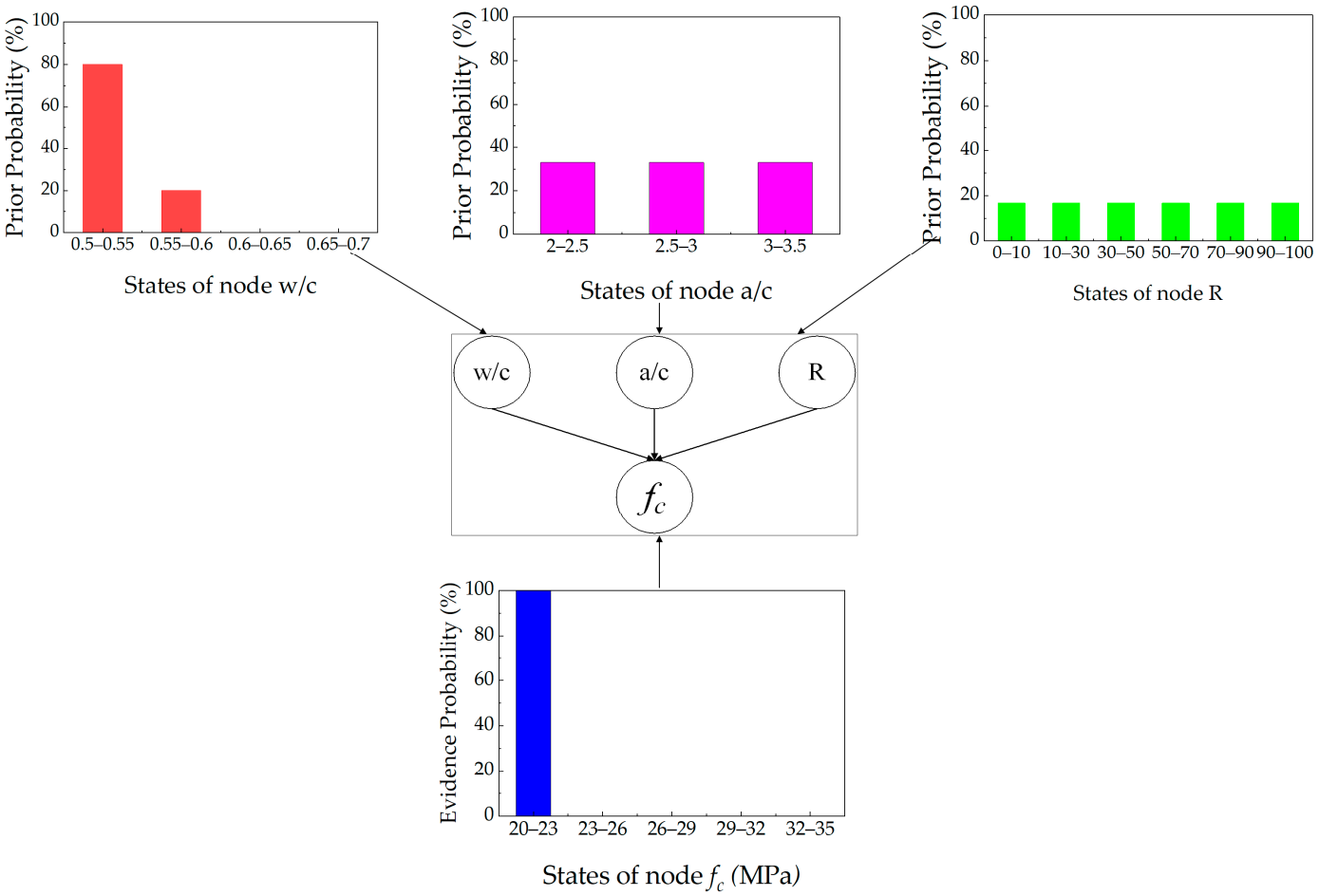


Figure 10. Prior distributions of the parent nodes and evidence of the child node.

Figure 11 presents the posterior probabilities of w/c, a/c , and R after updating the f_c node. In comparison with Figure 10, the posterior probabilities of w/c, a/c, and R changed with the evidence. The probability values of w/c have a maximum value of 68.7% for states [0.5–0.55], which is low compared to the prior probability of 80%. The values of the node a/c changed to a maximum value of 66.4% at state [2–2.5], which is a high value compared to the prior value of 33.3%. The posterior probabilities of R are large for the states [70–90%] and [90–100%]; these results are 30.53% and 47.51%, which are high compared to 16.7% and 16.7% of prior values, respectively. These values could provide some insights into the concrete formulation (w/c, a/c, and R) required to obtain a target compressive strength. For example, the mean values of parent nodes in this example are 0.54, 2.54, and 80.7% for w/c, a/c and R, respectively. Furthermore, the results could be improved by filtering the data used to build the CPT as well as optimizing the discretization of the nodes in the BNs, as mentioned previously. Further work should be addressed in both areas.

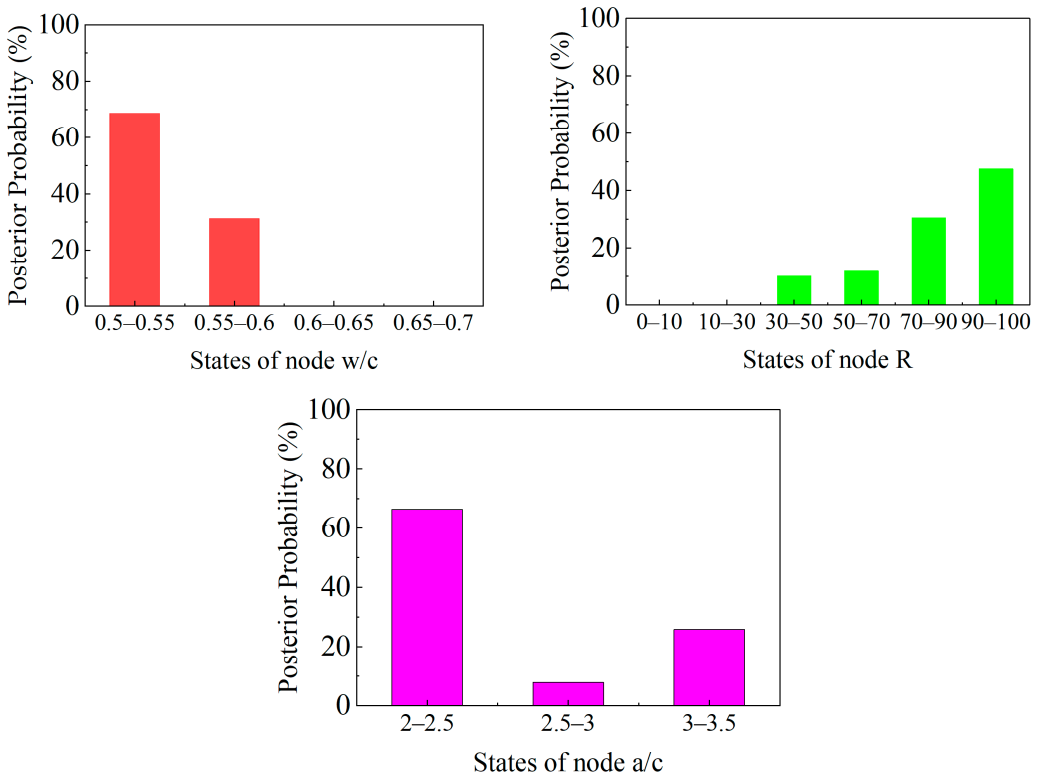


Figure 11. Posterior probabilities of the parent nodes.

Table 4 summarizes the mean values of all parent nodes with various cases of f_c evidence. The following targets (evidence) for f_c were considered: [20–23 MPa]: 100%; [23–26 MPa]: 100%; [26–29 MPa]: 100%; [29–32 MPa]: 100%; [32–35 MPa]: 100%. It can be observed in Table 4 that when increasing f_c, the mean values of w/c decrease slightly: 1.7% in comparison to the larger value obtained for the target f_c of [20–23 MPa]. This agrees with the fact that when the w/c ratio is reduced, the strength of the cement paste will increase, leading to an increase of f_c. For a/c there is a clear trend: when rising the a/c, the f_c also increases. This is explained by the positive correlation between a/c and f_c estimated for the data in Figure 5. Furthermore, when the a/c ratio has low values as 2–3.5 in the model, the aggregate matrix strength depends on the low strength of the RAC [35]. Mean values of R also decrease from 80.7% to 10.8% which is consistent with the trend in experiment data. Furthermore, R also has a great influence on f_c, since CRA has high water absorption and lower crushing value compared with natural aggregates. The influence of w/c is much smaller than that of R, a/c after updating, because R, a/c has a larger relationship with f_c as observed in in Figure 5 [16]. These results illustrate a practical use of the BN for optimal mix design. Further studies could also include durability indicators and other properties targets of RAC.

Table 4. Mean values of parent nodes with various f_c targets.

Target f_c (MPa)	Mean Values		
	w/c	*a/c*	*R* (%)
20–23	0.540	2.54	80.7
23–26	0.539	2.79	68.3
26–29	0.535	2.86	46.4
29–32	0.534	2.92	14.6
32–35	0.528	3.25	10.8

5. Sensitivity Analysis

A sensitivity analysis was carried out to investigate further the main factors that affect the prediction of compressive strength. Entropy reduction or mutual information is used in this sensitivity analysis:

$$\mathrm{I}(X; f_c) = \mathrm{H}(f_c) - \mathrm{H}(f_c|\mathrm{X}) \quad (4)$$

where X can be any input (w/c, a/c or R), $\mathrm{I}(X; f_c)$ is the mutual information between the input X and the output f_c, $\mathrm{H}(f_c)$ is the entropy of the output variable, and $\mathrm{H}(f_c|\mathrm{X})$ is the conditional entropy of f_c given the input X. The results for this BN are presented in Table 5. Entropy reduction measures the amount of information gained about the output by observing a particular input parameter. This metric captures the degree of dependence between the input parameter (i.e., w/c, a/c or R) and the output (f_c). In Table 5, w/c has the larger mutual information, followed by a/c and R. Higher entropy reduction values indicate that w/c is the more informative parameter for predicting the output. However, the small differences found for a/c and R indicate that these variables also have a meaningful impact on the determination of f_c.

Table 5. Results of the sensitivity analysis.

$\boldsymbol{X}$	$\mathbf{I}(\boldsymbol{X};\boldsymbol{f_c})$
w/c	0.15
a/c	0.13
R	0.11

6. Conclusions

This study proposed a BN model based on the literature and own experimental data to predict the compressive strength when w/c, a/c, and R are defined. The BN also served to find combinations of w/c, a/c, and R to reach a given compressive strength target. The main conclusions of the study are summarized as follows:

- The relative error between the mean values obtained by BN updating and the experiments is around 10%. The relative error gradually increases with replacement ratios, demonstrating the sensitivity of the model to variations in CRA composition. These findings proved the ability and efficiency of the proposed BN to predict compressive strength values by Bayesian updating.
- The ability to update data at both parent and child nodes and propagate evidence for the remaining nodes is very helpful in providing insights about a concrete mix proportion with a target compressive strength. The results show that mean values of w/c, R decrease from 0.54 to 0.528 and from 80.7% to 10.8%, respectively, when we target a compressive strength between the mean values of 21.5 MPa and 33.5 MPa. For these values, the mean of a/c increases from 2.54 to 3.25. Although the model predicted that great replacement ratios result in low compressive strengths, the application of Bayesian updating allows the identification of combinations of aggregate-to-cement and water-to-cement ratios that could lessen these impacts. The BN's capacity to account for specific constraints for the RAC formulations emphasizes its potential to reduce the environmental footprint of concrete construction by facilitating the effective utilization of recycled aggregates while attaining the intended mechanical qualities.
- The performance of the proposed BN can be improved by including other parent nodes, such as the quality of the original concrete, the cement content, the admixture kinds, or the curing conditions, which were not considered in this study. Furthermore, an optimal discretization of the considered random variables in the BN would

improve the accuracy of the predictions. Both improvements require a more comprehensive database.

- Most of the RAC in the dataset used to build the model have compressive strengths between 20 and 35 MPa. The framework application might be expanded to include high-performance RAC or different strength ranges. The model performance could be improved by adding more extensive datasets and utilizing sophisticated data categorization methods like K-means clustering, especially for larger replacement ratios when compressive strength variability is more noticeable.

Bayesian networks are effective tools for reasoning and modeling in the face of uncertainty. They depict variable dependencies using probabilistic relationships, enabling advanced inference and clear representation. Through the integration of data and expert knowledge, BNs enable the application of sparse data and effectively handle missing data through marginalization. However, BNs have some drawbacks. They can be computationally costly, particularly for large networks where the conditional probability table grows too big and accurate inference becomes challenging as the number of nodes and states for each node increase. Relying on expert input might result in bias, and BNs require accurate data and probability distributions to estimate parameters appropriately by Bayesian updating. Significant expertise and effort are required to build and validate network architectures, particularly when the number of nodes is very large. Furthermore, the assumptions about conditional independence and a predetermined structure do not necessarily align with practical situations. These difficulties show how important it is to thoroughly plan and implement BNs.

Author Contributions: Conceptualization, T.-D.N., R.C., P.-Y.M. and E.B.-A.; methodology, T.-D.N. and E.B.-A.; software, T.-D.N.; validation, R.C., P.-Y.M. and E.B.-A.; formal analysis, T.-D.N., R.C., P.-Y.M. and E.B.-A.; investigation, T.-D.N.; resources, E.B.-A.; data curation, T.-D.N.; writing—original draft preparation, T.-D.N.; writing—review and editing, R.C., P.-Y.M. and E.B.-A.; visualization, T.-D.N.; supervision, R.C., P.-Y.M. and E.B.-A.; funding acquisition, E.B.-A. All authors have read and agreed to the published version of the manuscript.

Funding: We acknowledge the French Ministry for Europe and Foreign Affairs and Campus France for funding the PhD scholarship of T.-D. Nguyen.

Data Availability Statement: Data used in this paper is provided in Appendices A and B.

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

Appendix A

Table A1. Dataset for the BN analysis.

w/c	*a/c*	*R*	f_c	Ref.
0.5	2.8	20	30.2	[31]
0.7	3.3	50	27.7	
0.7	2.2	100	20.4	
0.65	2.3	100	22.1	
0.5	3.2	0	32.7	
0.5	2.6	100	25.1	
0.54	3.4	0	32.8	
0.5	2.8	30	32.6	

Table A1. *Cont.*

w/c	*a/c*	*R*	f_c	Ref.
0.5	2.8	50	30.4	[30]
0.54	3.2	0	23.5	
0.66	3	100	25.7	
0.66	3.2	80	28	
0.66	3	100	25.1	
0.66	3.1	85	27.5	[32]
0.66	2.9	90	26.1	
0.66	2.9	70	27.4	
0.66	2.9	70	27.7	
0.66	2.8	50	25	
0.54	3.1	0	30.8	
0.55	3.2	50	27.5	
0.6	3.3	50	26.6	
0.6	3.3	50	25.7	[29]
0.5	3.5	0	28.3	
0.5	3.5	20	27.2	
0.5	3.5	40	26.5	
0.5	3.5	60	25.4	
0.5	3.2	20	26.4	
0.5	3.1	40	25.9	
0.65	3.3	0	28.2	
0.65	3.4	0	28.9	
0.66	3.3	20	28.7	
0.68	3.1	50	24.4	
0.68	2.8	100	23.1	
0.65	3.4	0	27.9	
0.65	3.3	20	26.7	
0.65	3.1	50	26.4	
0.65	2.8	100	22.1	
0.55	2.9	100	22.3	
0.65	3.3	100	24.8	
0.6	3.1	0	30	
0.6	3	25	26.7	
0.6	2.5	50	21.5	
0.6	2.5	75	21.4	
0.6	2.4	100	20	
0.66	2.9	0	27	
0.66	2.8	30	24	
0.55	2.8	0	26	
0.55	2.7	30	25	
0.51	2.6	50	24	

Table A1. *Cont.*

w/c	a/c	R	f_c	Ref.
0.5	2.7	0	31	
0.5	2.6	30	25	
0.58	3.1	0	32.8	
0.67	2.8	0	23.6	
0.68	3	20	26.5	
0.67	3	50	21	
0.7	2.3	100	20	
0.58	2.8	80	23.5	
0.51	2.3	20	31.4	
0.61	2.3	60	26	
0.62	2.3	50	25.5	
0.5	2.3	10	31.4	
0.56	2.3	20	28.2	
0.52	2.2	0	29.9	
0.5	3.5	20	33	
0.5	3.2	50	29.1	
0.68	3.2	100	23.5	
0.68	3.4	100	24.2	
0.68	3.4	100	24.7	
0.5	3	30	27.6	

Appendix B

Materials and Methods

We used a cement CEM II/A-LL 42.5 R with a specific density of 3150 kg/m^3 and 5/20 mm CRA provided by a recycling platform in La Rochelle, France. Table A2 provides the physical properties of the aggregate used in the study. Natural sand, which has a density of 2670 kg/m^3, was used in the experiment, while CRA and CNA have densities of 2870 kg/m^3 and 2354 kg/m^3, respectively. CRA has a water porosity of 7.95% while that of CNA is 0.5%.

We considered a cement content of 290 kg/m^3, and mixes maintained a 19 $\pm$ 1 mm slump range after an hour using a superplasticizer to regulate the mix proportion slump. All concretes were tested at the age of 28 days in terms of compressive strength. In total, we cast 18 concrete specimens (11 cm diameter and 22 cm height), including three samples for every test of mix proportion. Table A3 presents the data collected from our experiments.

Table A2. Physical properties of the aggregates used.

Physical Tests	Natural Sand	Coarse Natural Aggregate	Coarse Recycled Aggregate
Oven-dry particle density (kg/m^3)	2670	2870	2354
Water absorption (%)	1.2	0.5	7.95

Table A3. Dataset collected from authors' experiment.

w/c	*a/c*	*R*	f_c (MPa)
0.7	3.3	0	28.7
0.7	3.3	0	28
0.7	3.3	0	27.3
0.7	3.2	20	26.4
0.7	3.2	20	25.9
0.7	3.2	20	25.67
0.7	3.1	40	24.3
0.7	3.1	40	25
0.7	3.1	40	23.9
0.7	3	60	23
0.7	3	60	22.2
0.7	3	60	21.6
0.7	3	80	21.3
0.7	3	80	21.6
0.7	3	80	21.1
0.7	2.9	100	20
0.7	2.9	100	20.07
0.7	2.9	100	20.02

References

1. Bairagi, N.K.; Vidyadhara, H.S.; Ravande, K. Mix Design Procedure for Recycled Aggregate Concrete. *Constr. Build. Mater.* **1990**, *4*, 188–193. [CrossRef]
2. Wu, Z.; Yu, A.T.W.; Shen, L.; Liu, G. Quantifying Construction and Demolition Waste: An Analytical Review. *Waste Manag.* **2014**, *34*, 1683–1692. [CrossRef] [PubMed]
3. Limbachiya, M.; Meddah, M.S.; Ouchagour, Y. Use of Recycled Concrete Aggregate in Fly-Ash Concrete. *Constr. Build. Mater.* **2011**, *27*, 439–449. [CrossRef]
4. Han, T.; Siddique, A.; Khayat, K.; Huang, J.; Kumar, A. An Ensemble Machine Learning Approach for Prediction and Optimization of Modulus of Elasticity of Recycled Aggregate Concrete. *Constr. Build. Mater.* **2020**, *244*, 118271. [CrossRef]
5. Zhu, P.; Hao, Y.; Liu, H.; Wang, X.; Gu, L. Durability Evaluation of Recycled Aggregate Concrete in a Complex Environment. *J. Clean. Prod.* **2020**, *273*, 122569. [CrossRef]
6. Deng, Z.; Huang, H.; Ye, B.; Xiang, P.; Li, C. Mechanical Performance of RAC under True-Triaxial Compression after High Temperatures. *J. Mater. Civ. Eng.* **2020**, *32*, 04020194. [CrossRef]
7. Kong, D.; Lei, T.; Zheng, J.; Ma, C.; Jiang, J.; Jiang, J. Effect and Mechanism of Surface-Coating Pozzalanics Materials around Aggregate on Properties and ITZ Microstructure of Recycled Aggregate Concrete. *Constr. Build. Mater.* **2010**, *24*, 701–708. [CrossRef]
8. Kou, S.C.; Poon, C.S. Properties of Self-Compacting Concrete Prepared with Coarse and Fine Recycled Concrete Aggregates. *Cem. Concr. Compos.* **2009**, *31*, 622–627. [CrossRef]
9. Andreu, G.; Miren, E. Experimental Analysis of Properties of High Performance Recycled Aggregate Concrete. *Constr. Build. Mater.* **2014**, *52*, 227–235. [CrossRef]
10. Singh, R.; Nayak, D.; Pandey, A.; Kumar, R.; Kumar, V. Effects of Recycled Fine Aggregates on Properties of Concrete Containing Natural or Recycled Coarse Aggregates: A Comparative Study. *J. Build. Eng.* **2022**, *45*, 103442. [CrossRef]
11. Soares, D.; de Brito, J.; Ferreira, J.; Pacheco, J. Use of Coarse Recycled Aggregates from Precast Concrete Rejects: Mechanical and Durability Performance. *Constr. Build. Mater.* **2014**, *71*, 263–272. [CrossRef]
12. Dantas, A.T.A.; Batista Leite, M.; de Jesus Nagahama, K. Prediction of Compressive Strength of Concrete Containing Construction and Demolition Waste Using Artificial Neural Networks. *Constr. Build. Mater.* **2013**, *38*, 717–722. [CrossRef]

13. Duan, Z.H.; Kou, S.C.; Poon, C.S. Prediction of Compressive Strength of Recycled Aggregate Concrete Using Artificial Neural Networks. *Constr. Build. Mater.* **2013**, *40*, 1200–1206. [CrossRef]
14. Deshpande, N.; Londhe, S.; Kulkarni, S. Modeling Compressive Strength of Recycled Aggregate Concrete by Artificial Neural Network, Model Tree and Non-Linear Regression. *Int. J. Sustain. Built Environ.* **2014**, *3*, 187–198. [CrossRef]
15. Dabiri, H.; Kioumarsi, M.; Kheyroddin, A.; Kandiri, A.; Sartipi, F. Compressive Strength of Concrete with Recycled Aggregate; a Machine Learning-Based Evaluation. *Clean. Mater.* **2022**, *3*, 100044. [CrossRef]
16. Nguyen, T.-D.; Cherif, R.; Mahieux, P.-Y.; Lux, J.; Aït-Mokhtar, A.; Bastidas-Arteaga, E. Artificial Intelligence Algorithms for Prediction and Sensitivity Analysis of Mechanical Properties of Recycled Aggregate Concrete: A Review. *J. Build. Eng.* **2023**, *66*, 105929. [CrossRef]
17. Kabir, G.; Sumi, R.S.; Sadiq, R.; Tesfamariam, S. Performance Evaluation of Employees Using Bayesian Belief Network Model. *Int. J. Manag. Sci. Eng. Manag.* **2018**, *13*, 91–99. [CrossRef]
18. Tesfamariam, S.; Bastidas-Arteaga, E.; Lounis, Z. Seismic Retrofit Screening of Existing Highway Bridges with Consideration of Chloride-Induced Deterioration: A Bayesian Belief Network Model. *Front. Built Environ.* **2018**, *4*, 67. [CrossRef]
19. Kabir, G.; Demissie, G.; Sadiq, R.; Tesfamariam, S. Integrating Failure Prediction Models for Water Mains: Bayesian Belief Network Based Data Fusion. *Knowl.-Based Syst.* **2015**, *85*, 159–169. [CrossRef]
20. Al-Taai, S.R.; Azize, N.M.; Thoeny, Z.A.; Imran, H.; Bernardo, L.F.A.; Al-Khafaji, Z. XGBoost Prediction Model Optimized with Bayesian for the Compressive Strength of Eco-Friendly Concrete Containing Ground Granulated Blast Furnace Slag and Recycled Coarse Aggregate. *Appl. Sci.* **2023**, *13*, 8889. [CrossRef]
21. Caspeele, R.; Taerwe, L. Numerical Bayesian Updating of Prior Distributions for Concrete Strength Properties Considering Conformity Control. *Adv. Concr. Constr.* **2013**, *1*, 85–102. [CrossRef]
22. Abebe, Y.; Tesfamariam, S. Storm Sewer Pipe Renewal Planning Considering Deterioration, Climate Change, and Urbanization: A Dynamic Bayesian Network and GIS Framework. *Sustain. Resilient Infrastruct.* **2020**, *8*, 70–85. [CrossRef]
23. Franchin, P.; Lupoi, A.; Noto, F.; Tesfamariam, S. Seismic Fragility of Reinforced Concrete Girder Bridges Using Bayesian Belief Network: BBN Model of Seismic Fragility of RC Girder Bridges. *Earthq. Eng. Struct. Dyn.* **2016**, *45*, 29–44. [CrossRef]
24. Schoefs, F.; Tran, T.-B. Reliability Updating of Offshore Structures Subjected to Marine Growth. *Energies* **2022**, *15*, 414. [CrossRef]
25. Kabir, G.; Sadiq, R.; Tesfamariam, S. A Fuzzy Bayesian Belief Network for Safety Assessment of Oil and Gas Pipelines. *Struct. Infrastruct. Eng.* **2016**, *12*, 874–889. [CrossRef]
26. Tran, T.-B.; Bastidas-Arteaga, E.; Schoefs, F.; Bonnet, S. A Bayesian Network Framework for Statistical Characterisation of Model Parameters from Accelerated Tests: Application to Chloride Ingress into Concrete. *Struct. Infrastruct. Eng.* **2018**, *14*, 580–593. [CrossRef]
27. Niedermayer, D. An Introduction to Bayesian Networks and Their Contemporary Applications. In *Innovations in Bayesian Networks*; Holmes, D.E., Jain, L.C., Eds.; Studies in Computational Intelligence; Springer: Berlin/Heidelberg, Germany, 2008; Volume 156, pp. 117–130. ISBN 978-3-540-85065-6.
28. Guo, H.; Dong, Y.; Bastidas-Arteaga, E. Mixed Bayesian Network for Reliability Assessment of RC Structures Subjected to Environmental Actions. *Struct. Saf.* **2024**, *106*, 102392. [CrossRef]
29. Khan, K.; Ahmad, W.; Amin, M.N.; Aslam, F.; Ahmad, A.; Al-Faiad, M.A. Comparison of Prediction Models Based on Machine Learning for the Compressive Strength Estimation of Recycled Aggregate Concrete. *Materials* **2022**, *15*, 3430. [CrossRef]
30. Duan, J.; Asteris, P.G.; Nguyen, H.; Bui, X.-N.; Moayedi, H. A Novel Artificial Intelligence Technique to Predict Compressive Strength of Recycled Aggregate Concrete Using ICA-XGBoost Model. *Eng. Comput.* **2021**, *37*, 3329–3346. [CrossRef]
31. Suescum-Morales, D.; Salas-Morera, L.; Jiménez, J.R.; García-Hernández, L. A Novel Artificial Neural Network to Predict Compressive Strength of Recycled Aggregate Concrete. *Appl. Sci.* **2021**, *11*, 11077. [CrossRef]
32. Quan Tran, V.; Quoc Dang, V.; Si Ho, L. Evaluating Compressive Strength of Concrete Made with Recycled Concrete Aggregates Using Machine Learning Approach. *Constr. Build. Mater.* **2022**, *323*, 126578. [CrossRef]
33. Sosa, M.E.; Villagrán Zaccardi, Y.A.; Zega, C.J. A Critical Review of the Resulting Effective Water-to-Cement Ratio of Fine Recycled Aggregate Concrete. *Constr. Build. Mater.* **2021**, *313*, 125536. [CrossRef]
34. Güneyisi, E. Axial Compressive Strength of Square and Rectangular CFST Columns Using Recycled Aggregate Concrete with Low to High Recycled Aggregate Replacement Ratios. *Constr. Build. Mater.* **2023**, *367*, 130319. [CrossRef]
35. Poon, C.S.; Lam, C.S. The Effect of Aggregate-to-Cement Ratio and Types of Aggregates on the Properties of Pre-Cast Concrete Blocks. *Cem. Concr. Compos.* **2008**, *30*, 283–289. [CrossRef]
36. Nili, M.; Sasanipour, H.; Aslani, F. The Effect of Fine and Coarse Recycled Aggregates on Fresh and Mechanical Properties of Self-Compacting Concrete. *Materials* **2019**, *12*, 1120. [CrossRef]
37. Verian, K.P.; Ashraf, W.; Cao, Y. Properties of Recycled Concrete Aggregate and Their Influence in New Concrete Production. *Resour. Conserv. Recycl.* **2018**, *133*, 30–49. [CrossRef]
38. Patil, S.V.; Balakrishna Rao, K.; Nayak, G. Prediction of Recycled Coarse Aggregate Concrete Mechanical Properties Using Multiple Linear Regression and Artificial Neural Network. *JEDT* **2023**, *21*, 1690–1709. [CrossRef]

39. Kim, J. Influence of Quality of Recycled Aggregates on the Mechanical Properties of Recycled Aggregate Concretes: An Overview. *Constr. Build. Mater.* **2022**, *328*, 127071. [CrossRef]
40. Kabir, G.; Tesfamariam, S.; Loeppky, J.; Sadiq, R. Predicting Water Main Failures: A Bayesian Model Updating Approach. *Knowl.-Based Syst.* **2016**, *110*, 144–156. [CrossRef]
41. Normand, S.-U.; Tritchler, D. Parameter Updating in a Bayes Network. *J. Am. Stat. Assoc.* **1992**, *87*, 1109–1115. [CrossRef]
42. Demissie, G.; Tesfamariam, S.; Sadiq, R. Prediction of Soil Corrosivity Index: A Bayesian Belief Network Approach. In Proceedings of the 12th International Conference on Applications of Statistics and Probability in Civil Engineering, Vancouver, BC, Canada, 12–15 July 2015. [CrossRef]
43. Imounga, H.M.; Bastidas-Arteaga, E.; Rostand, M.P.; Ekomy Ango, S. Bayesian Updating of a Cracking Model for Reinforced Concrete Structures Subjected to Static and Cyclic Loadings. In *Fracture, Fatigue, Failure and Damage Evolution, Volume 3*; Xia, S., Beese, A., Berke, R.B., Eds.; Conference Proceedings of the Society for Experimental Mechanics Series; Springer International Publishing: Cham, Switzerland, 2021; pp. 45–49. ISBN 978-3-030-60958-0.
44. Alexandridou, C.; Angelopoulos, G.N.; Coutelieris, F.A. Mechanical and Durability Performance of Concrete Produced with Recycled Aggregates from Greek Construction and Demolition Waste Plants. *J. Clean. Prod.* **2018**, *176*, 745–757. [CrossRef]
45. McNeil, K.; Kang, T.H.-K. Recycled Concrete Aggregates: A Review. *Int. J. Concr. Struct. Mater.* **2013**, *7*, 61–69. [CrossRef]
46. Akbarnezhad, A.; Ong, K.C.G.; Tam, C.T.; Zhang, M.H. Effects of the Parent Concrete Properties and Crushing Procedure on the Properties of Coarse Recycled Concrete Aggregates. *J. Mater. Civ. Eng.* **2013**, *25*, 1795–1802. [CrossRef]
47. Tran, T.-B.; Bastidas-Arteaga, E.; Aoues, Y. A Dynamic Bayesian Network Framework for Spatial Deterioration Modelling and Reliability Updating of Timber Structures Subjected to Decay. *Eng. Struct.* **2020**, *209*, 110301. [CrossRef]
48. Liao, Q.-Z.; Xue, L.; Lei, G.; Liu, X.; Sun, S.-Y.; Patil, S. Statistical Prediction of Waterflooding Performance by K-Means Clustering and Empirical Modeling. *Pet. Sci.* **2022**, *19*, 1139–1152. [CrossRef]