Significance of Eco-Cement Constituents to Its Mechanical Properties, by Machine Learning Algorithms

Natia R. Anastasi(✉) and Nikolaos P. Bakas

Civil Engineering Department, School of Architecture, Engineering, Land and Environmental Sciences, Neapolis University Pafos, 2 Danais Avenue, 8042 Paphos, Cyprus
natia.anastasi@nup.ac.cy

Abstract. A novel proposed material for the potential replacement of cement in some of its applications was evaluated. This new material labeled Eco-Cement comprised of Biomass from the dairy and poultry industries, Urea, Cement Kiln Dust, Rice Husk Ash, Sand, and Water and was manufactured in a range of weight ratios. In this work, a comprehensive analysis of the ingredients in varying weight percentages of the novel material was manufactured and the corresponding strength and strains of the material were studied. A variety of concrete pastes using amounts of sought ratios were produced, molded into blocks, and allowed to cure under laboratory conditions. Unconfined compression tests were performed using a deformation control compressive strength machine. The strength and strains were evaluated from the initial zero load step incrementally, until the failure of each specimen. The resulting database was analyzed by utilizing Linear Regression, Random Forests, and the Gradient Boosting machine learning methods. Extensive sensitivity analysis with the machine learning algorithms, reveal certain patterns, which were established with three different methods. Furthermore, we present the analysis of the corresponding literature with Bibliometric techniques.

Keywords: Cement · Biomass · Compressive Strength · Machine Learning

1 Introduction

The need to use unwanted side products or trash in the preparation of useful materials is a general drive of modern industrial research. This is elaborated in such processes as the use of recycled tires in road construction [7, 10, 12], recycled glass in the production of heat insulation products [4] or more intriguing applications, like the use of recycled bottles in the manufacturing of concrete [11, 15]. Concrete is one of the most abundant materials on earth. It is composed of cement, water, and aggregates. The manufacturing of cement is amongst the processes that are most severely detrimental to the environment. On the contrary, the material studied herein, Eco-Cement [1], is a composite material of Bacteria, Urea, Calcium Carbonate, Cement Kiln Dust (CKD), the hydraulic agent Rice Husk Ash (RHA), and Sand. It utilizes waste products of numerous industries such as the cement manufacturing industry, the dairy industry, the poultry growing industry making it an overall environmentally desired product (Fig. 1).

M. Papadaki et al. (Eds.): EMCIS 2022, LNBIP 464, pp. 33–39, 2023.
https://doi.org/10.1007/978-3-031-30694-5_3

2 Experimental Design

CKD, Cement Kiln Dust is a solid waste of the cement manufacturing industry. It is a super-fine grained solid that accumulates during the manufacturing of cement on the air pollution control filters. The physical and chemical characteristics of this material vary from plant to plant and depend heavily on the raw materials used in the manufacturing of Portland cement. It is highly alkaline and it is considered to be a mixture of primarily unreactive starting materials. Studies have been performed in order to investigate the use of CKD as a cementitious material in concrete and mortar. It was found that whilst substitution of certain percentages of cement with CKD did not have advert effects on the mechanical characteristics of the resulting material, with substitution above certain limits the effects on said properties were detrimental. The diminished mechanical abilities of CKD-containing concrete are due to the low levels of calcite, the calcium carbonate polymorph speculated to have the highest strength tolerances.

Precipitation of calcium carbonate is facilitated though a plethora of processes, amongst them biomineralization. Bacteria have shown to induce calcium carbonate precipitation through urea hydrolysis. Urease is a Nickel-containing enzyme that catalyzes the hydrolysis of urea into ammonia and carbon dioxide. At specifically tuned pH the hydrolysis products are ammonium and carbonate.

$$CO(NH_2)_2 + H_2O \rightarrow 2NH_4^+ + CO_3^{-2} \tag{1}$$

The produced carbonate ions precipitate in the presence of calcium ions as calcium carbonate crystals:

$$Ca^{+2} + CO_3^{-2} \rightarrow CaCO_3 \tag{2}$$

3 Linear Regression

We applied basic statistical analysis to the obtained dataset, and afterwards a machine learning investigation. We normalize each one of the independent and dependent variables v_i by subtracting its mean value μ_i and dividing by its standard deviation σ_i, and we get each normalized ones

$$v_i^n = (v_i - \mu_i)/\sigma_i \tag{3}$$

Initially, we checked the multiple correlations among the independent variables, and obtained a high correlation among Water and CKD (0.9646), and Water with Biomass (0.8532), as depicted in Fig. 2. This was due to the mixing procedure, with mixtures of high CKD, and Biomass needed more water to get to a solid-state. Accordingly, as the aim of the regression analysis is to assess the impact of each independent variable on the dependent, and not predictive modeling, we exclude water from the model. The new model has five variables, which are Urea, Biomass, CKD, RHA, and Sand. In order to check the multicollinearity, we perform linear regression, with each independent variable as response and the remaining independent variables as predictors, and compute the Variance Inflation Factor (VIF)

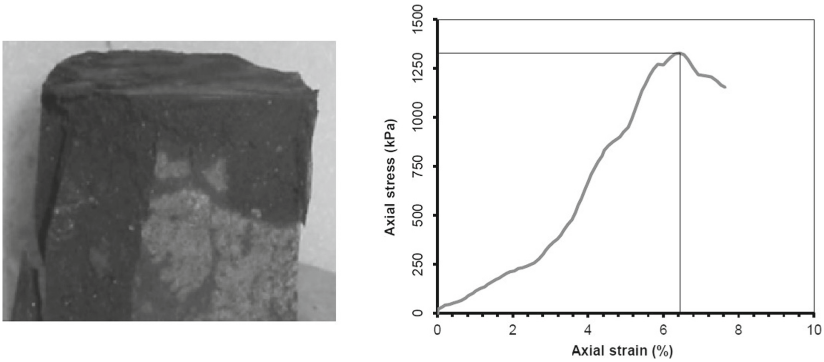

Fig. 1. Example strength test: (left) Photograph of the specimen after failure, (right) Stress-strain curve.

$$VIF_i = /(1 - R_i^2) \tag{4}$$

for each independent variable i. In the new model, we get the VIF values for each variable: Urea (2.261), Biomass (5.551), CKD (2.791), RHA (2.858), Sand (2.255). It is suggested that the independent variable has a VIF $<$ 10 [13], so we keep these five variables.

We use compressive strength as the dependent variable, and the regression analysis results are presented in Table 1, and Fig. 5. Interestingly, we obtained a low p-value for Biomass (0.0148), with high and positive weight (0.747) indicating Biomass as a significant factor for the compressive strength. Furthermore, the CKD exhibited a p-value of 0.0267, with a negative weight (−0.4792), indicating that the CKD has a negative effect on the compressive strength. This is because of the highly exothermic reaction of the hydration of CKD raises the temperature to levels above the tollerances of the enzyme, resulting in denaturing and less concentration of active enzyme participating in biocementation [6]. These numerical results are verified by the Machine Learning investigation in the following section.

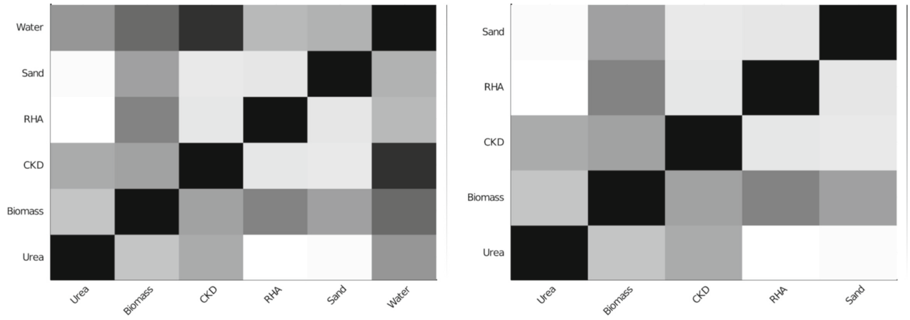

Fig. 2. Pearson Correlation Factors among the independent variables before (left) and after (right) the exclusion of Water due to multicollinearity.

4 Machine Learning Models

In order to quantify the significance of each ingredient to Eco-Cement's mechanical properties, a variety of regression analyses were performed. The analyses were accomplished utilizing MIT's Julia programming language [2] and custom codes written by the authors. The dependent variable was the compressive strength (Fig. 3) and the independent, the ingredients of Eco-Cement, divided with each specimen's volume. The dataset contained limited observations (N = 55), hence three different regression methods were utilized and compared: Multiple Linear Regression (MLR), Random Forests [3, 16], and Gradient Boosting [5, 8, 18]. Accordingly, a modified version of the Profile method [9, 14] is utilized, in order to investigate each variable's contribution to the dependent variable. In particular, each input variable varies within its given (raw) range while all the other input variables are kept constant in a certain value. This constant takes three discrete values: 25% Percentile, Median, and 75% Percentile. In Table 2, the accuracy metrics for each regression method are presented: the Pearson Correlation Coefficient (COR), the Mean Absolute Error (MAE), the Root Mean Squared Error (RMSE), the Mean Absolute Percentage Error (MAPE), the Maximum Absolute Percentage Error (MAXAPE), and the Alpha metric (Table 1).

Table 1. Regression analysis metrics for compressive strength

METHODS	COR	MAE	RMSE	MAPE	MAXAPE	Alpha
RandomForest	0.661	172.553	220.357	0.177	1.387	0.217
GradientBoosting	0.901	101.121	131.363	0.110	0.936	0.617
LinearRegression	0.471	177.965	236.693	0.181	1.516	0.222

The metrics for the three methods indicate low predictive performance. The reasons for this performance may be attributed to the relatively small sample database and a poor transcription of solution chemistry [17] to solid state.

Through detailed sensitivity analysis, however we were able to reveal some patterns on how each independent variable influences the dependent. In Fig. 3, the sensitivity curves for CKD are plotted. Linear Regression Exhibits a clear decreasing pattern, while Random Forests and Gradient Boosting a decreasing trend for low values of CKD and almost constant for the rest. This can be speculatively attributed to a more exothermic reaction at higher concentrations of CKD resulting in the possible denaturing on the enzyme and affecting the cementation process. Furthermore, in Fig. 4, for higher values of Urea we obtain lower values of compressive strength, however with a lower variation. Again, we can speculate that higher concentrations of urea, result in an increase of the pH hence causing denaturing of the enzyme and affecting biocementation.

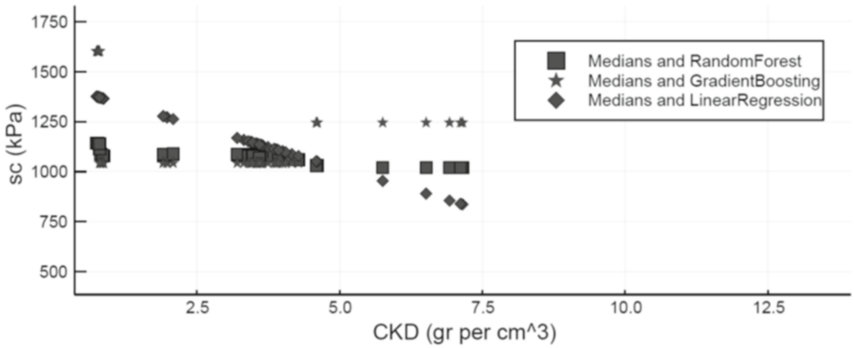

Fig. 3. Sensitivity curves for CKD.

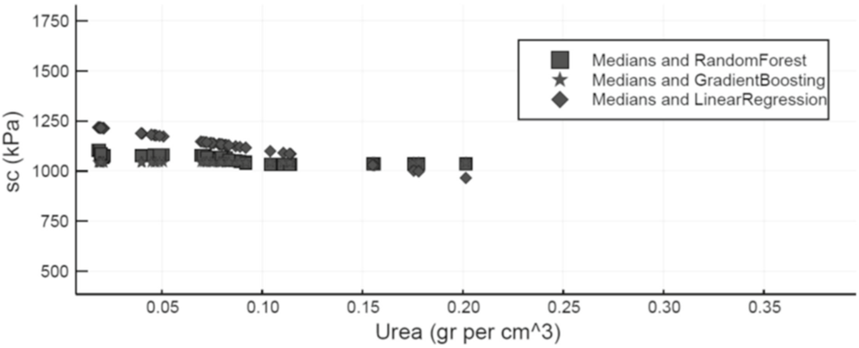

Fig. 4. Sensitivity curves for urea.

5 Conclusions

The composition of a new material using waste byproducts of various industries, achieves compressive strengths that fall within the range of a material used as mortar. An optimal recipe is not easily achieved as there is a non-straightforward behavior that relates to the compressive strength and the amounts of the components. Statistical analysis and machine learning algorithms extracted specific patterns of the importance of Eco-Cement ingredients, with respect to its compressive strength, especially for the Urea and CKD. The difficulty in obtaining a clear dependency of the compressive strength, based on its constituents, is a well-known problem for a variety of materials, as even if we repeat the same experiment with the same composition of a material, the resulting strength might diverge. Hence this study serves as an initial point of investigation for an ecofriendly material, which could be used as a mortar, with significant economic and environmental advantages.

Acknowledgements. This work was made under the frame of the ECOCE-MENT project (FP7 - Grant 282922).

References

1. Anastasi, N., et al.: Ecocement: A novel composite material for the construction industry. Identification of an Optimal Recipe Using Neural Networks (2016)
2. Bezanson, J., Edelman, A., Karpinski, S., Shah, V.B.: Julia: A fresh approach to numerical computing. SIAM Rev. **59**, 65–98 (2017)
3. Breiman, L.: Random forests. Mach. Learn. **45**, 5–32 (2001)
4. Breit, W., Schnell, J.: Monolithic external architectural lightweight concrete components providing highly efficient thermal insulation-an experimental building. Betonwerk und Fertigteil-Technik **80**(2), 128–130 (2014)
5. Chen, T., Guestrin, C.: XGBoost: A scalable tree boosting system. In: Proceedings of the 22nd ACM SIGKDD International Conference on Knowledge Discovery and Data Mining. pp. 785–794. KDD '16, ACM, New York, NY, USA (2016). https://doi.org/10.1145/2939672.2939785
6. Czapik, P., Zapała-Sławeta, J., Owsiak, Z., Stkepień, P.: Hydration of cement by- pass dust. Constr. Build. Mater. **231**, 117139 (2020)
7. Elchalakani, M.: High strength rubberized concrete containing silica fume for the construction of sustainable road side barriers. Structures **1**, 20–38 (2015). https://doi.org/10.1016/j.istruc.2014.06.001,citedBy43
8. Friedman, J.H.: Stochastic gradient boosting. Comput. Stat. Data Anal. **38**, 367–378 (2002)
9. Gevrey, M., Dimopoulos, I., Lek, S.: Review and comparison of methods to study the contribution of variables in artificial neural network models (2003). https://doi.org/10.1016/S0304-3800(02)00257-0
10. Huang, Y., Bird, R.N., Heidrich, O.: A review of the use of recycled solid waste materials in asphalt pavements. Resour. Conserv. Recycl. **52**, 58–73 (2007). https://doi.org/10.1016/j.resconrec.2007.02.002,citedBy284
11. Irwan, J.M., et al.: The mechanical properties of pet fiber reinforced concrete from recycled bottle wastes. Adv. Mat. Res. **795**, 347–351 (2013). https://doi.org/10.4028/www.scientific.net/AMR.795.347, cited By 14; Conference of 2nd International Conference on Sustainable Materials, ICoSM 2013; Conference Date: 26 March 2013 Through 27 March 2013; Conference Code:100411
12. Myhre, M., Mackillop, D.A.: Rubber recycling. Rubber Chem. Technol. **75**, 429–474 (2002). https://doi.org/10.5254/1.3547678,citedBy106
13. Neter, J., Kutner, M.H., Nachtsheim, C.J., Wasserman, W.: Applied linear statistical models (1996)

14. Olden, J.D., Jackson, D.A.: Illuminating the "black box": a randomization approach for understanding variable contributions in artificial neural networks. Ecol. Model. (2002). https://doi.org/10.1016/S0304-3800(02)00064-9
15. Pelisser, F., Montedo, O.R.K., Gleize, P.J.P., Roman, H.R.: Mechanical properties of recycled pet fibers in concrete. Mater. Res. **15**, 679–686 (2012). https://doi.org/10.1590/S1516-14392012005000088
16. Sadeghi, B.: Decisiontree.jl (2013)
17. Whiffin, V.S.: Microbial caco3 precipitation for the production of biocement (2004)
18. Xu, B., Chen, T.: Xgboost.jl (2014)