Bayesian Networks Prediction of Compressive Strength of Recycled Aggregate Concrete

Tien-Dung NGUYEN[*,1,2], Rachid CHERIF[1], Pierre-Yves MAHIEUX[1], Emilio BASTIDAS-ARTEAGA[1]

[1] *LaSIE UMR CNRS 7356, University of La Rochelle, Avenue Michel Crépeau, 17042 La Rochelle Cedex 1, France,* tien-dung.nguyen@univ-lr.fr, rachid.cherif@univ-lr.fr, pierre-yves.mahieux@univ-lr.fr, ebastida@univ-lr.fr

[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, Viet Nam,* ntdung@dut.udn.vn

ABSTRACT Many studies proposed machine learning approaches for prediction models analysing the impact factors on recycled aggregate concrete (RAC) compressive strength. However, most machine learning algorithms require a large dataset size for the model's generalisation capability. Few studies have used Bayesian Networks (BNs) based probabilistic inference techniques towards this aim. This paper uses BNs to predict the compressive strength of recycled aggregate concrete. The BNs approach utilised available data of three input parameters: water-to-cement ratio, aggregate-to-cement ratio, and recycled aggregate replacement ratio to compute the output's prior and posterior probability of RAC's compressive strength. The results highlight the potential applicability of BNs in predicting the compressive strength of RAC.

Keywords Recycled aggregate concrete, Bayesian networks, Compressive strength.

I. INTRODUCTION

To use fewer raw materials, recycled aggregate (RA) offers a good substitute for natural aggregate (NA). When structures like bridges, buildings, roads, and trains are destroyed, as well as occasionally when disasters like floods, wars, and earthquakes occur, the resulting debris can be utilised to create RA. However, the mortar of the RA that has been bonded to it has decreased in quality because of increased porosity and water absorption (Zhu et al., 2020). This results in the quality of recycled aggregate concrete (RAC) being inferior to those of the original concrete, most obviously in its compressive strength (Eguchi et al., 2007). Many factors, such as the mixed design elements or the characteristics of the recycled aggregate (RA), have an impact on the compressive strength of RAC (Abdollahzadeh et al., 2016).

Moreover, its characteristics are also impacted by the interfacial transition zones (ITZs) (Poon et al., 2004). Nearly all investigations found that the compressive strength of RAC reduced as replacement increased (Alexandridou et al., 2018; Faella et al., 2016), while some have shown the opposite, particularly as later age (Basheer et al., 2005; Gholampour and Ozbakkaloglu, 2020) because RA particles have non-hydrated cement adhered to their surfaces. These residues react with water to increase the late strength development, resulting in a complicated relationship between mix ratio, RA physical characteristics, and compressive strength.

Machine learning techniques have been created to forecast the compressive strength of RAC. They include Support Vector Machines (SVM), Radial Basis Functions (RBF), Neuro-Fuzzy Inference Systems (ANFIS), Genetic Programming (GP), and Artificial Neural Networks (ANNs). Dantas et al. (Dantas et al., 2013) presented the ANNs for predicting the compressive strength of RAC at the age of 3, 7, 28 and 91 days using 24 input parameters. Deshpande et al. (Deshpande et al., 2014) modelled compressive strength of RAC by ANNs, Model Tree (MT) and Non-linear Regression (NTR), which utilising 9 mandatory input parameters and 5 input non-dimensional parameters. Several researchers have examined the impacts of combining various ratios of the components that affect concrete's compressive strength (CS) and their estimation using all machine learning approaches (Nguyen et al., 2023). However, due to the uncertain nature of the materials and the many parameters involved in the problem, compressive strength prediction of RAC is still a challenging task.

All machine learning algorithms need a sizable dataset to generalise the models (Bui et al., 2018). Due to time constraints, the cost of the tests, and the implementation challenges in practice, learning can only be conducted with a limited number of sites in actual practice. As a result, other statistical and/or probabilistic methodologies could be used to quantify uncertainty while making the most effective use of the currently accessible data. Therefore, it makes sense to approach this topic using a Bayesian approach. To estimate the compressive strength of RAC, Bayesian networks (BNs) were used in the present study. The NETICA software package was used to create and run the BNs model. The model phases' prior and posterior probabilities were calculated using test data from numerous published publications. The Python programming language is used to create the BBNs-based updating prediction model.

II. BAYESIAN NETWORKS AND ITS APPLICATION

A. Bayesian Networks (BNs)

A BNs is often a particular kind of graphical model represented as a directed acyclic graph (DAG). Nodes in a DAG represent variables or states and are graphical representations of things that happen in the real world. Drawing an edge between two nodes indicates a causal relationship between them. A directed edge will lead from the cause factor towards the effect variable if there is a causal connection between the nodes. There is a Probability Density Function (PDF) for each parameter in the DAG, and the concept and scale of the PDF are determined by the edges connecting to the variables.

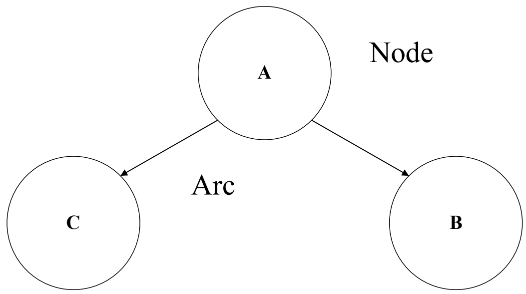

FIGURE 1. A simple Bayesian Network

A straightforward BN is depicted in figure 1; it has three nodes that stand in for the three considered parameters A, B, and C, with B and C being children of the parent node A. Conditional probability distributions for the child nodes are dependent on the parent node. A marginal probability distribution exists for the parent node. Given the prior and the conditional probabilities p(A) and p(B|A), Bayes' rule enables the computation of the posterior probability p(A| B):

$$P(A/B)=\frac{P(B/A)P(A)}{P(B)} \tag{1}$$

B. Application to predict compressive strength of RAC

When employing AI techniques to predict the mechanical properties of RAC, the input data, chosen model, and hyperparameters are all crucial components in reaching the necessary accuracy (Ahmad et al., 2021). Most studies in the literature on the topic included crucial input qualities like cement content, water content, and W/C, as well as the testing age. In contrast, a few studies utilised physical traits like maximum aggregate size and water absorption of aggregate in prediction models (Nguyen et al., 2023). All the machine learning approaches utilised input variables related to mixing design, physical properties, and age of recycled aggregate concrete in the available literature. In this study, BNs also choose suitable input parameters to predict RAC compressive strength (Nunez et al., 2021). Figure 2 depicts the BNs architecture for predicting the compressive strength of RAC considered in this study. It can be observed that the input variables are aggregate–cement ratio (A/C), water–cement ratio (W/C), replacement ratio (R), and compressive strength as output variable. The NETICA package is used to create the BN-based prediction model.

Before defining the architecture of the BN, it is important to analyse if correlation between the considered variables (nodes) exists. Figure 3 shows the Pearson correlation coefficients that measure the linear connection between two variables or to quantify this relationship. It ranges from -1 to 1. A negative linear correlation between two variables is shown by a value of -1. No linear association between two variables is indicated by a 0, and a complete positive linear relationship between two variables is indicated by a value of 1. As shown in figure 3, the values vary from -0.29 to 0.4, resulting in W/C, A/C, and R being three independent input variables.

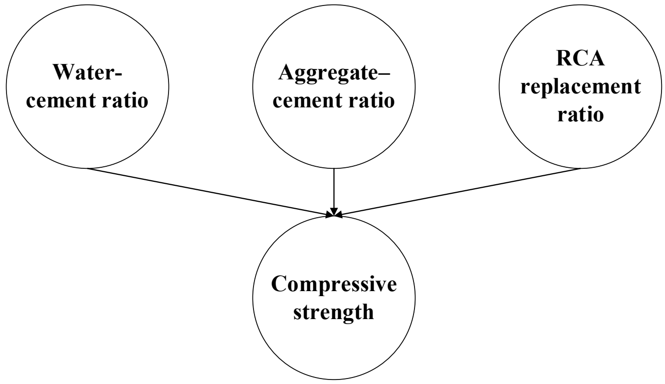

FIGURE 2. **Proposed BNs for predicting compressive strength of RAC**

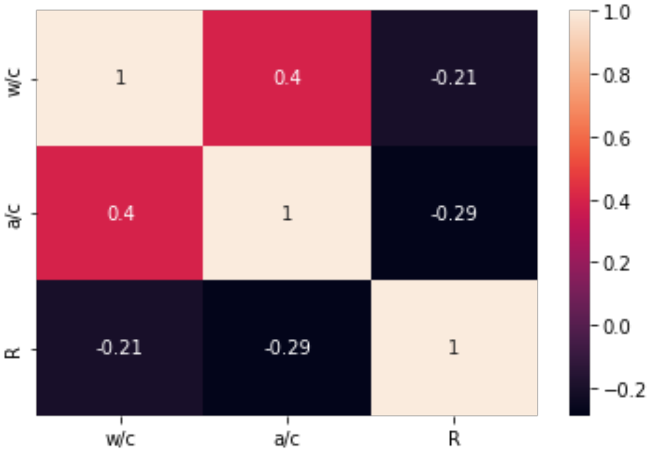

FIGURE 3. **Correlation matrix of input parameters**

C. Assessment of prior and conditional probabilities

Based on data from numerous previously published works (Suescum-Morales et al., 2021), all parent nodes are continuous and partitioned into various states within established constraints at the beginning of this phase. The intervals (upper and lower bounds) for each parameter should contain the parameter's theoretically/physically permissible values. These ranges can be determined using pre-existing databases, similar research cases, or expert knowledge. The boundaries of each node are given in table 1.

TABLE 1. **The discretisation of nodes for the model**

Nodes	Number of states	Boundaries
W/C	5	[0.35, 0.6]
A/C	3	[2, 3.2]
R	4	[0, 100]
Compressive strength	5	[25, 50]

Next, the prior probability was calculated directly from the relative frequency of the database. The details of the discretisation of the different nodes are summarised in Table 1 and the relative frequencies are shown in figure 4. The selected number of nodes and states results in a conditional probability table having a large number of rows. The minimum values for output parameters is 300 to fulfil the CPT.

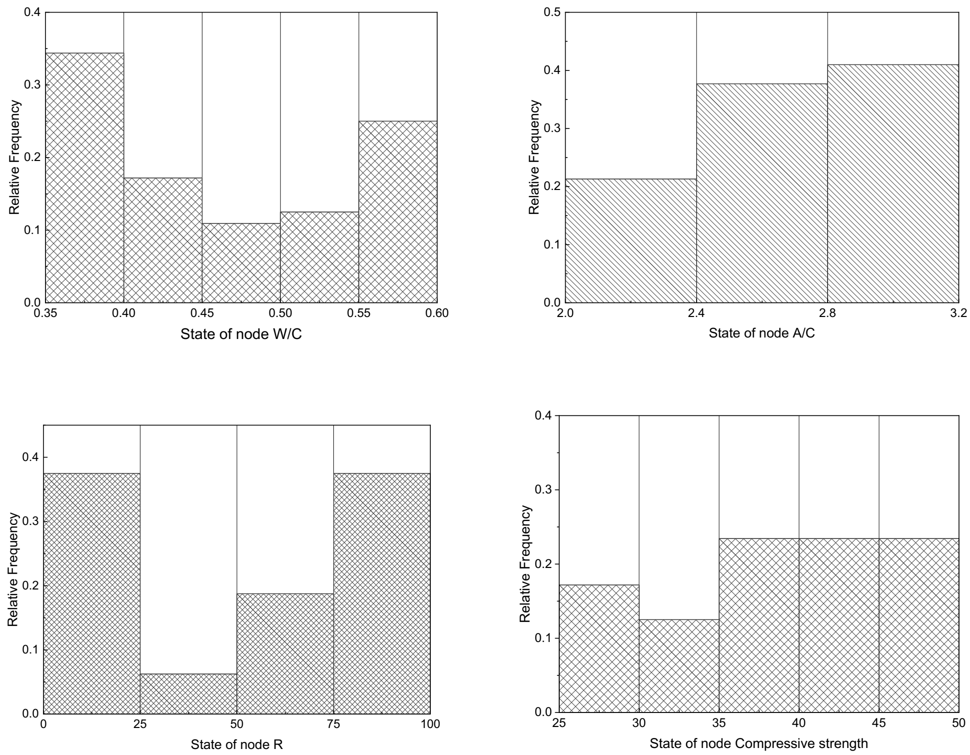

FIGURE 4. Histogram of all nodes in BNs

Figure 4 also provides the parent and child nodes' prior probabilities and states of all nodes in the BNs. The table summarises the discretisation details and a priori knowledge of the various nodes. We can see from figure 4 that the water-cement ratio values in prior distribution are concentrated mainly in two states, 0.35-0.4 and 0.55-0.6; these are 34.3% and 25.1%, respectively. Regarding the aggregate-cement ratio, values in the distribution focus on 40.9% at state 2.8-3.2. The replacement ratio values in the distribution emphasise 37.5% and 37.5% at two states, 0-25 and 75-100, respectively. The values of output prior distribution mainly concentrate on three states, 35-40 MPa, 40-45 MPa and 45-50 MPa; these are 23.4%,23.4% and 23.4%, respectively. Even if the prior probabilities of the parent nodes have different shapes, the distribution of the compressive strength is almost uniform. This highlights the non-linear relations between parent and child nodes that should be accurately represented by the BNs.

The relationship in a BN is quantified by a set of tabulated conditional probabilities that display one probability distribution for each possible combination of parent values. This study's conditional probability table (CPT) sizes would be (5x3x4) rows and 5 columns, respectively. By calculating the frequency with which the child node's value occurs when the parent node's value is given, the conditional probability values of the child node can be calculated. After that, the CPT values were reviewed and adjusted as necessary to satisfy engineering judgment. Table 2 shows the conditional probability table of the BNs model. These prior probabilities and conditional probabilities have been used for Bayesian inference.

TABLE 2. Conditional probability table of the BNs model

No. of rows	Node and states			States of node Compressive strength and Conditional probability				
	W/C	**A/C**	**R**	25-30	30-35	35-40	40-45	45-50
1	0.35-0.4	2-2.4	0-25	0	0.1667	0	0	0.8333
…	…	…	…	…	…	…	…	…
13	0.4-0.45	2-2.4	0-25	0	1	0	0	0
…	…	…	…	…	…	…	…	…
25	0.45-0.5	2-2.4	0-25	0.2	0.2	0.2	0.2	0.2
…	…	…	…	…	…	…	…	…
37	0.5-0.55	2-2.4	0-25	0.2	0.2	0.2	0.2	0.2
…	…	…	…	…	…	…	…	…
49	0.55-0.6	2-2.4	0-25	0.2	0.2	0.2	0.2	0.2
…	…	…	…	…	…	…	…	…
60	0.55-0.6	2.8-3.2	75-100	0	0	0.33	0.33	0.34

D. Assessment of posterior probabilities

The output (child) node, which has a marginal probability distribution, can be determined by the following equation:

$$P(CS) = \sum_{\frac{W}{C},\frac{A}{C},R} P((CS)\left|\frac{W}{C},\frac{A}{C},R)P\left(\frac{W}{C},\frac{A}{C},R\right) \; with \; P\left(\frac{W}{C},\frac{A}{C},R\right) = P(\frac{W}{C})P(\frac{A}{C})P(R) \right. \quad (2)$$

NETICA software package is useful to implement this BN and to compute the posterior probability of compressive strength results. The posterior probabilities of an output parameter are shown in figure 5. The results for the child node show that the states 25–30, 35–40 and 40-45 had the highest posterior probabilities of the output parameter, at 21.1%, 25.7%, and 21.1, respectively. The posterior results increase compared to the prior probability for the states 25-30 and 35-40, whereas the posterior probability of states 40-45 decreases by 2.3%.

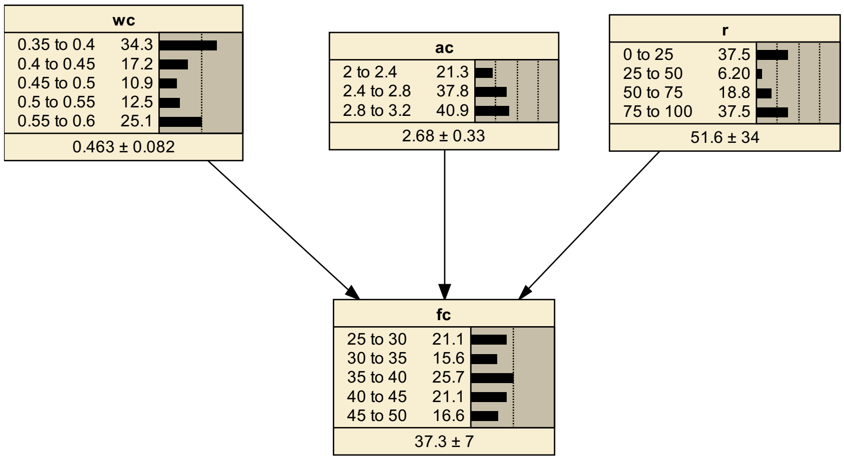

FIGURE 5. Posterior probability of all nodes in BNs

E. Assessment of belief or probability updating

This BBNs-based evaluation approach has many advantages, including changing CPTs and node beliefs whenever new knowledge or data becomes available. The performance of the BBNs model can be adapted to new information/data for its nodes or parameters, allowing for the construction of a model by using the best data/information available and later incorporation of evidence. The results allow to use the best data currently available as evidence of observational data or variable correlations for distribution updating.

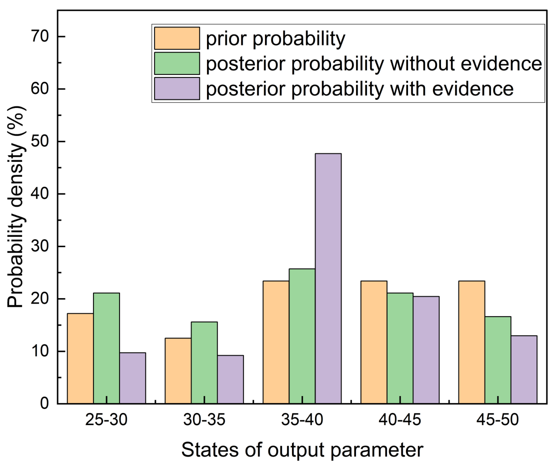

FIGURE 6. Histogram of the output node (compressive strength)

This study used data from a literature review (Casuccio et al., 2008; Kou et al., 2007) to evaluate beliefs or update probabilities to assess the effects of new information or evidence. In the approach, the output data was utilised to update the model. Because the output node had relationship with all input nodes, the posterior probability of all nodes was updated with observational data. It can be observed from figure 6 that the posterior probability values with evidence of output node have increased by 22% at the state 35-40. At other states, the posterior probability values decreased slightly at the states 40–45 and 45-50, these results are slower than without evidence by 0.5% and 2.5%, respectively. Table 3 summarises the model's posterior probability values of all parent nodes. It can be observed that the posterior probability of all input nodes has changed with new information. More specifically, the posterior probability of node W/C has maximum values for two primary states, 0.35-0.4 and 0.55-0.6, by 33.29% and 29.96% compared to the prior probability at 34.3% and 25.1%, respectively. The posterior probability of node A/C had a maximum value at the state 2.4-2.8 of 42.38% compared to the prior value at the state 2.8-3.2 of 40.9%. The posterior values of node replacement ratio focus on the states 0-25 and 75-100, these results are 33.24% and 34.32% compared to 37.5% and 37.5% of prior values, respectively.

TABLE 3. **The posterior probabilities of the parent nodes**

Parameters	States	Prior probability (%)	Posterior probability (%)
Water to cement ratio	0.35-0.4	34.3	33.29
	0.4-0.45	17.2	17.89
	0.45-0.5	10.9	8.71
	0.5-0.55	12.5	10.15
	0.55-0.6	25.1	29.96
Aggregate to cement ratio	2.0-2.4	21.3	20.48
	2.4-2.8	37.8	42.38
	2.8-3.2	40.9	37.14
Replacement ratio (%)	0-25	37.5	33.24
	25-50	6.2	7.76
	50-75	18.8	24.68
	75-100	37.5	34.32

The Bayesian model in this paper can predict compressive strength from the values of input parameters. In figure 7, we assume that the prior probability of the W/C node is 10%, 80%, and 10% for three states 0.35-0.4, 0.4-0.45, 0.45-0.5, respectively. The prior probability of the A/C node is 10%, 80%, and 10% for three states 2-2.4, 2.4-2.8, 2.8-3.2, respectively. The prior probability of the R node is 10%, 80%, and 10% for three states 0-25, 25-50, and 50-75, respectively. The Bayesian model shows that the probability of output has the highest value at the state 35-40 MPa.

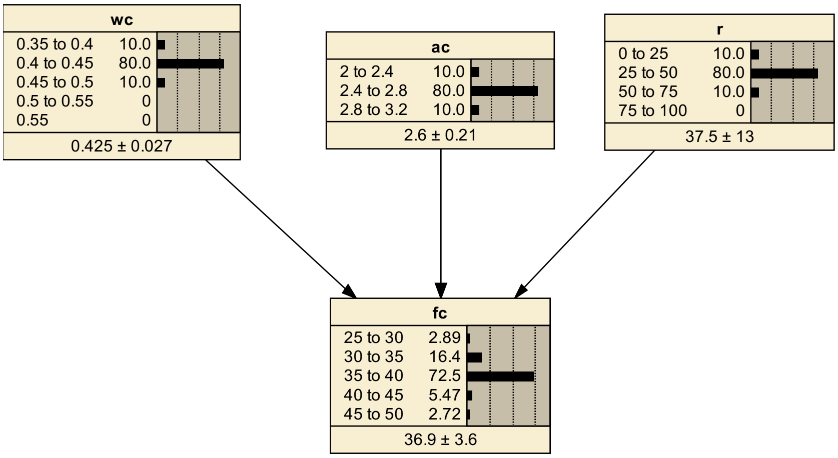

FIGURE 7. Posterior probability of all nodes in BNs

III. CONCLUSION

Nearly all machine learning techniques were used to forecast RAC's compressive strength but needed a larger amount of input data. The BNs model, however, offers numerous benefits that other approaches do not. BNs can quantify uncertainty while maximising the usage of the currently available knowledge. The proposed model, grounded in probability theory, can be a fresh approach to illustrate the data by a probability distribution (the prior and posterior probability of input and output information). BNs are a suitable method if there is inadequate data or information. This model can still be developed using expert knowledge or a medium data set and information. According to the findings in the study, the suggested BBNs can also update the model whenever new information/data becomes available. The application is the strength of this method compared to other hybrid forms. We can quickly update information without spending much time on a new model.

REFERENCES

Abdollahzadeh, G., Jahani, E., & Kashir, Z. (2016). Predicting of compressive strength of recycled aggregate concrete by genetic programming. Computers and Concrete, 18(2), 155–163. https://doi.org/10.12989/CAC.2016.18.2.155

Ahmad, A., Farooq, F., Ostrowski, K. A., Śliwa-Wieczorek, K., & Czarnecki, S. (2021). Application of Novel Machine Learning Techniques for Predicting the Surface Chloride Concentration in Concrete Containing Waste Material. Materials, 14(9), 2297. https://doi.org/10.3390/ma14092297

Alexandridou, C., Angelopoulos, G. N., & Coutelieris, F. A. (2018). Mechanical and durability performance of concrete produced with recycled aggregates from Greek construction and demolition waste plants. Journal of Cleaner Production, 176, 745–757. https://doi.org/10.1016/j.jclepro.2017.12.081

Basheer, L., Basheer, P. A. M., & Long, A. E. (2005). Influence of coarse aggregate on the permeation, durability and the microstructure characteristics of ordinary Portland cement concrete. Construction and Building Materials, 19(9), 682–690. https://doi.org/10.1016/j.conbuildmat.2005.02.022

Bui, D.-K., Nguyen, T., Chou, J.-S., Nguyen-Xuan, H., & Ngo, T. D. (2018). A modified firefly algorithm-artificial neural network expert system for predicting compressive and tensile strength of high-performance concrete. Construction and Building Materials, 180, 320–333. https://doi.org/10.1016/j.conbuildmat.2018.05.201

Dantas, A. T. A., Batista Leite, M., & de Jesus Nagahama, K. (2013). Prediction of compressive strength of concrete containing construction and demolition waste using artificial neural networks. Construction and Building Materials, 38, 717–722. https://doi.org/10.1016/j.conbuildmat.2012.09.026

Deshpande, N., Londhe, S., & Kulkarni, S. (2014). Modeling compressive strength of recycled aggregate concrete by Artificial Neural Network, Model Tree and Non-linear Regression. International Journal of Sustainable Built Environment, 3(2), 187–198. https://doi.org/10.1016/j.ijsbe.2014.12.002

Eguchi, K., Teranishi, K., Nakagome, A., Kishimoto, H., Shinozaki, K., & Narikawa, M. (2007). Application of recycled coarse aggregate by mixture to concrete construction. Construction and Building Materials, 21(7), 1542–1551. https://doi.org/10.1016/j.conbuildmat.2005.12.023

Faella, C., Lima, C., Martinelli, E., Pepe, M., & Realfonzo, R. (2016). Mechanical and durability performance of sustainable structural concretes: An experimental study. Cement and Concrete Composites, 71, 85–96. https://doi.org/10.1016/j.cemconcomp.2016.05.009

Gholampour, A., & Ozbakkaloglu, T. (2020). Long-term behavior of recycled aggregate concrete. In Advances in Construction and Demolition Waste Recycling (pp. 301–320). Elsevier. https://doi.org/10.1016/B978-0-12-819055-5.00015-2

Nguyen, T.-D., Cherif, R., Mahieux, P.-Y., Lux, J., Aït-Mokhtar, A., & Bastidas-Arteaga, E. (2023). Artificial intelligence algorithms for prediction and sensitivity analysis of mechanical properties of recycled aggregate concrete: A review. Journal of Building Engineering, 66, 105929. https://doi.org/10.1016/j.jobe.2023.105929

Nunez, I., Marani, A., Flah, M., & Nehdi, M. L. (2021). Estimating compressive strength of modern concrete mixtures using computational intelligence: A systematic review. Construction and Building Materials, 310, 125279. https://doi.org/10.1016/j.conbuildmat.2021.125279

Poon, C. S., Shui, Z. H., & Lam, L. (2004). Effect of microstructure of ITZ on compressive strength of concrete prepared with recycled aggregates. Construction and Building Materials, 18(6), 461–468. https://doi.org/10.1016/j.conbuildmat.2004.03.005

Suescum-Morales, D., Salas-Morera, L., Jiménez, J. R., & García-Hernández, L. (2021). A Novel Artificial Neural Network to Predict Compressive Strength of Recycled Aggregate Concrete. Applied Sciences, 11(22), 11077. https://doi.org/10.3390/app112211077

Zhu, P., Hao, Y., Liu, H., Wang, X., & Gu, L. (2020). Durability evaluation of recycled aggregate concrete in a complex environment. Journal of Cleaner Production, 273, 122569. https://doi.org/10.1016/j.jclepro.2020.122569