PROPER ORTHOGONAL DECOMPOSITION FOR GENERATION OF ORGANIZED DATA IN CONCRETE TECHNOLOGY

A Manoj[1], K S Babu Narayan[1]

1. National Institute of Technology Karnataka, Surathkal, India

ABSTRACT. Concrete undoubtedly is the most widely used construction material owing to its mouldability to any conceivable shape. Workability, strength and durability are the major properties of concrete which are decided by a great number of variables. Also there is the interplay of these important characteristics. Notwithstanding the advancements made in concrete technology, getting the right performance levels is considered more an art than science.

Proper Orthogonal Decomposition (POD) is a powerful tool in situations where dependence and interdependence of variables can be studied, interpreted and judiciously selected from among the vast data base to make relevant data less exhaustive and more meaningful.
This paper presents utility of POD in reorganizing and rationalizing data in concrete technology.

Keywords: Concrete, Strength, Slump, POD.

A Manoj is a Research scholar in Computational Mechanics in the Department of Civil Engineering, National Institute of Technology Karnataka, Surathkal, India. His research interest includes data analysis and math modelling in Concrete Technology and computational mechanics.

Dr K S Babu Narayan is a Professor of Civil Engineering at National Institute of Technology Karnataka, Surathkal, India. His research interests are Structural Analysis and Design, Structural Optimization, Math Modelling and Unsaturated Soil Mechanics.

INTRODUCTION

Concrete is a heterogeneous material formed by mixing cement, fine aggregate, coarse aggregate and water, to form a solid mass [1]. Concrete is the most popular construction material because of its mouldability.

Compressive strength, workability and durability are the very important properties that decide the performance levels. Therefore, the influence of the ingredients, mix proportions, mixing methods, placement & compaction and curing should be well understood [2]. Statistical tools are of immense help in predicting the behavioral aspects [3, 4].

POD is a popular dimensionality reduction technique, which helps in hierarchizing the various influencing variables based on their variability [5]. Therefore, this technique is helpful in development of mathematical models that are operational and require less computational effort and time.

POD ALGORITHM

Algorithm for POD [5, 6] has the following sequential steps in reorganizing and rationalizing data for subsequent use.

- **Step 1:** Relevant data acquisition is the first step in POD.
- **Step 2:** The data acquired shall be checked for size, completeness and outliers.
- **Step 3:** Creation of synthetic variables from available data may be considered if established relationship exists, to reduce time and effort.
- **Step 4:** As numerical values of different variables are likely to have different ranges to compare, discriminate and select components in POD, normalization is practiced.

Z-score standardization is a popular normalization method used in POD.

$$z_i = \left(\frac{x_i - \bar{x}_k}{\sigma_k}\right) \quad (1)$$

where,

z_i = Standardized variable

x_i = Original variable

$\bar{x}_k$ = Variable mean

σ_k = Standard deviation

$$= \sqrt{\frac{1}{N}\sum_{i=1}^{N}(x_i - \bar{x}_k)^2} \quad (2)$$

N = Total no. of observations

K = Total no. of variables

- **Step 5:** For the normalized data correlation matrix is assembled. Check for singularity, sample adequacy and sphericity of data are made.
- **Step 6:** From the correlation matrix eigenvalues and eigenvectors are extracted. Eigenvectors indicate the direction in which the greatest variations are seen. Eigenvalues quantify the relative amount of variation explained by the components.
  Correlation matrix is always a symmetric matrix, the eigenvalues are always real and eigenvectors are orthogonal to each other.
- **Step 7:** Data reduction and hierarchization is done, based on the end objective of the exercise. Scree plots, eigenvalues and eigenvectors help in decision making as to how many axes need to be considered.

ILLUSTRATIVE EXAMPLE

An available data set [7] consisting of 6 original variables namely compressive strength (σ), water content (W), cement content (C), fine aggregate (FA), coarse aggregate (CA) and slump (S) and four synthetic variables (W/C, CA/C, FA/C and C/(CA+FA)) for 97 observations has been subjected to POD. Variation of compressive strength and slump have been investigated with reference to other variables.

IMPLEMENTATION & INTERPRETATION

- Figure 1 shows the scatter matrix plot for the data analyzed. The plot provides graphical explanation for dependence, independence and interdependence.

- The correlation matrix for the normalized data is given in Table 1. Coefficient of correlation matrix reveals the measure of linear dependency of each variable with another. Sample adequacy check by KMO test with a value 0.548(>0.5) indicates adequacy and Bartlett test has revealed that the data set is not spherical.

Table 1 Correlation Matrix

	σ_c	C	Sl	W/C	CA/C	FA/C	W	FA	CA	C/(FA+CA)
σ_c	1	0.69	0.22	-0.65	-0.58	-0.52	0.31	-0.07	-0.27	0.70
C	0.69	1	0.21	-0.91	-0.90	-0.70	0.55	0.00	-0.41	0.98
Sl	0.22	0.21	1	0.09	-0.51	0.31	0.68	0.73	-0.80	0.26
W/C	-0.65	-0.91	0.09	1	0.71	0.79	-0.16	0.26	0.15	-0.88
CA/C	-0.58	-0.90	-0.51	0.71	1	0.45	-0.73	-0.31	0.74	-0.92
FA/C	-0.52	-0.70	0.31	0.79	0.45	1	-0.13	0.69	-0.21	-0.69
W	0.31	0.55	0.68	-0.16	-0.73	-0.13	1	0.47	-0.66	0.54
FA	-0.07	0.00	0.73	0.26	-0.31	0.69	0.47	1	-0.75	-0.01
CA	-0.27	-0.41	-0.80	0.15	0.74	-0.21	-0.66	-0.75	1	-0.50
C/(FA+CA)	0.70	0.98	0.26	-0.88	-0.92	-0.69	0.54	-0.01	-0.50	1

Figure 1 Scatter matrix plot of Variables

- Component eigenvalues, percentage variation explained by the component and cumulative variation explained by the successive components are tabulated and presented in Table 2.
-

Table 2 Eigenvalues

COMPONENT	EIGEN VALUES	% OF VARIANCE	CUMULATIVE %
1	5.361	53.609	53.609
2	3.238	32.378	85.987
3	0.535	5.352	91.339
4	0.485	4.850	96.188
5	0.210	2.098	98.286
6	0.142	1.418	99.704
7	0.027	0.273	99.977
8	0.001	0.014	99.991
9	0.001	0.008	99.999
10	8.985E-5	0.001	100.000

- Scree plot presented in Fig. 2 helps in selection of number of components for further interpretation and analysis.

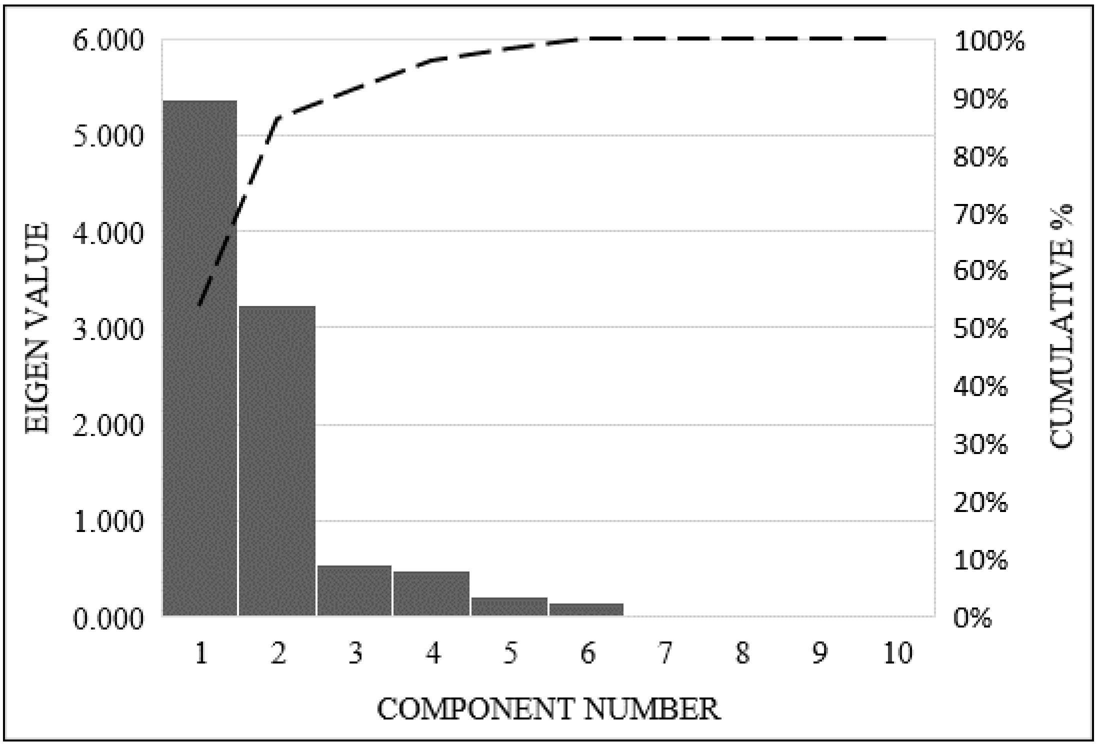

Figure 2 Scree plot

Here, it can be seen that the gradient of the plot shows a steep drop up to component 3. There after the change in the gradient is insignificant suggesting first 3 components are significant. Also from Table 2 it is evident that 91% of variation is explained by first 3 components which are arranged in Table 3.

Table 3 Eigenvectors

	COMPONENTS		
	1	2	3
σ_c	0.316	-0.126	0.624
C	0.412	-0.128	-0.038
Sl	0.194	0.439	0.028
W/C	-0.341	0.289	-0.211
FA/C	-0.245	0.431	0.307
CA/C	-0.419	-0.081	0.107
FA	0.076	0.513	0.222
W	0.295	0.264	-0.627
CA	-0.270	-0.393	-0.126
C/(FA+CA)	0.418	-0.107	-0.011

- Vector loadings also known as component scores is presented in Figure 3.

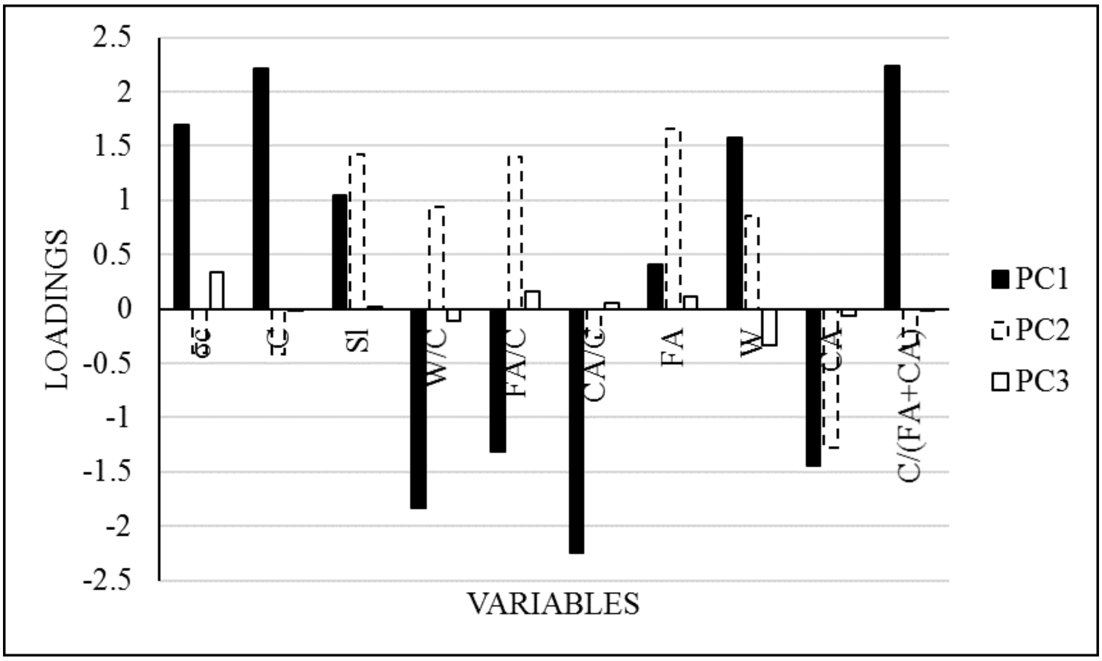

Figure 3 Vector Loadings

- Fig. 4 is the Component plot for the data after decomposition.

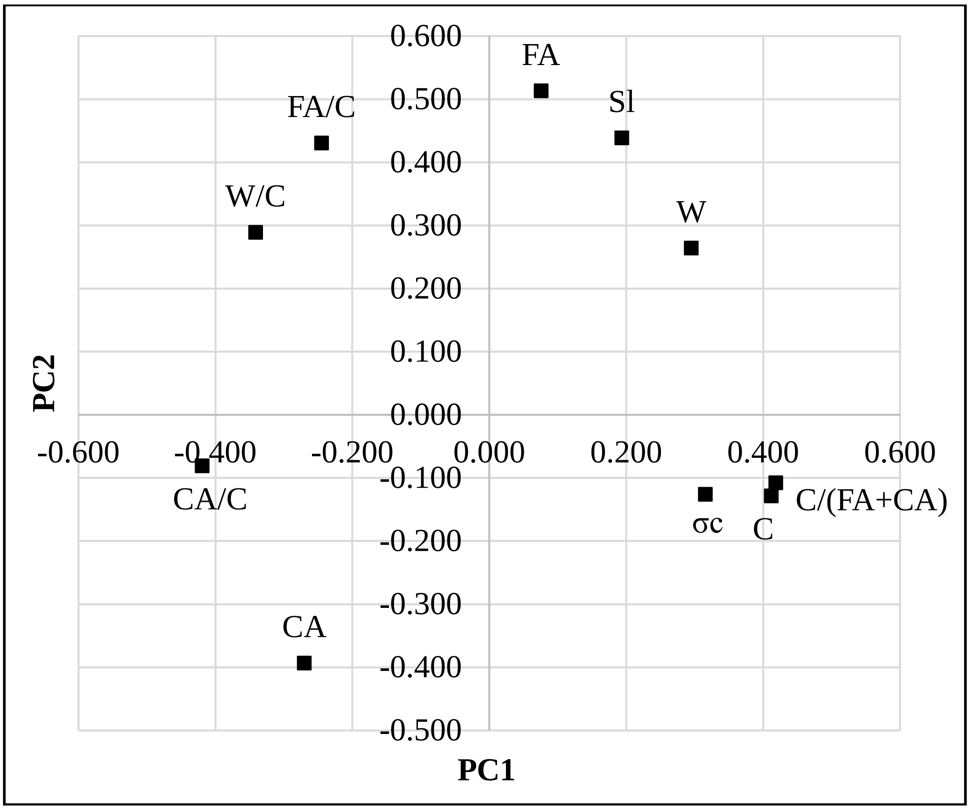

Figure 4 Component plot

- Data reduction to these principal axes can be done by utilizing the eigenvectors on original data matrix resulting in score plot, as shown in Figure 5. This plot can be employed in further deduction of the data.

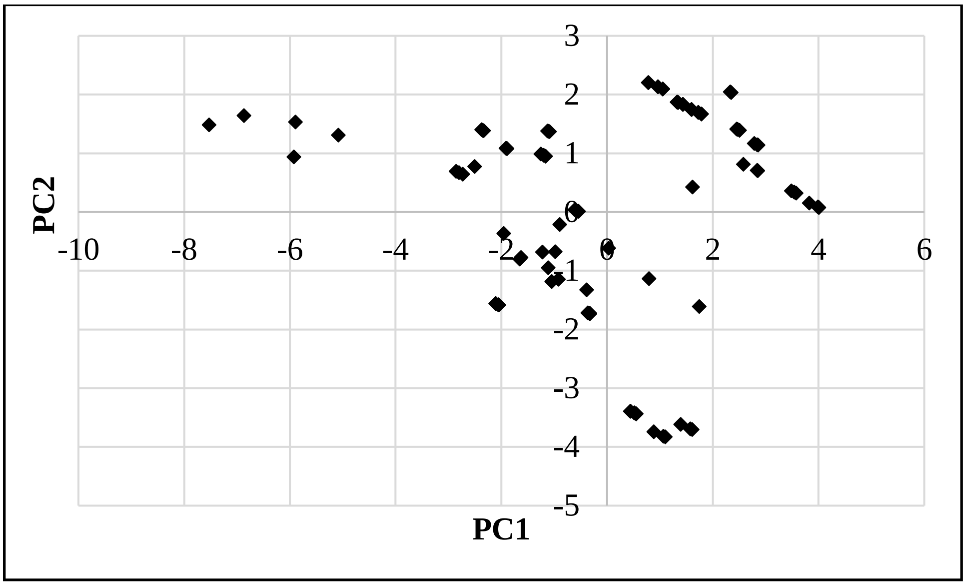

Figure 5 Score plot

CONCLUDING REMARKS

POD for the data indicates,

- Increase in compressive strength with cement content and ratio of cement to aggregate.
- Decrease in strength with increase in water to cement ratio. This is consistent with the recognized Abraham's law.
- Strength is negatively correlated to fine aggregate to cement content and Coarse aggregate to cement content ratios.
- Slump as an indicator of workability is directly influenced by water and fine aggregate content as intuitively expected.
- Higher coarse aggregate to cement ratio makes the mix harsher and hence reduces the slump.

ACKNOWLEDGEMENTS

We profusely thank Dr. Subhash C Yaragal, Professor in Civil Engineering Department, NITK and his research team for providing data employed in illustration of POD in Concrete technology.

REFERENCES

1. NEVILLE A M AND BROOKS J J, Concrete Technology, Pearson Edu Ltd., London, 2010.

2. PAN Y, PRADO A, PORRAS R, HAFEZ O M, AND BOLANDER J E, Lattice modeling of early-age behavior of structural concrete, Materials, February 2017, Vol. 10, No. 3, pp 231.
3. EBERLY COLLEGE OF SCIENCE, Stat 505, Applied Multivariate Statistical Analysis, PennState, Website:https://onlinecourses.science.psu.edu/stat505/node/1
4. ISLAM M S AND ALAM S, Principal component and multiple regression analysis for steel fiber reinforced concrete (SFRC) beams, International Journal of Concrete Structures and Materials, December 2013, Vol. 7, No. 4, pp 303-317.
5. Jolliffe I T, Principal component analysis, Springer, New York, 2002.
6. LIANG Y C, LEE H P, LIM S P, LIN W Z, LEE K H AND WU C G, Proper orthogonal decomposition and its applications - Part I: Theory, Journal of Sound and Vibration, 2002, Vol. 252, No. 3, pp 527-544.
7. PRADEEP B K, SHASHANK B N AND KARTHIK S U, Application of Principal Component Analysis in Civil Engineering, Thesis, NITK, Surathkal, 2012.