Vol. 11, 2022, 203-214

DOI: 10.17512/bozpe.2022.11.23

Construction of optimized energy potential
Budownictwo o zoptymalizowanym potencjale energetycznym

ISSN 2299-8535 e-ISSN 2544-963X

A principal component analysis in concrete design

Janusz Kobaka[1] (*orcid id: 0000-0002-3596-3187*)
Jacek Katzer[1*] (*orcid id: 0000-0002-4049-5330*)
[1] University of Warmia and Mazury in Olsztyn, Poland

Abstract: Over the last 200 years, ordinary concrete has evolved from four basic ingredient materials (gravel, sand, cement, and water) to multicomponent complex composites. The number and variety of the additives, admixtures, non-conventional aggregates, fillers, and fibres currently used for concrete production have continued to grow rapidly. Regrettably, the methods for de-signing concrete mixes have not evolved at a similarly fast pace. Keeping the above facts in mind, the authors utilised a principal component analysis (PCA) to design modern concrete mixes. As an initial approach, 550 cast and tested concrete mixes were analysed. The main aim of the presented study was to prove the usefulness of the PCA methodology for the fast classification of concrete mix compositions. The acquired knowledge should be useful for the effective design of multicomponent modern concrete mixes.

Keywords: principal component analysis, PCA, concrete designing, concrete mix

Access to the content of the article only on the bases of the Creative Commons licence CC BY-NC-ND 4.0

Please, quote this article as follows:

Kobaka J., Katzer J., A principal component analysis in concrete design, Construction of Optimized Energy Potential (CoOEP), Vol. 11, 2022, 203-214, DOI: 10.17512/bozpe.2022.11.23

Introduction

Since 1824, when Joseph Aspdin obtained a patent for the invention of Portland cement (To Joseph Aspdin, of Leeds, in the County of York …, 1826), ordinary concrete has played a key role in civil engineering. Initially, concrete was produced using only four ingredients (sand, gravel, water, and cement). The composition of a mix was prepared on a trial-and-error basis, or based on the previous experience of an engineer. Things became more sophisticated after Abrams law (Singh et al., 2015) was formulated, and after it was further developed by Bolomey (Bolomey, 1927). The water/cement (w/c) ratio is still used for forecasting the compressive strength of hardened concrete (de Brito et al., 2018), as follows:

[*] Corresponding author: jacek.katzer@uwm.edu.pl

$$f_c = A_1\left(\frac{1}{\frac{w}{c}} - 0.5\right)\ if\ \frac{w}{c} > 0.4 \tag{1}$$

$$f_c = A_2\left(\frac{1}{\frac{w}{c}} + 0.5\right)\ if\ \frac{w}{c} \leq 0.4 \tag{2}$$

where A_1 and A_2 depend on the quality of the aggregates and cement.

The linear relationship between the w/c ratio and compressive strength of concrete (Salem & Pandey, 2015) enables the design of concrete mixes. The above presented Bolomey equations are a key element of the currently used design methods for concrete mixes.

Over the last 200 years, alongside developments in concrete technologies, numerous new types of concrete have entered the construction industry (Abdul-Razzaq, 2017; Aygün et al., 2021; Xuan et al., 2016). These advanced concretes can be divided into the following groups: high-performance concrete, self-compacting concretes, fibre-reinforced concretes, and water-tight concretes. With the emergence, development, and widespread use of these new advanced concrete types, the limitations of traditional mix design methods have become unquestionable. Modern concrete mixes are rich in content and consist of multiple materials instead of the traditional four ingredients (cement, fine aggregate, coarse aggregate, and water). New components such as fly ash, slag, silica fume (and other micro-fillers), superplasticisers, stabilisers, organic admixtures, fibres, and ceramic (red and white) (Katzer et al., 2020) are commonly used in these mixes (Atiş & Karahan, 2009; Hornakova et al., 2019; Kansal & Goyal, 2021; Zhang et al., 2021). Non-conventional aggregates such as waste aggregates and lightweight aggregates add another level to the complexity of modern mix design problems. Since the end of the 20th century, new design methods have been proposed to address these problems (e.g. the Mehta/Aitcin method (Mehta & Aietcin, 1990; Alves et al., 2004), Laboratory Central des Ponts et Chausses method (Rossi, 1997; Benaicha et al., 2019), compactness method (Zhang et al., 2015), overall calculation method (Xie, 2012; Shen et al., 2011) and genetic algorithm method (Gheibi et al., 2018; Kumar, 2013)). Some of these methods have proven to be quite successful for specific cases of concrete mixes. However, none of them are universal enough to be harnessed. Varied quantities of numerous materials are contained in any given modern concrete, thereby disabling an effective assessment of their influence on the properties of hardened cement composites. Moreover, traditional design methods focus solely on the compressive strength of hardened concrete. Currently, various other properties of both the fresh concrete mix and hardened concrete are expected to be precisely designed. New requirements associated with exposure classes (e.g. in accordance with EN 206 (Concrete. Specification, Performance, Production and Conformity, 2013)) have placed additional pressure on the accurate prediction of watertightness, absorptivity, and/or freeze-thaw resistance.

In the authors opinion, a new approach is needed for designing modern multicomponent concrete mixes. Considering the growing number of ingredients in a concrete mix (which may be even larger in the future owing to the possible in-situ production of cement (Halbiniak et al., 2020)), a more statistical approach should be considered. Principal component analysis (PCA) and projection to latent structures (also known as the discrete Karhunen-Loève transform) seem to be best suited for the task (Kobaka, 2021). PCA has been successfully used in multiple scientific areas such as signal denoising, classification tasks, physics, chemistry, material science, image processing, neuroscience, and quantitative finance (Zarzycki, 2017). It was also recently used for the classification of steel fibres used for reinforcing concrete (Zarzycki et al., 2017). In this case, harnessing PCA enabled fast classification of the fibres. A quick selection of the key geometric and mechanical characteristics of the fibres was also enabled. Most recently, PCA has been utilised for the analysis and comparison of lunar soil simulants (LSSs) with lunar soil samples (Zarzycki & Katzer, 2018; Kobaka et al., 2019). The PCA of LSSs has allowed for the selection of the most adequate LSS for a specific civil engineering research programme.

The main aim of the presented research project was to demonstrate the usefulness of the PCA methodology for the fast classification of concrete mix compositions. The acquired knowledge should be useful for the effective design of multicomponent concrete mixes in the future.

1. Research programme

PCA algorithms are considered to be specific implementations of linear algebra. They enable significant data reduction, and the acquisition of concealed and unapparent information from large raw datasets. Currently, PCA is gaining significant popularity in all scientific areas where it is essential to obtain strong patterns from large or complex datasets. It is a very efficient analytical instrument for finding concealed and unapparent correlations within a dataset; moreover, it enables statistical representations of such datasets (Pieters, 2002). Using PCA, one can compute the optimal presentation of multivariate datasets. The raw data is projected into a space described by eigenvectors of a data variance-covariance matrix (Johnson et al., 1985). Utilisation of PCA results in a reduction in the dimensions of multivariate datasets. Usually, two or three components are obtained, and can be easily visualised using traditional two- or three-dimensional plots. These new virtual variables capture the vast majority of the original information. The entire transformation is characterised by a minimal loss of information (Kassambara, 2017).

In this study, the analysis was conducted for three datasets of concrete mix recipes (Table 1). A total of 550 recipes were considered. Each recipe was differentiated according to the number of ingredients and their volumes. The ingredients of each concrete mix were given in kg/m^3. The ingredients were treated as variables using the PCA.

Table 1. Concrete mixes datasets used for the principal component analysis (PCA)

Number of mixtures	Composition	Properties
425*	cement, coarse aggregate, fine aggregate, blast furnace slag, fly ash, water, superplasticiser	compressive strength after 28 days
103**	cement, coarse aggregate, fine aggregate, slag, fly ash, water, superplasticiser	slump test, flow table test
22***	cement, pebble gravel (2-16, 2-8, 8-16 mm, granite grit (2-8, 8-16 mm), sand, water, superplasticiser	water permeability

Data sourced from: * (see Information division), ** (see Information division), *** local concrete production plant located in Poland.

2. Research results

During the first stage of the research program, 425 concrete recipes were analysed using PCA. The cases were grouped into three groups (A, B, and C) according to their compressive strength (Fig. 1). The resultant chart explains approximately 51% of cases.

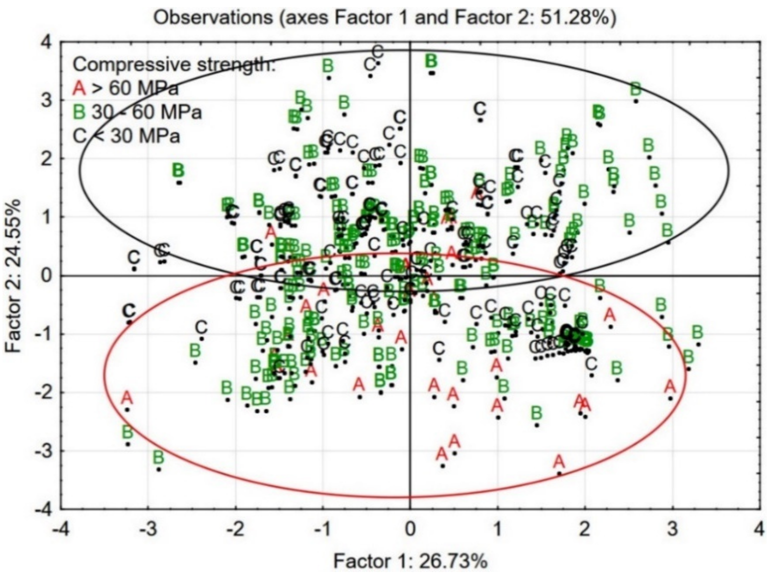

Fig. 1. Principal component analysis (PCA) with object grouping in two-dimensional space on the basis of concrete composition. A, B, C – compressive strength groups (see description in the text) (*own research*)

Group A consists of concretes characterised by compressive strengths ranging from 60 to 90 MPa (marked in red in Fig. 1). The majority of these cases (surrounded by the red ellipse) are placed at the bottom half of the chart, principally at the bottom right quarter, which corresponds to a high amount of cement, high amount of coarse aggregate, low amount of water, and low presence of blast furnace slag and fly ash (Fig. 2 and Table 2). The high compressive strength, bound up with a high amount of cement together with a low water content, is justified by the strong dependence of this property on the w/c ratio.

Group B (marked in green in Fig. 1) represents concrete mixes in the middle range of compressive strength from 30 to 60 MPa. These mixes are scattered throughout the chart. They are relatively evenly present in all quarters of the chart.

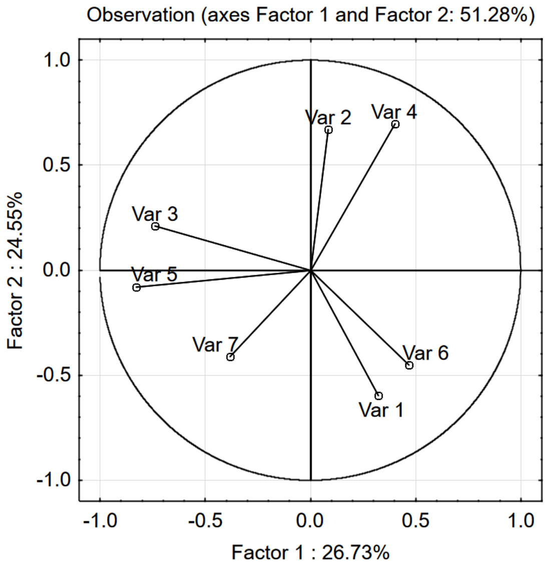

Fig. 2. PCA projection of variables set in 2D factor loading space (variables designation assignment – see Table 2) (*own research*)

Table 2. Contributions of the variables in PCA factors for compressive strength values (*own research*)

Variable designation	Designation assignment	Contribution of the variables [%]	
		Factor 1	Factor 2
Var 1	Cement	5.6	**20.8**
Var 2	Blast Furnace Slag	0.4	**26.2**
Var 3	Fly Ash	**29.1**	2.6
Var 4	Water	8.6	**28.3**
Var 5	Superplasticiser	**36.8**	0.4
Var 6	Coarse Aggregate	**11.7**	11.8
Var 7	Fine Aggregate	7.8	9.9

Group C (marked black in Fig. 1) represents mixes characterised by the lowest compressive strength (from 8 to 30 MPa). These mixes are mostly present in the top half of the chart (surrounded by a large black ellipse), and corresponds to a high amount of water and low amounts of cement and natural aggregate (coarse and fine); the latter were partially replaced by blast furnace slag and fly ash (Fig. 2 and Table 2). The most significant contributing variables for Factor 1 are the superplasticiser, fly ash, and coarse aggregate (Table 2). These three variables together represent 77.6% of the contributions of the variables. In the case of Factor 2, water, blast furnace slag, and cement represent 75.3% of the contributions of the variables. The PCA allows for the separation of most cases characterised by extremely different compressive strengths.

During the second stage of the research program, 103 concrete recipes were analysed using PCA (Fig. 3). The cases are grouped according to the flow test values, in accordance with (BS EN 12350-5 2009). The concrete mixes characterised by flow consistency F1 (up to 340 mm) are marked in red in Figure 3. Most of the cases in this group of mixes and group B are located in the red ellipse. These mixes are characterised by a low amount of fine aggregates, water, and cement (Fig. 4 and Table 3). The concrete mixes characterised by flow consistencies F3 and F4 are spread evenly throughout the chart. The concrete mixes characterised by flow consistencies F5 and F6 are mostly located in the black ellipse. The C mixes are characterised by low amounts of coarse aggregates and high amounts of fine aggregates, or high amounts of water and cement (Fig. 4 and Table 3).

The most significant contributing variables for Factor 1 are the coarse aggregate, water, and fly ash (Table 3). The three variables together constitute 70.1% of the contributions of the variables. In the case of Factor 2, the slag and superplasticiser constitute 82.8% of the contributions of the variables. Factor 2 is the most influential factor.

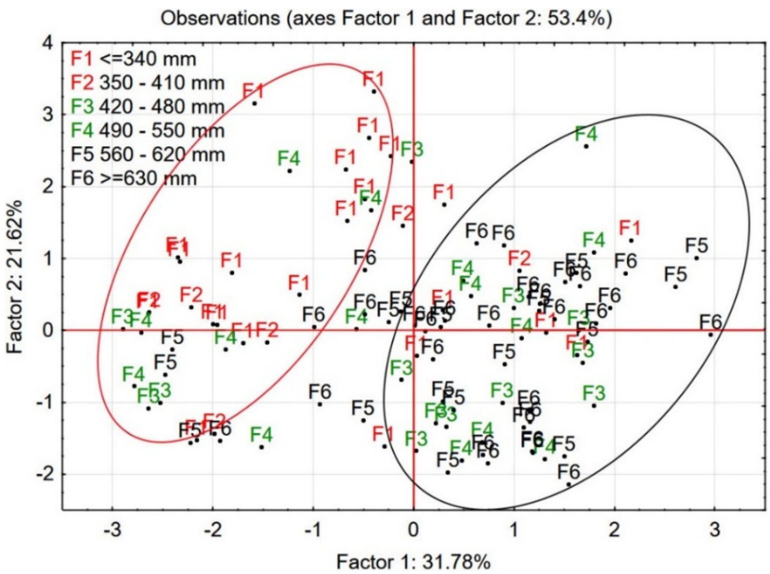

Fig. 3. PCA with object grouping in two-dimensional space on the basis of concrete composition with marked groups of flow test values (*own research*)

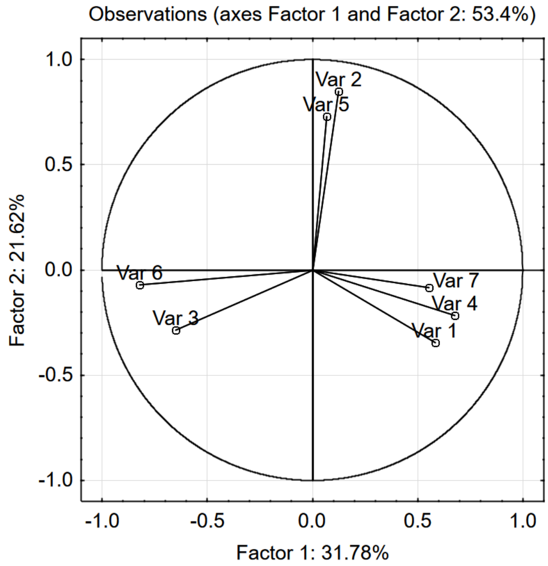

Fig. 4. PCA-projection of variables set in 2D factor loading space (variables designation assignment – Table 3) (*own research*)

Table 3. Contributions of the variables in PCA factors flow test values (*own research*)

Variable designation	Designation assignment	Contribution of the variables [%]	
		Factor 1	Factor 2
Var 1	Cement	15.3	7.9
Var 2	Slag	0.7	**47.7**
Var 3	Fly ash	**19.0**	5.4
Var 4	Water	**20.7**	3.1
Var 5	Superplasticiser	0.2	**35.1**
Var 6	Coarse aggregate	**30.4**	0.3
Var 7	Fine aggregate	13.8	0.5

During the third step of the research program (Table 1), 22 concrete recipes were studied (Fig. 5). These cases were grouped according to the water penetration depth (BS EN 12390-8:2009 2009). The concrete mixes characterised by low water penetration (up to 30 mm) are marked in Figure 5 in red (A, B, and C letters in the chart). Some of these cases are located in the left part of the chart, and are surrounded by the red circle which corresponds to the presence of granite grit comprised of two fractions: 2-8 mm and 8-16 mm (Fig. 6 and Table 4). The rest of the cases marked red and surrounded by the red circle are located in the bottom part of the chart, and correspond to the presence of pebble gravel comprised of two fractions: 2-8 and 8-16 mm. A large cluster of cases, characterised by high water penetration in the range of 30 to 130 mm (marked in black), is located at the top-right part of the chart.

These cases correspond to the presence of pebble gravel (2-16 mm), and a low cement content. The trends in the distribution of the cases presented in Figures 5 and 6 can be explained simply as follows: it is easier to design a watertight aggregate composition with three separate aggregates: fine aggregate (sand 0-2 mm), medium-grade aggregate (gravel 2-8 mm), and coarse aggregate (gravel 8-16 mm), rather than two types of aggregate: sand (0-2 mm) and gravel (2-16 mm).

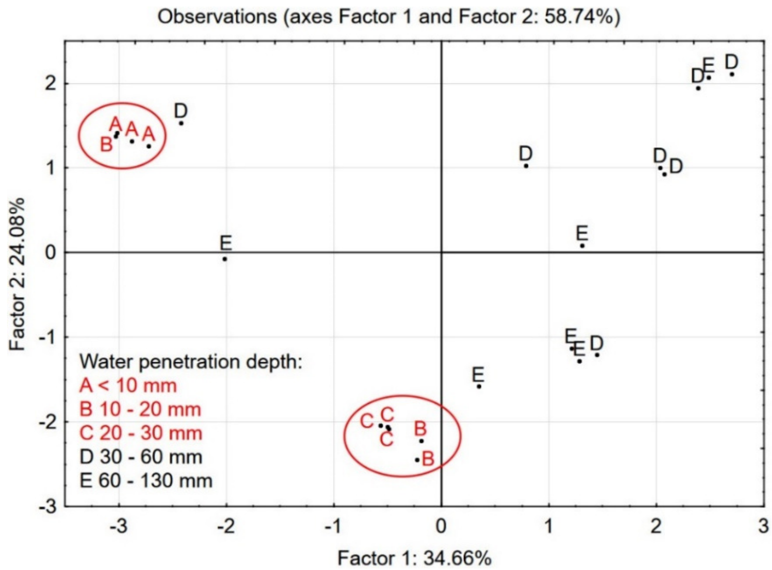

Fig. 5. PCA with object grouping in two-dimensional space on the basis of concrete composition with marked groups of water penetration depth values (*own research*)

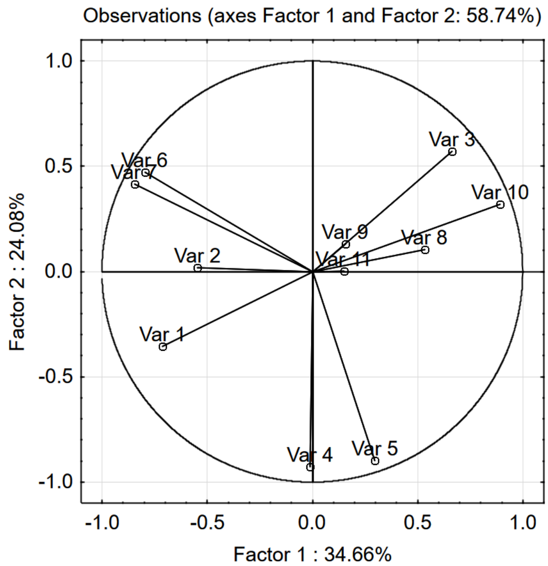

Fig. 6. PCA projection of variables set in 2D factor loading space (variables designation assignment – Table 4) (*own research*)

There are also some cases characterised by high water penetration marked in black at the bottom right section of the chart, in close proximity to the water penetration--resistant cases. Their presence can be explained by the incorrect proportions of aggregates used.

Table 4. Contributions of the variables in PCA factors for water permeability (*own research*)

Variable designation	Designation assignment	Contribution of the variables [%]	
		Factor 1	Factor 2
Var 1	Cement	13.3	4.8
Var 2	Sand 0-2	7.8	0.0
Var 3	Pebble gravel 2-16 mm	11.6	**12.3**
Var 4	Pebble gravel 2-8 mm	0.0	**32.6**
Var 5	Pebble gravel 8-16 mm	2.3	**30.5**
Var 6	Granite-grit 2-8 mm	**16.6**	8.4
Var 7	Granite-grit 8-16 mm	**18.8**	6.5
Var 8	Water	7.5	0.4
Var 9	Silica fume	0.6	0.7
Var 10	Fly ash	**20.9**	3.9
Var 11	Sum of concrete admixtures	0.6	0.0

The most significant contributing variables for Factor 1 are fly ash, granite grit 8-16 mm, and granite grit 2-8 mm (Table 4). The three variables together constitute 56.3% of the contributions of the variables. In the case of Factor 2, pebble gravel (2-8 mm), pebble gravel (8-16 mm), and pebble gravel (2-16 mm) constitute 75.4%, and are the most influential.

3. Discussion

To achieve certain favourable properties of concrete, a combination of basic and additional ingredients are generally incorporated into a concrete mix. Deciding their proportions at the mix design stage is far from easy. The numerous desired properties for both fresh mix and hardened concrete, such as in regards to the consistency, workability, strength, durability, permeability, and abrasion resistance, influence the mix design procedure. During the conducted research, the compressive strengths of over 400 different concrete mixes were analysed using PCA. The results prove that the commonly used process for mix design is very inefficient. Many of the mixes are characterised by high amounts of water and low amounts of cement (Figs. 1 and 2,

Table 1), which may improve the consistency, but reduces the strength. Moreover, excessive amounts of slag or fly ash may lower the compressive strength. Another PCA analysis performed on the data of over 100 concrete mixes displayed a considerably strong influence of the fine aggregate and water amounts on the flow test values (Figs. 3 and 4, Table 3). Only one of the PCA analyses (performed on 22 cases) showed a strict grouping of cases. An analysis of watertight concrete mixes showed a direct relationship between the mix design and water penetration depth. The major influence on water penetration was from the adequate selection of the aggregate composition. The concrete mixes made with aggregates, composed of two fractions (2-8 mm and 8-16 mm), resulted in a far lower water penetration depth than a concrete mix made with a continuous aggregate fraction (2-16 mm) (Figs. 5 and 6, Table 4). Considering the analysis results from the compressive strength and flow tests, one may conclude that the commonly used concrete mix design methods are inadequate for meeting the technological needs. The scattered nature of the spacing of object groupings in the two-dimensional spaces presented in Figures 3 and 5 suggests that in many cases, the mix design for a modern multicomponent concrete is based more on a trial-and-error basis than on a real mix design. A PCA analysis only forms a proof of concept for harnessing a PCA approach for the design of modern concrete mixes. Using numerous existing sets of concrete mixes which have been designed, cast, and tested for PCA would provide more precise identification of the factors playing key roles in achieving particular properties in fresh and hardened concrete. Such research programs should provide clear directions for the development of design methods for modern multicomponent concrete mixes.

Conclusions

The conducted analysis allows certain conclusions to be formed, as follows:

- It is possible to determine the properties of a concrete, such as the compressive strength, consistency, and water permeability, based only on the amounts of its ingredients.
- The method for determining the concrete properties based on the amount of ingredients and PCA method is not sensitive enough for the medium ranges of the values of the given properties.
- The method using PCA may be suitable for a preliminary estimation of concrete properties at the concrete mix design stage.
- The proposed methodology allows for fast and robust object classifications to be obtained, in contrast to a univariate approach.
- A minimum number of key parameters should be selected for an accurate concrete mix assessment with a minimum experimental setup, and is possible using the proposed approach.
- Analyses should be conducted using larger sets of concrete mixes, with particular focus on high-performance concretes.

Bibliography

Abdul-Razzaq, K.S. (2017) Proposing a new type of structural slurry infiltrated concrete (SSICON). *Journal of Building Engineering*, 13, 254-265.

Alves, M.F., Cremonini, R.A. & Dal Molin, D.C.C. (2004) A comparison of mix proportioning methods for high-strength concrete. *Cement and Concrete Composites*, 26(6), 613-621.

Atiş, C.D. & Karahan, O. (2009) Properties of steel fiber reinforced fly ash concrete. *Construction and Building Materials*.

Aygün, B., Şakar, E., Agar, O., Sayyed, M.I., Karabulut, A. & Singh, V.P. (2021) Development of new heavy concretes containing chrome-ore for nuclear radiation shielding applications. *Progress in Nuclear Energy*, 133.

Benaicha, M., Belcaid, A., Alaoui, A.H., Jalbaud, O. & Burtschell, Y. (2019) Rheological characterization of self-compacting concrete: New recommendation. *Structural Concrete*, 20(5), 1695-1701.

Bolomey, J. (1927) Determination of the compressive strength of mortars and concretes. *Bulletin Technique de La Suisse Romande*, 16, 22-24.

de Brito, J., Kurda, R. & da Silva, P.R. (2018) Can we truly predict the compressive strength of concrete without knowing the properties of aggregates? *Applied Sciences* (Switzerland), 8(7).

BS EN 12350-5 (2009) *Testing Fresh Concrete – Part 5: Flow Table Test*. British Standard.

BS EN 12390-8:2009 (2009) *Testing Hardened Concrete – Part 8: Depth of Penetration of Water under Pressure*. British Standard.

Concrete. Specification, Performance, Production and Conformity (2013), European Committee for Standardization (CEN)).

Gheibi, M., Karrabi, M., Shakerian, M. & Mirahmadi, M. (2018) Life cycle assessment of concrete production with a focus on air pollutants and the desired risk parameters using genetic algorithm. *Journal of Environmental Health Science and Engineering*, 16(1), 89-98.

Halbiniak, J., Katzer, J., Major, M. & Major, I. (2020) A proposition of an in situ production of a blended cement. *Materials*, 13(10), 2289.

Hornakova, M., Katzer, J., Kobaka, J. & Konecny, P. (2019) Lightweight SFRC benefitting from a pre-soaking and internal curing process. *Materials*, 12(24).

Johnson, P.E., Smith, M.O. & Adams, J.B. (1985) Quantitative analysis of planetary reflectance spectra with principal components analysis. *Journal of Geophysical Research*.

Kansal, C.M. & Goyal, R. (2021) *Analysing mechanical properties of concrete with nano silica, silica fume and steel slag*. Materials Today: Proceedings available from <https://linkinghub.elsevier.com/retrieve/pii/S2214785320407503>.

Kassambara, A. (2017) Practical guide to principal component methods in R: PCA, M (CA), FAMD, MFA, HCPC, Factoextra. STHDA 2, 1-155.

Katzer, J., Kobaka, J. & Ponikiewski, T. (2020) Influence of crimped steel fibre on properties of concrete based on an aggregate mix of waste and natural aggregates. *Materials*, 13(8).

Kobaka, J. (2021) Principal component analysis as a statistical tool for concrete mix design. *Materials*, 14(10), 2668.

Kobaka, J., Katzer, J. & Zarzycki, P.K. (2019) Pilbara craton soil as a possible lunar soil simulant for civil engineering applications. *Materials*, 12(23), 1-11.

Kumar, D.C. (2013) Genetic algorithm based optimum design of prestressed concrete beam. *International Journal for Computational Civil and Structural Engineering*, 3(3), 644-654.

Mehta, P.K. & Aietcin, P.C.C. (1990) Principles underlying production of high-performance concrete. *Cement, Concrete and Aggregates*, 12(2), 70-78.

Pieters, C. (2002) Statistical analysis of the links among lunar mare soil mineralogy, chemistry, and reflectance spectra. *Icarus*, 155(2), 285-298.

Rossi, P. (1997) Strain rate effects in concrete structures: The LCPC experience. *Materials and Structures/Materiaux et Constructions*.

Salem, M.A.A. & Pandey, R.K. (2015) Effect of cement-water ratio on compressive strengthand density of concrete. *International Journal of Engineering Research and Technology* (*IJERT*), 4(2).

Shen, W., Li, F. & Lu, C. (2011) Study on improving overall calculation method of mix design method for fly ash Hpc. *Advanced Materials Research*, 163-167, 1414-1418.

Singh, S.B., Munjal, P. & Thammishetti, N. (2015) Role of water/cement ratio on strength development of cement mortar. *Journal of Building Engineering* [online], 4, 94-100.

To Joseph Aspdin, of Leeds, in the County of York, Bricklayer, for His New Invented Improvement in the Mode of Producing an Artificial Stone (1826). *Journal of the Franklin Institute*, 1(3), 146-149.

Xie, X. (2012) The design of mix proportion of high-performance concrete for passenger dedicated railway based on overall calculation method. *Applied Mechanics and Materials*, 193-194, 814-821

Xuan, D., Zhan, B. & Poon, C.S. (2016) Development of a new generation of eco-friendly concrete blocks by accelerated mineral carbonation. *Journal of Cleaner Production*, 133, 1235-1241.

Zarzycki, P.K. (2017) Application of multivariate statistical techniques in chromatographic and spectroscopic fingerprinting of complex food and environmental samples. *Journal of AOAC International*.

Zarzycki, P.K. & Katzer, J. (2018) *Assessment of lunar soil simulants based on multivariate statistics*. Earth and Space 2018: Engineering for Extreme Environments – Proceedings of the 16th Biennial International Conference on Engineering, Science, Construction, and Operations in Challenging Environments.

Zarzycki, P.K., Katzer, J. & Domski, J. (2017) Fast classification of fibres for concretebased on multivariate statistics. *Computers and Concrete*, 20(1).

Zhang, H., Zou, K., Ji, X., Zhang, C., Tang, F. & Wu, X. (2015) *Mixture design methods for high performance concrete a review*. Proceedings of the 2015 International Conference on Advanced Engineering Materials and Technology. Paris, Atlantis Press.

Zhang, J., Huang, Y., Ma, G. & Nener, B. (2021) Mixture optimization for environmental, economical and mechanical objectives in silica fume concrete: A novel frame-work based on machine learning and a new meta-heuristic algorithm. *Resources, Conservation and Recycling*, 167, 105395.