Genetic Algorithm Combining Operation Tree (GAOT) for concrete performance

Li Chen[1, a], Chang-Huan Kou[1,b] and Shih-Wei Ma[1,c]

[1]Department of Civil Engineering, Chung Hua University, Hsinchu, Taiwan, 30012, R.O.C.

[a]lichen@chu.edu.tw, [b]chkou@chu.edu.tw, [c]d09404006@chu.edu.tw

Keywords: GAOT. Slump Flow. High-Performance Concrete.

Abstract. Genetic Algorithm Combining Operation Tree (GAOT) was constructed to estimate the slump flow of high-performance concrete (HPC) by using seven concrete ingredients. HPC is a highly complex material; because modeling its behavior is extremely difficult, robust optimization techniques are required. GAOT is a type of evolutionary algorithm that simultaneously optimizes functions and their associated coefficients and is suitable for automatically discovering relationships between nonlinear systems. In a case study, it was observed that for estimating HPC slump flow, the GAOT is more accurate than regression model and back-propagation neural networks (BPNN).

Introduction

High performance concrete (HPC) is a highly complex material manufactured of four to ten different components and its behavior is thus difficult to model. HPC incorporates supplementary cementitious materials, including fly ash, blast furnace slag, and chemical admixtures, such as superplasticizers [1]. Traditional approaches model performance of concrete assuming an analytical equation followed by a regression analysis using experimental data to determine unknown coefficients in the equation [2]. In recent years, artificial neural networks (ANNs) performed exceptionally as regression tools because they are more direct than traditional statistical methods [3]. Moreover, ANNs are highly nonlinear and can capture complex interactions between input/output variables in a system, without prior knowledge about the nature of interactions. However, such "black box" models do not easily generate analytical equation forms [4]. The references [5, 6] proposed using genetic operation trees (GOT) to study concrete strength.

The main purpose of this paper is based on GOT to improve the modeling techniques of estimating HPC parameters by Genetic Algorithm Combining Operation Tree (GAOT). Because of the complex nonlinear relationship between several ingredients and the slump flow of HPC, GOT usually suffers from premature convergence which cannot acquire satisfying solutions. The results show GAOT is better, compared with regression model and back-propagation neural networks (BPNN).

Genetic Algorithm Combining Operation Tree (GAOT)

Genetic Algorithm (GA). Genetic algorithm (GA) is an optimization method that employs a unique search algorithm which can jump from a local optimum to close to the global optimum. The concept of GA was derived from Darwin's theories on natural selection and survival of the fittest. GA can generate an optimal solution by considering the optimization problem as an evolution problem. [7]

Operation Tree (OT). Operation tree is a tree structure which represents a mathematical formula. A five-layered operation tree model is shown in Fig. 1. N_1 is the root node denoted a mathematical operation (addition, subtraction, multiplication, division, natural logarithm, or exponentiation). N2–N15 are interior nodes denoted a variable, a constant, or a mathematical operation. N_{16} - N_{31} are leaf nodes denoted a variable or a constant [8]. Fig. 2, shows an example of operation tree model is following Eq. 1.
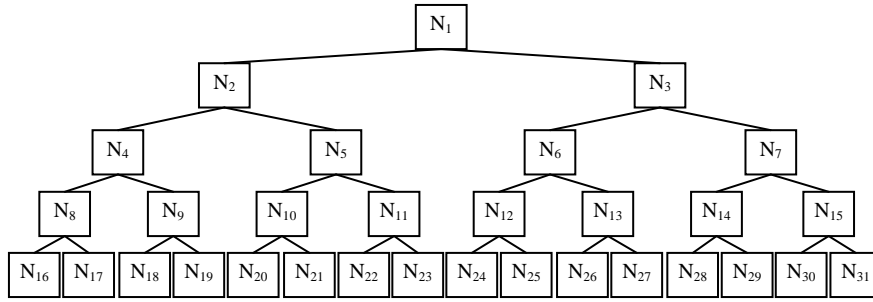
$$Y = (A + B)^{C/D} \tag{1}$$
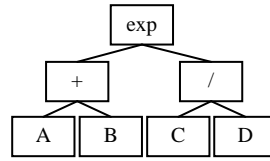

Fig. 1 A five-layered operation tree

Fig. 2 the parse tree of Eq. 1

Mathematical operations, variables and constants in the root, branches, and leaves of the operation tree. When the tree-style structure is set up to represent a specific mathematical formula, operation tree can generate predicted output value for each data by substituting the input values of data into the variables on branches or leafs of the tree-style structure. The conventional regression analysis requires predetermined formula structure and is only allowed to adjust the regression coefficients in the predetermined structure. The disadvantage can be overcome by operation tree. Operation tree is a tree-style data structure which represents a flexible mathematical formula and optimizing the structure to fit the data best is a discrete optimization problem. Therefore, the optimization of operation tree cannot be solved with conventional mathematical programming. Genetic algorithm, that can solve discrete optimization problem, is adopted in this study to optimize the operation trees to fit the data best.

Genetic Algorithm Combining Operation Tree. This study employed operation tree to express a regression formula, and genetic algorithms to optimize the operation tree to fit the data set to produce a self-organized regression formula. Operation tree performance can be evaluated with root mean squared error (RMSE) between predicted and actual output values. A five-layered operation tree was employed in this study as shown in Fig. 1. Table 1 lists the gene codes of mathematical operations, variables, and constants.

Table 1 genetic code of mathematical operations, variables, and constants

Code	1	2	3	4	5	6	7
Meaning	+	-	×	÷	x^y	ln	C
Code	8	9	10	11	12	13	14
Meaning	x_1	x_2	x_3	x_4	x_5	x_6	x_7

Case Study

Concrete is a complex material widely used due to its strong capacity to withstand compression [9]. In addition to the three basic ingredients in conventional concrete, i.e., Portland cement, fine and coarse aggregates, and water, HPC incorporates supplementary cementitious materials, including fly ash, blast furnace slag, and chemical admixtures, such as superplasticizers [1]. The use of fly ash, blast furnace slag, and other replacement materials contributes to better workability [10]. Therefore, the model consisted of seven input variables, including cement (*C*), fly ash (*FL*), slag (*SL*), water (*W*),

superplasticizer (*SP*), coarse aggregate (*CA*), and fine aggregate (*FA*). Table 2 demonstrates the statistical properties of concrete evaluated in this study.

Table 2 Statistical properties of seven components

Variables (kg/m^3)	Minimum	Maximum	Mean	Standard deviation	CC
Cement (*C*)	137.0	374.0	248	78.49	0.19
Fly ash (*FL*)	0.0	193.0	94	60.17	-0.33
Slag (*SL*)	0.0	260.0	126	85.00	-0.06
Water (*W*)	160.0	240.0	194	20.11	0.63
Superplasticizer (*SP*)	4.4	19.0	11.4	2.79	-0.18
Coarse aggregate (*CA*)	708.0	1049.9	853	87.96	-0.33
Fine aggregate (*FA*)	640.6	902.0	753	63.03	0.19

Data Set. Experimental data from References [11,12,13], were used to construct the slump model. The slump test assessed the fresh concrete. Similar to the slump-cone test, the slump flow can be deduced by measuring the diameter of the slumped fresh concrete and the minimum slump flow is the bottom diameter (20 cm) of the slump-cone [10]. To collect training and testing data systematically, mix proportions were examined using a mixture experiment. Ninety-four concrete samples from the above investigations were evaluated, each containing seven components and one output value, the slump flow (from 20 to 78 cm).

Results. Through 400 generations (8 * 50 * 400 = 160,000 evaluations), all subpopulations are converging to the optimal solution CC = 0.897, and RMSE = 8.92 (cm) for training & testing set (Table 3).The best relationship among seven ingredients and slump flow of HPC were acquired through GAOT which can be represented as Eq. 2. The corresponding parse tree of Eq. 2 is a complex ten layered tree type structure shown in Fig. 3. All seven input variables are included and shown in the following equation.

$$Slump\ flow = -18.1 + 0.867 \times \left(0.465 \times W + \ln\left(\frac{11.015 \times \left(W + \frac{SP}{C + FL + SL + CA + FA}\right)}{\frac{W}{C} - \frac{FL}{50.445} + (CA + FA) \times (C + FL + SL)} \right) \right) \quad (2)$$

Table 3 Statistical properties of seven components

Models	Training & testing Sets		Validating Sets	
	CC	RMSE	CC	RMSE
GAOT	0.897*	8.92*	0.824*	9.45*
RA	0.724	10.36	0.464	15.17
BPNN	0.890	9.02	0.805	11.93

Note: the symbol * represents the best result of these models.

The predicted values shown in Fig. 4 are close to the ideal (450) line, indicating that the model accurately predicts experimental results for both the training & testing set and the validating set. The linear regression analysis (RA) including four independent variables was given by:

$$Slump\ flow = -7.24 - 0.04C - 0.14FL - 0.03SL + 0.48W + 0.39SP - 0.02CA + 0.01FA \quad (3)$$

The CC of the RA was 0.464 and RMSE was 15.17 cm for the validating set indicating that the performance of the linear model was not as accurate as that GAOT (Table 3). Fig. 5 demonstrates the scatter diagram of predicted slump flow values versus values observed in the laboratory for the RA. The CC and RMSE of the BPNN were 0.890 and 9.02 cm for the training & testing set, Fig. 6

demonstrates the scatter diagram of predicted slump flow values versus values observed in the laboratory for the BPNN.
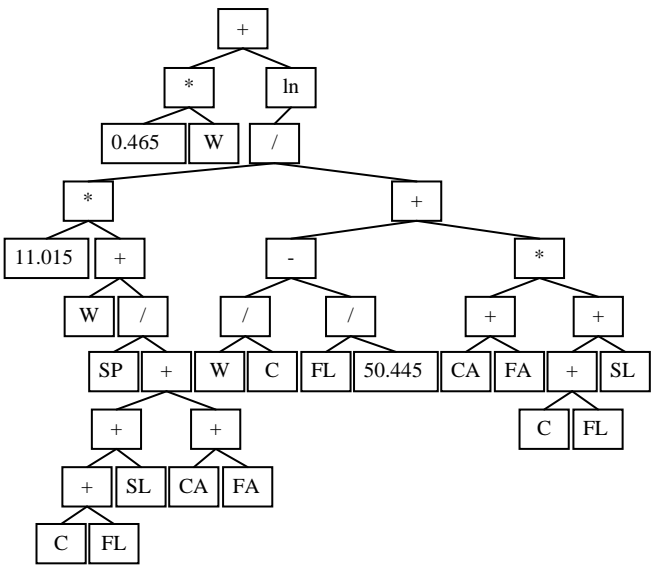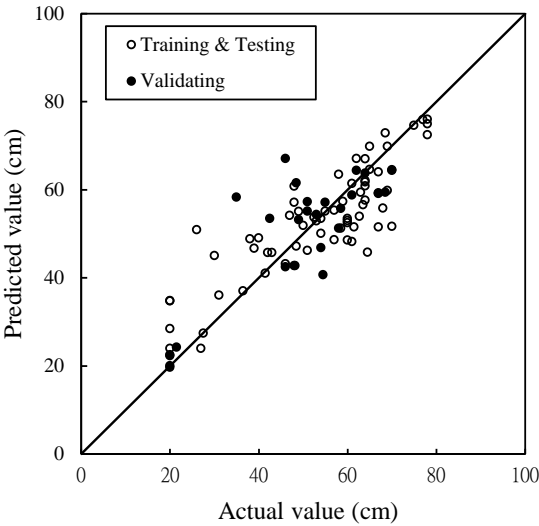

Fig. 3 The parse tree of Eq. 2
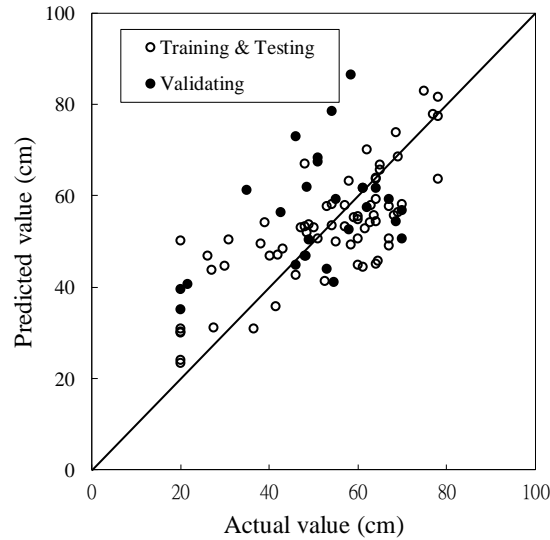

Fig. 4 The scatter plot of Eq. 2
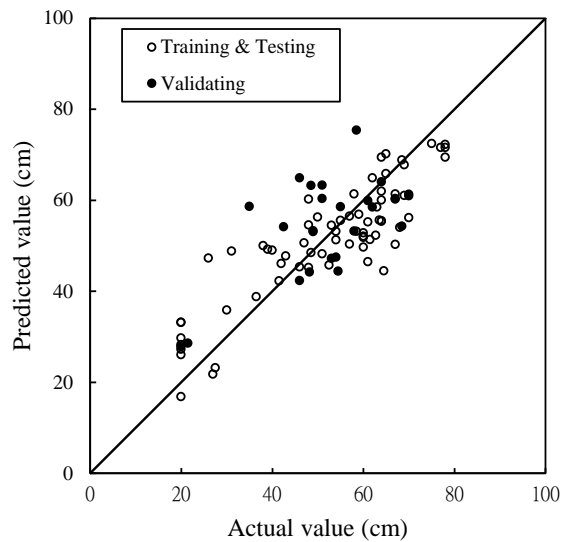
Fig. 5 The scatter plot of RA

Fig. 6 The scatter plot of BPNN

Conclusion

A new and improved version of GOT called GAOT was introduced for modeling the slump flow of HPC.The results showed that the presented model was not as simple as basic formulas, but provided an appropriate model for estimating the slump flow of HPC. We concluded that compared with the regression model, the GAOT is a more accurate model for identifying systems. The results also indicated that the response trace plot of BPNN is similar to those of GAOT. Therefore, the GAOT is recommended as a superior option for modeling slump flow.

References

[1] I.C. Yeh: Exploring concrete slump model using artificial neural networks. J. Comput Civil Eng. 20 (3) (2006), p. 217-221.

[2] M.Y. Mansour: Predicting the shear strength of RC beams using ANN. Eng Struct. 26 (2004), p. 781-799.

[3] I.C. Yeh: Analysis of strength of concrete using design of experiments and neural networks. Journal of Materials in Civil Engineering, ASCE 18 (4) (2006), p. 597-604.

[4] A. Baykasoğlu et al.: Prediction of cement strength using soft computing techniques. Cement and Concrete Research 34 (11) (2004), p. 2083-2090.

[5] I.C. Yeh, L.C. Lien: Knowledge discovery of concrete material using genetic operation trees. Expert Syst. Appl 36 (3) (2009), p. 5807- 5812.

[6] C.H. Peng et al.: Building strength models for high-performance concrete at different ages using genetic operation trees nonlinear regression, and neural networks. Engineering with Computers 26 (1) (2010), p. 61-73.

[7] K.T. Chen, C.H. Kou, L. Chen, S.W. Ma: Application of genetic algorithm combining operation tree (GAOT) to stream-way transition. Proceedings of the 2012 International Conference on Machine Learning and Cybernetics, IEEE (2012), p. 1774–1778, Xian, China.

[8] L. Lien et al.: Modeling strength of high performance concrete using genetic algorithms and operation tree. Journal of technology 21 (1) (2006), p. 41-54.

[9] H.C. Tsai, Y.H. Lin: Predicting high-strength concrete parameters using weighted genetic programming. Engineering with Computers 27 (2011), p. 347-355.

[10] I.C. Yeh: Modeling slump flow of concrete using second-order regressions and artificial neural networks. Cement & Concrete Composites 29 (2007), p. 474-480.

[11] I.C. Chen: Optimization the mixture design of high-performance concrete by neural networks. Master's thesis. Dept. of Civ. Engrg. and Engrg. Informatics, Chung Hua University, Hsin Chu, Taiwan (in Chinese), 2001.

[12] J.W. Chen: Modeling and comparison the workability of high-performance concrete by regression analysis and neural networks. Master's thesis. Dept. of Civ. Engrg. and Engrg. Informatics, Chung Hua University, Hsin Chu, Taiwan (in Chinese), 2002.

[13] L.C. Lien: Applications of genetic algorithms in reinforced learning. Master's thesis. Dept. of Civ. Engrg. and Engrg. Informatics, Chung Hua University, Hsin Chu, Taiwan (in Chinese), 2005.