Structural Engineering and Mechanics, Vol. 38, No. 1 (2011) 1-25
DOI: http://dx.doi.org/10.12989/sem.2011.38.1.001

Nonlinear modeling of shear strength of SFRC beams using linear genetic programming

A.H. Gandomi[1a], A.H. Alavi[2b] and G.J. Yun*[1]

[1]*Department of Civil Engineering, The University of Akron, Akron, OH, USA*
[2]*School of Civil Engineering, Iran University of Science and Technology, Tehran, Iran*

(Received July 3, 2010, Accepted November 18, 2010)

Abstract. A new nonlinear model was developed to evaluate the shear resistance of steel fiber-reinforced concrete beams (SFRCB) using linear genetic programming (LGP). The proposed model relates the shear strength to the geometrical and mechanical properties of SFRCB. The best model was selected after developing and controlling several models with different combinations of the influencing parameters. The models were developed using a comprehensive database containing 213 test results of SFRC beams without stirrups obtained through an extensive literature review. The database includes experimental results for normal and high-strength concrete beams. To verify the applicability of the proposed model, it was employed to estimate the shear strength of a part of test results that were not included in the modeling process. The external validation of the model was further verified using several statistical criteria recommended by researchers. The contributions of the parameters affecting the shear strength were evaluated through a sensitivity analysis. The results indicate that the LGP model gives precise estimates of the shear strength of SFRCB. The prediction performance of the model is significantly better than several solutions found in the literature. The LGP-based design equation is remarkably straightforward and useful for pre-design applications.

Keywords: fiber-reinforced concrete beams; linear genetic programming; SFRC beam; shear strength; formulation.

1. Introduction

Fiber-reinforced concrete (FRC) may be defined as a composite materials made with Portland cement, aggregate, and incorporating discrete discontinuous fibers (Chanh 2004). Shear failure of reinforced concrete beams occurs when the principal tensile stress within the shear span exceeds the tensile strength of concrete and a diagonal crack propagates through the beam web. The addition of steel fibers into the concrete mix is known to increase its shear strength (Gencoglu 2007). If sufficient fibers are added, a brittle shear failure can be suppressed in favor of more ductile behavior. The increased shear strength and ductility of steel fiber-reinforced concrete beams (SFRCB) stems from the post-cracking tensile strength of FRC (Kwak *et al.* 2002). Use of fibers is

*Corresponding author, Professor, E-mail: gy3@uakron.edu
[a]PhD Candidate, E-mail: a.h.gandomi@gmail.com
[b]Researcher, E-mail: ah_alavi@hotmail.com

beneficial, since they are distributed throughout the concrete volume and provide equal resistance to stress in all direction. This is advantageous in structures particularly designed to resist earthquake and wind loading (Li *et al.* 1992). Fiber reinforcement may also significantly reduce construction

Fig. 1 Different types of steel fibers (Wang 2006)

Table 1 Evaluated properties in the literature for the shear strength of SFRC beams

Author(s)	No. of tests	Properties and parameters				
		V_f (%)	L_f/D_f	ρ (%)	f_c' (MPa)	a/d
Batson (1972)	42	X	X		X	X
Roberts and Ho (1982)	6	X		X		X
Jindal (1984)	1					
Swamy and Bahia (1985)	4	X				
Uomoto *et al.* (1986)	24	X	X	X	X	X
Kadir and Saeed (1986)	4		X			
Mansur *et al.* (1986)	9	X		X		
Kaushik *et al.* (1987)	9	X	X			
Murthy and Venkatacharyulu (1987)	7	X	X			X
Narayanan and Darwish (1987)	29	X	X	X	X	X
Lim *et al.* (1987)	6	X				X
Narayanan and Darwish (1988)	7	X			X	X
Li *et al.* (1992)	6		X	X		
Tan *et al.* (1992)	5	X				X
Ashour *et al.* (1992)	8	X		X		X
Swamy *et al.* (1993)	6			X		X
Imam (1994)	5					X
Shin *et al.* (1994)	6	X				X
Adebar *et al.* (1997)	6	X	X			
Oh *et al.* (1998)	2	X				
Casanova (1999)	2					
Noghabai (2000)	11	X	X			
Kwak *et al.* (2002)	4	X			X	X
Cucchiara *et al.* (2004)	4	X				X
Total	213	18	9	7	5	14

V_f (%): Fiber volume fraction; L_f/D_f : fiber aspect ratio; ρ (%): Flexural steel reinforcement ratio; f_c' (MPa): Cylinder compressive strength of concrete; *a/d*: Shear span-depth ratio.

time and costs, especially in an era of high labor costs and possibly even labor shortages. A possible reason is that conventional stirrups require relatively high labor input to bend and fix in place. Fiber concrete can also be easily placed in thin or irregularly shaped sections, such as architectural panels, where it may be very difficult to place stirrups. Another potential area of fiber use is in high-strength concrete. The use of high-strength concrete is attractive for longer spans and taller structures (Khuntia 1999). The relative brittleness and lack of ductility of high-strength concrete can be overcome by inclusion of fibers in high-strength concrete mix (Khuntia *et al.* 1999).

The increase in shear strength can vary drastically depending on the beam geometry and material properties. Li *et al.* (1992) found that the material strength of steel FRC is significantly affected by the volume ratio, aspect ratio, and steel fiber shape. Different types of steel fibers are shown in Fig. 1. These fibers have different mechanical properties such as tensile strength, grade of mechanical anchorage, and capability of stress distribution and absorption. Hence they have different influence on concrete properties (Wang 2006). The literature describes numerous experimental studies of rectangular, fiber-reinforced beams without stirrups. In these studies, several aspects of the FRC beams are evaluated. The geometrical and material properties used in some of these works are summarized in Table 1.

Numerous reports published over the past 30 years have considered the possibility of utilizing FRC by assigning the function of shear reinforcement to the fibers. The achievable advantages are emphasized by ACI Committee 544 (1988) and by RILEM TC 162-TDF (2000). Over the years, several empirical or semi-empirical relations have been developed to determine the ultimate shear strength of SFRCB. Different prediction equations proposed by researchers are presented in Table 2.

Several computer-aided data mining approaches have been developed by extending developments in computational software and hardware. Pattern recognition system, as an example, learns adaptively from experiences and extracts various discriminators. Computational intelligence (CI) (Schwefel *et al.* 2002) techniques are well-known pattern recognition methods. Artificial neural networks (ANNs) are the most widely-used branches of the CI methods. They have successfully been applied to the evaluation of the SFRC beams (Adhikary and Mutsuyoshi 2006). Although

Table 2 Different prediction equations for the ultimate shear strength of SFRC beams

No.	Author(s)	Equation
1	Swamy *et al.* (1993)	$v_{frc} = 0.37\,\tau V_f \dfrac{L_f}{D_f} + v_c \times \begin{cases} 1 & \dfrac{a}{d} \geq 2 \\ 2\dfrac{d}{a} & \dfrac{a}{d} < 2 \end{cases}$
2	Sharma (1986)	$v_{frc} = k f_t \left(\dfrac{d}{a}\right)^{1/4}$ where $k = \begin{cases} 1 & \text{If } f_t \text{ is obtained by direct tension test} \\ 2/3 & \text{If } f_t \text{ is obtained by indirect tension test} \\ 4/9 & \text{If } f_t \text{ is obtained by modulus of rupture} \end{cases}$
3	Narayanan and Darwish (1987)	$v_{frc} = e\left(0.24 f_{spfc} + 80\rho\dfrac{d}{a}\right) + v_b$ $\quad e = \begin{cases} 1 & \dfrac{a}{d} > 2.8 \\ 2.8\dfrac{d}{a} & \dfrac{a}{d} \leq 2.8 \end{cases}$

Table 2 Continued

No.	Author(s)	Equation
4	Ashour *et al.* (1992)	$v_{frc} = \begin{cases} (2.11\sqrt[3]{f_c}+7F)\left(\rho\frac{d}{a}\right)^{1/3} & \text{For } \frac{a}{d}\geq 2.5 \\ (2.11\sqrt[3]{f_c}+7F)\left(\rho\frac{d}{a}\right)^{1/3}\frac{5d}{2a}+v_b\left(2.5-\frac{a}{d}\right) & \text{For } \frac{a}{d}<2.5 \end{cases}$
5	Ashour *et al.* (1992)	$v_{frc} = (0.7\sqrt{f_c'}+7F)\frac{d}{a}+17.2\rho\frac{d}{a}$
6	Kwak *et al.* (2002)	$v_{frc} = 2.1ef_{spfc}^{0.7}\left(\rho\frac{d}{a}\right)^{0.22}+0.8v_b^{0.97} \quad e = \begin{cases} 1 & \frac{a}{d}>3.5 \\ 3.5\frac{d}{a} & \frac{a}{d}\leq 3.5 \end{cases}$
7	Kwak *et al.* (2002)	$v_{frc} = 3.7ef_{spfc}^{2/3}\left(\rho\frac{d}{a}\right)^{1/3}+0.8v_b \quad e = \begin{cases} 1 & \frac{a}{d}>3.4 \\ 3.5\frac{d}{a} & \frac{a}{d}\leq 3.4 \end{cases}$
8	Li *et al.* (1992)	$v_{frc} = \begin{cases} 1.25+4.68(f_f f_{spfc})^{3/4}\left(\rho\frac{d}{a}\right)^{1/3}d^{-1/3} & \text{For } FRC \\ 0.53+5.47(f_f f_{spfc})^{3/4}\left(\rho\frac{d}{a}\right)^{1/3}d^{-1/3} & \text{For } FRM \end{cases}$
9	Khuntia *et al.* (1999)	$v_{frc} = (0.167\alpha+0.25F)\sqrt{f_c'} \quad \alpha = \begin{cases} 1 & \frac{a}{d}\geq 2.5 \\ \min\left(2.5\frac{d}{a},3\right) & \frac{a}{d}<2.5 \end{cases}$
10	Shin *et al.* (1994)	$v_{frc} = \begin{cases} 0.19f_{spfc}+93\rho\frac{d}{a}+0.34\tau F & \text{For } \frac{a}{d}\geq 3 \\ 0.22f_{spfc}+217\rho\frac{d}{a}+0.34\tau F & \text{For } \frac{a}{d}<3 \end{cases}$
11	Mansur *et al.* (1986)	$v_{frc} = 0.41\left(\tau V_f\frac{L_f}{D_f}\right)+\left(0.16\sqrt{f_c}+17.2\frac{\rho Vd}{M}\right)$

v_{frc}: ultimate shear strength, MPa; τ: average fiber matrix interfacial bond stress = 4.15 MPa; v_c: concrete contribution to shear strength and calculated according to ACI design code = $0.167\sqrt{f_c'}$; f_{spfc} computed value of split-cylinder strength of fiber concrete, MPa; f_t: Splitting tensile strength, MPa; f_f: modulus of rupture MPa; α_1 = coefficient representing the fraction of bond mobilized at first matrix cracking = 0.5 and α_2: efficiency factor of fiber orientation in the uncracked state of the composite = 1 (Naaman and Reinhardt 2003); V: shear force at section, N; M: bending moment at section, N.mm; v_b: 0.41 τF, MPa; F: fiber factor $(d_f(L_f/D_f)\ V_f)$; d_f: bond factor.

ANNs are successful in prediction, they are not usually able to produce practical prediction equations. Furthermore, they require the structure of the network (e.g., number of inputs, transfer functions, number of hidden layers, etc.) to be identified a priori. The ANN method is mostly appropriate to be used as a part of a computer program.

Genetic programming (GP) (Koza 1992, Banzhaf *et al.* 1998) is another alternative approach for behavior modeling of structural engineering problems. GP is a relatively new branch of the CI techniques inspired from Darwin's evolution theory. It may generally be defined as a supervised machine learning technique that searches a program space instead of a data space (Banzhaf *et al.* 1998). The programs generated by traditional GP are represented as tree structures and expressed in the functional programming language (Koza 1992). The main advantage of the GP-based approaches over the conventional statistical and ANN techniques is their ability to generate prediction equations without assuming prior form of the existing relationship. Many researchers have employed GP and its variants to discover any complex relationships among experimental data (e.g., Johari *et al.* 2006, Alavi *et al.* 2010a, Gandomi *et al.* 2010a, Mousavi *et al*. 2010). Linear genetic programming (LGP) (Brameier and Banzhaf 2007) is a new subset of GP with a linear structure similar to the DNA molecule in biological genomes. LGP is a machine learning approach that uses sequences of imperative instructions as genetic material. More specifically, LGP operates on genetic programs that are represented as linear sequences of instructions of an imperative programming language (Brameier and Banzhaf 2001, 2007). Based on the numerical experiments, the LGP approach is able to significantly outperform similar techniques (Oltean and Grossan 2003, Baykasoglu *et al.* 2008). Unlike tree-based GP and other soft computing tools like ANNs, the LGP applications to solve problems in civil engineering are restricted to relatively fewer areas. Some of these limited scientific efforts include predicting limestone compressive and tensile strength (Baykasoglu *et al.* 2008), formulation for compressive strength of CFRP confined concrete cylinders (Gandomi *et al.* 2010b), prediction of circular pile scour (Guven *et al.* 2009), and formulation of geotechnical engineering systems (Alavi and Gandomi 2011).

In this study, the LGP approach was utilized to derive an empirical prediction model for the shear capacity of SFRC beams without stirrups. The derived model relates the shear strength to a couple of influencing parameters. LGP is useful in deriving prediction equations for the shear capacity by directly extracting the knowledge contained in the experimental data. The proposed model was developed based on several published shear tests on SFRC beams. The results made by the developed model were further compared with those obtained by several researchers.

2. Genetic programming

Evolutionary algorithms (EAs) (Bäck 1996) are a subset of evolutionary computations. They use biology-inspired mechanisms to optimize a solution with regard to desired result. CI includes EAs and all of their different branches with ANNs and fuzzy logic. The CI techniques have wide ranging applications for approximating the nonlinearities. Developments in the computer hardware during the last two decades have made it much easier for these techniques to grow into more efficient frameworks. In addition, it has been proven that several CI approaches may be used as efficient tools in problems where conventional approaches fail or perform poorly. GP is a robust branch of the CI techniquesthat creates computer programs to solve a problem using the principle of Darwinian natural selection. The breakthrough in GP came in the late 1980s with the experiments

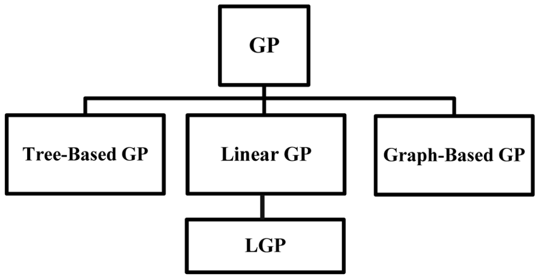


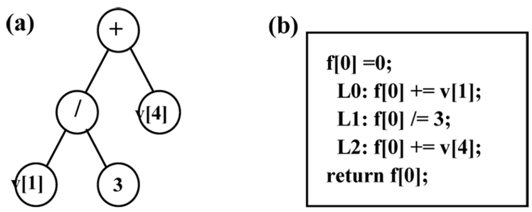
Fig. 2 Different types of genetic programming

on symbolic regression (Koza 1992). GP is an extension of genetic algorithms (GAs). This classical GP technique is also referred to as tree-based GP (Koza 1992). The main difference between the GA and GP approaches is that in GP the evolving programs (individuals) are parse trees rather than fixed-length binary strings. The traditional optimization techniques, like GAs, are in particular used to minimize an existing objective function. In other words, to perform GA-based analyses, an objective function should be defined in advance. Conversely, GP is able to generate prediction equations without any need to establish a pre-defined function or to assume prior form of the relationship. GP is relatively a new field of pattern recognition methods in contrast with GA. A survey of the literature reveals the growing interest of the research community in GP.

In addition to traditional tree-based GP, there are other types of GP where programs are represented in different ways (see Fig. 2). These are linear and graph-based GP (Banzhaf *et al.* 1998, Poli *et al.* 2007). The emphasis of the present study is placed on the linear GP techniques.Several linear variants of GP have recently been proposed such as linear genetic programming (LGP) (Brameier and Banzhaf 2007) and multi-expression programming (MEP) (Oltean and Dumitrescu 2002). The linear variants of GP make a clear distinction between the genotype and phenotype of an individual. In these variants, individuals are represented as linear strings that are decoded and expressed like nonlinear entities (trees) (Oltean and Grosan 2003). There are some main reasons for using linear GP. Basic computer architectures are fundamentally the same now as they were twenty years ago, when GP began. Almost all architectures represent computer programs in a linear fashion. In other words, computers do not naturally run tree-shaped programs. Hence, slow interpreters have to be used as part of tree-based GP. Conversely, by evolving the binary bit patterns actually obeyed by the computer, the use of an expensive interpreter (or compiler) is avoided and GP can run several orders of magnitude faster (Poli *et al.* 2007).

2.1 Linear genetic programming

LGP is a subset of GP with a linear representation of individuals. The main characteristic of LGP in comparison with traditional tree-based GP is that expressions of a functional programming language (like LISP) are substituted by programs of an imperative language (like C/C++) (Brameier and Banzhaf 2001, 2007). Fig. 3 presents a comparison of the program structures in LGP and tree-based GP. A linear genetic program can be seen as a data flow graph generated by multiple usage of register content. That is, on the functional level the evolved imperative structure denotes a special directed graph. In the tree-based GP, the data flow is more rigidly determined by the tree structure



y = f[0] = (v[1] / 3) + v[4]

Fig. 3 Comparison of the GP program structures (a) Tree-based GP, (b) LGP (after Alavi *et al.* 2010b)

```
void LGP (double r[5])
{ ...
r[0] = r[5] + 70;
r[5] = r[0] - 50;
if (r[1] > 0)
if (r[5] > 2)
r[4] = r[2] × r[1];
r[2] = r[5] + r[4];
r[0] = sin(r[2]);
}
```

Fig. 4 An excerpt of a linear genetic program

of the program (Brameier and Banzhaf 2001, Alavi *et al.* 2010b).

In the LGP system described here, an individual program is interpreted as a variable-length sequence of simple C instructions. The instruction set or function set of LGP consists of arithmetic operations, conditional branches, and function calls. The terminal set of the system is composed of variables and constants. The instructions are restricted to operations that accept a minimum number of constants or memory variables, called registers (r), and assign the result to a destination register, e.g., $r_0 := r_1 + 1$. A part of a linear genetic program in C code is represented in Fig. 4. In this figure, register $r[0]$ holds the final program output.

Automatic Induction of Machine code by Genetic Programming (AIMGP) is a particular form of LGP. The programs evolved by AIMGP are stored as linear strings of native binary machine code and directly executed by the processor during fitness calculation. The absence of an interpreter and complex memory handling results in a significant speedup in the AIMGP execution compared to tree-based GP (Brameier and Banzhaf 2007). This machine-code-based LGP approach searches for the computer program and the constants at the same time. Here are the steps the machine-code-based LGP follows for a single run (Francone and Deschaine 2004, Brameier and Banzhaf 2007):

I. Initializing a population of randomly generated programs and calculating their fitness values.
II. Running a Tournament. In this step four programs are selected from the population randomly. They are compared and based on their fitness, two programs are picked as the winners and two as the losers.
III. Transforming the winner programs. After that, two winner programs are copied and transformed probabilistically as follows:

- Parts of the winner programs are exchanged with each other to create two new programs (crossover); and/or
- Each of the tournament winners are transformed randomly to create two new programs (mutation).

IV. Replacing the loser programs in the tournament with the transformed winner programs. The winners of the tournament remain without change.

V. Repeating steps two through four until termination or convergence conditions are satisfied.

A basic representation of the LGP algorithm is presented in Fig. 5.

Crossover occurs between instruction blocks. Fig. 6 demonstrates a two-point linear crossover used in LGP for recombining two tournament winners. As it is seen, a segment of random position and arbitrary length is selected in each of the two parents and exchanged. If one of the two children would exceed the maximum length, crossover is aborted and restarted with exchanging equally sized segments (Brameier and Banzhaf 2001). The mutation operation occurs on a single instruction

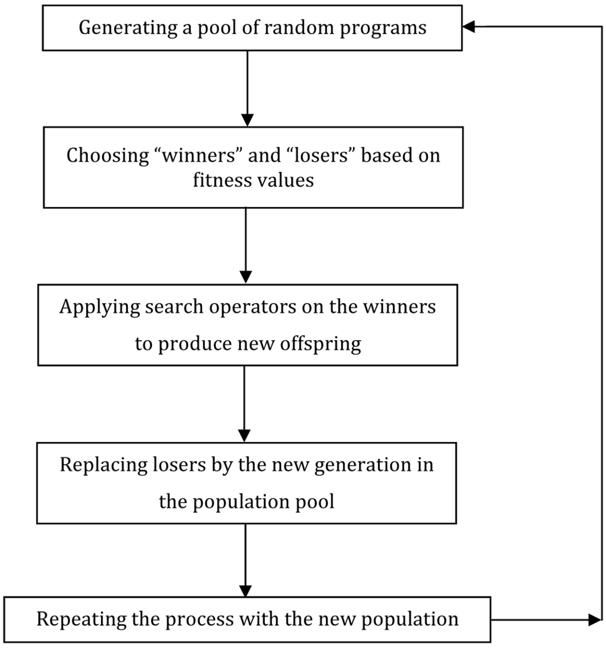


Fig. 5 A basic representation of the LGP algorithm (Regmi *et al.* 2004)

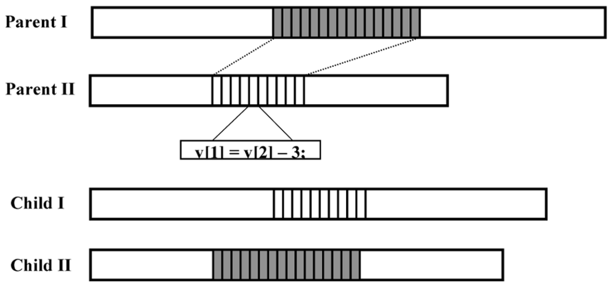


Fig. 6 Crossover in LGP

set. Inside instructions, mutation randomly replaces the instruction identifier (a variable or a constant) by equivalents from valid ranges. Comprehensive descriptions of the basic parameters used to direct a search for a linear genetic program can be found in (Brameier and Banzhaf 2007). According to Francone and Deschaine (2004), the LGP system can be regarded as an efficient modeling tool for complex problems for several reasons including:

- Its speed permits conducting many runs in realistic timeframes. This leads to deriving consistent, high-precision models with little customization;
- It is well-designed to prevent overfitting and to evolve robust solutions; and
- The solutions evolved by the LGP system can be executed very quickly when called by an optimizer.

3. Modeling of shear strength of SFRC beams

The main purpose of this study is to derive an alternative formulation for the shear strength of SFRCB without stirrupsusing the LGP approach. The most important factors commonly used in the previous models and codes (Zsutty 1971, Narayanan *et al.* 1987, Sharma 1986, Mansur *et al.* 1986, Ashour *et al.* 1992, Li *et al.* 1992, Swamy *et al.* 1993, Shin *et al.* 1994, Khuntia *et al.* 1999, Kwak *et al.* 2002, Adhikary and Mutsuyoshi 2006) were employed as the predictor variables. Consequently, formulation of the shear strength was considered to be as follows

$$v_u = f\left(\frac{a}{d}, f_c', F, \rho, \omega, e, v_b, f_t', f_{spfc}'\right) \tag{12}$$

where,

v_u = Average shear stress at shear failure, MPa
a/d = Shear span-depth ratio
a = shear span, mm
b = Beam width, mm
d = Effective depth of beam, mm
f_c' = Cylinder compressive strength of concrete, MPa
ρ = Flexural steel reinforcement ratio, %
F = Fiber factor $(d_f(L_f/D_f)\ V_f)$
L_f/D_f = Fiber aspect ratio
L_f = Fiber length, mm
D_f =Fiber diameter, mm
V_f = volume fraction of steel fibers, %
d_f = bond factor: 0.50 for round fibers, 0.75 for crimped fibers, and 1.00 for indented fibers (Narayanan and Darwish 1987)
ω = Reinforcement factor $(\rho(1 + 4F))$
e = Arch action factor: 1.0 for $a/d > a/d_{transition}$, and $a/d_{transition} \times d/a$ for $a/d \le a/d_{transition}$ ($a/d_{transition} = 3$)
v_b = 0.41 τF, MPa
τ = Average fiber matrix interfacial bond stress, taken as 4.15 MPa, based on recommendations of Swamy *et al.* (1974)
f_t' = Splitting tensile strength (0.79 f_c'), MPa
f_{spfc} = Computed value of split-cylinder strength of fiber concrete

Table 3 The considered combinations of the predictor variables

Combination #	Variables
1	$a/d, f_c', F, \rho, e$
2	$a/d, f_c', F, \rho$
3	$a/d, f_c', \omega, e$
4	$a/d, f_c', \omega$
5	$a/d, f_c', v_b, \rho, e$
6	$a/d, f_c', v_b, \rho$
7	$a/d, f_t', F, \rho, e$
8	$a/d, f_t', F, \rho$
9	$a/d, f_t', \omega, e$
10	$a/d, f_t', \omega$
11	$a/d, f_t', v_b, \rho, e$
12	$a/d, f_t', v_b, \rho$
13	$a/d, f_{spfc}, \rho, e$
14	$a/d, f_{spfc}, \rho$
15	$a/d, f_{spfc}, v_b, \rho, e$
16	$a/d, f_{spfc}, v_b, \rho$

$$f_{spfc} = \frac{f_{cuf}}{(20-F)} + 0.7 + 1.0\sqrt{F} \text{ (MPa)} \tag{13}$$

where f_{cuf} (MPa) is cube strength of fiber concrete. The above formulation of v_u incorporates the contributions of various geometrical and mechanical factors influencing the shear strength of SFRCB. To formulate the shear strength of SFRC beams, several models were developed and analyzed using different combinations of the predictor variables. Table 3 summarizes the considered input combinations. The selection of the parameters for each of the combination types was based on the shape of the classical relationships (see Table 2) and also on the basis of a trial study. Although a maximum of five parameters was considered for the cases, it was tried to include the effect of several variables, shown by Eq. (12), in each combination. As an example, the sixth combination consists of a/d, f_c', ρ, and v_b, where v_b is a function of τ and F and F is itself a function of d_f, L_f, D_f, and V_f. Thus, this combination incorporates the effect of several influencing variables, i.e., a, d, f_c', ρ, τ, F, d_f, L_f, D_f, and V_f. The same is true for other combinations. The advantage of using such combinations is deriving much simpler models since fewer parameters are involved in the model development.

3.1 Experimental database

The comprehensive database used for developing the models was gathered through an extensive literature review. The collected data included 213 tests results for the SFRC beams without shear reinforcement (Batson 1972, Roberts and Ho 1982, Jindal 1984, Swamy and Bahia 1985, Uomoto *et al.* 1986, Kadir and Saeed 1986, Mansur *et al.* 1986, Kaushik *et al.* 1987, Murthy and Venkatacharyulu 1987, Lim *et al.* 1987, Narayanan and Darwish 1987, 1988, Li *et al.* 1992, Tan *et*

Table 4 Descriptive statistics of the independent variables used in the model development

Parameter	a (mm)	d (mm)	f_c' (MPa)	ρ (%)	L_f/D_f	V_f (%)	d_f	v_u (MPa)
Mean	633.142	203.061	47.221	2.811	79.427	0.827	0.728	3.929
Standard Error	39.344	7.358	1.377	0.089	1.677	0.030	0.010	0.178
Median	508	175	40.85	2.9	67	0.75	0.75	3.13
Mode	392	127	33.21	3.08	100	1	0.75	2.43
Standard Deviation	574.203	107.390	20.100	1.305	24.478	0.435	0.139	2.600
Sample Variance	329709.521	11532.510	404.023	1.704	599.188	0.190	0.019	6.761
Kurtosis	10.992	3.261	1.534	7.761	-0.547	-0.064	0.198	3.825
Skewness	3.202	1.853	1.352	2.031	0.504	0.592	-0.038	1.919
Range	3517.870	490.000	90.900	9.500	108.333	1.780	0.500	13.610
Minimum	120	80	20.6	1	25	0.22	0.5	1.44
Maximum	3637.87	570	111.5	10.5	133.333	2	1	15.05
Confidence Level (95.0%)	77.555	14.505	2.715	0.176	3.306	0.059	0.019	0.351

al. 1992, Ashour *et al.* 1992, Swamy *et al.* 1993, Imam 1994, Shin *et al.* 1994, Adebar *et al.* 1997, Oh *et al.* 1998, Casanova 1999, Noghabai 2000, Kwak *et al.* 2002, Cucchiara *et al.* 2004). A summary of the number of tests employed herein and the related references is also given in Table 1. The ranges and statistics of the independentvariables involved in the model development are given in Table 4. As can be observed from this table, the database contains experimental results for both normal-strength concrete (NSC) and high-strength concrete (HSC) beams. To visualize the distribution of the samples, the data are presented by frequency histograms (Fig. 7). The histograms show a visual impression of the distribution of the experimental data. The cumulative number of the observations in the considered bins is illustrated by cumulative percentage line. As can be observed from Fig. 7, the distributions of the predictor variables are not uniform. The models would most probably provide better predictions for the cases where the densities of the variables are higher.

Overfitting is one of the principal problems in machine learning generalization. It is a case in which the error on the learning set is driven to a very small value, but when new data is presented to the model, the error is large. An efficient approach to prevent overfitting is to test other individuals from the run on a validation set to find a better generalization (Banzhaf *et al.* 1998). This technique was used in this study for improving the generalization of the models. For this purpose, the available data sets were randomly divided into learning, validation and testing subsets. The learning data were used for training (genetic evolution). The validation data were used to specify the generalization capability of the evolved programs on data they did not train on (model selection). In other words, the learning and validation data sets were used to select the best evolved programs and included in the training process. Thus, they were categorized into one group referred to as training data. The testing data were finally used to measure the performance of the models obtained by LGP on data that played no role in building the models. This technique provides decent results as long as the models perform well on the learning data sets (Banzhaf *et al.* 1998). In order to obtain a consistent data division, several combinations of the training and testing sets were considered. Of the 213 data, 107 data vectors were used for the learning process and 64 data were taken as the validation data. The remaining 42 sets (20%) were used for the testing of the derived models.

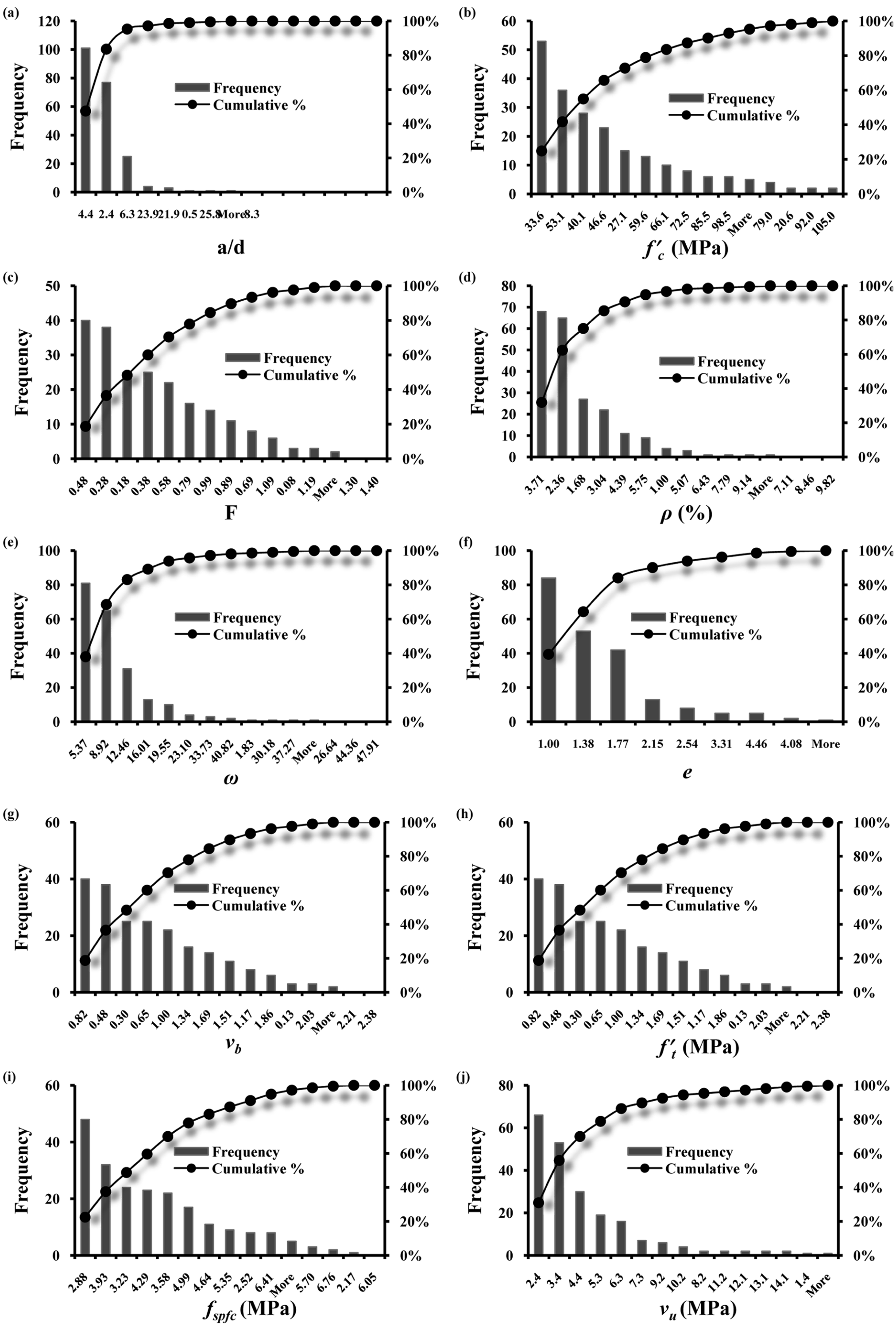


Fig. 7 Histograms of the parameters used in the model development

3.2 Model development using LGP

The available database was used for developing the LGP prediction models relating v_u to a/d, f_c', F, ρ, e, ω, e, v_b, f_t', and f_{spfc}. Various parameters involved in the LGP algorithm are shown in Table 5. The parameter selection will affect the model generalization capability of LGP. The number of programs in the population that LGP will evolve is set by the population size. A run will take longer with a larger population size. The maximum number of tournaments sets the outer limit of the tournaments that will occur before the program terminates the run. The proper number of population and tournaments depends on the number of possible solutions and complexity of the problem. Mutation and crossover rates are the probabilities that an offspring will be subjected to mutation and crossover, respectively. The lengths of the evolved programs in runs were controlled by initial and maximum program size parameters. The initial program size parameter sets the size of the programs in the first population at the start of each run. The maximum program size parameter sets the maximum length of the other programs evolved during each run.

Several runs were conducted to come up with a parameterization of LGP that provided enough robustness and generalization to solve the problem. The LGP parameters were changed for different runs. Three levels were set for the population size and two levels were considered for the crossover rate. The success of the LGP algorithm usually increases with increasing the initial and maximum program size parameters. In this case, the complexity of the evolved functions increase and the speed of the algorithm decreases. These parameters are measured in bytes. The initial and maximum program sizes were respectively set to optimal values of 80 and 256 bytes as tradeoffs between the running time and the complexity of the evolved solutions. One level was considered for the other parameters based on some previously suggested values (Baykasoglu *et al.* 2008, Gandomi *et al.* 2010b, Alavi and Gandomi 2011) and also after making several preliminary runs and observing the performance behavior. The number of demes presented in Table 5 is related to the way that the population of programs is divided. Note that demes are semi-isolated subpopulations that evolution proceeds faster in them in comparison to a single population of equal size (Brameier and Banzhaf

Table 5 Parameter settings for the LGP algorithm

Parameter	Settings
Population size	500, 1500, 3000
Maximum number of tournaments	9000000
Crossover rate (%)	50, 95
Homologous crossover (%)	95
Mutation rate (%)	90
Block mutation rate (%)	30
Instruction mutation rate (%)	30
Data mutation rate (%)	40
Maximum program size	256
Initial program size	80
Function set	$+, -, \times, /, \sqrt{}, \sin, \cos$
Number of demes	20
Fitness function	Squared error

2007). In this study, four basic arithmetic operators (+, −, ×, /) and basic mathematical functions ($\sqrt{}$, sin, cos) were utilized to get the optimum LGP models. There are $3 \times 2 = 6$ different combinations of the parameters. All of these combinations were tested and 10 replications for each combination were carried out. This makes 60 runs for each of the combinations of the predictor variables shown in Table 3. Therefore, the overall number of runs was equal to 60 × 16 (number of the input combinations) = 960. A fairly large number of tournaments were tested on each run to find models with minimum error. For each case, the program was run until there was no longer significant improvement in the performance of the models or the runs terminated automatically. Each run was observed while in progress for overfitting. In checking for overfitting, situations were checked in which the fitness of the samples for the training of LGP was negatively correlated with the fitness on the validation data sets. To evaluate the fitness of the evolved program, the average of the squared raw errors was used. For the runs showing signs of overfitting, the LGP parameters were progressively changed so as to reduce the computational power available to the LGP algorithm until observed overfitting was minimized. The resulting run was then accepted as the production run. The programs with the best performance on both of the training and validation data sets were finally selected as the outcomes of each run. For the LGP-based analysis, the Discipulus software (Conrads *et al.* 2004) was used which works on the basis of the AIMGP platform.

3.3 LGP-based formulation for shear strength of SFRCB

An extensive trial study was performed to select the most relevant input parameters for the LGP models. Several LGP models were developed using different combinations of the input parameters. The best model was chosen on the basis of a multi-objective strategy as follows:

i. The simplicity of the model, although this was not a predominant factor.
ii. Providing the best fitness value on the training (learning and validation) set of data.
iii. Providing the best fitness value on a test set of unseen data.

The first objective can be controlled by the user through the parameter settings (e.g., program size for LGP). For the other objectives, the following objective function (OBJ) was constructed as a measure of how well the model predicted output agrees with the experimentally measured output. The selection of the best LGP model was deduced by the minimization of the following function (Gandomi *et al.* 2010c)

$$OBJ = \left(\frac{No_{.Train} - No_{.Test}}{No_{.All}}\right)\frac{MAE_{Train}}{R^2_{Train}} + \frac{2No_{.Test}}{No_{.All}}\frac{MAE_{Test}}{R^2_{Test}} \tag{14}$$

where $No._{Train}$, $No._{Test}$, and $No._{All}$ are respectively the number of training, testing and whole number of data; Correlation coefficient (R) and mean absolute error (MAE) are given in the form of formulas as follows

$$R = \frac{\sum_{i=1}^{n}(h_i - \bar{h}_i)(t_i - \bar{t}_i)}{\sqrt{\sum_{i=1}^{n}(h_i - \bar{h}_i)^2 \sum_{i=1}^{n}(t_i - \bar{t}_i)^2}} \tag{15}$$

$$MAE = \frac{1}{2}\sum_{i=1}^{n}|h_i - t_i| \tag{16}$$

in which h_i and t_i are respectively actual and calculated outputs for the i^{th} output, $\bar{h}_i$ is the average of the actual outputs, and n is the number of sample. The R value alone is not an appropriate indicator of prediction accuracy of a model as by shifting the output values of a model equally, R will not change significantly. The constructed objective function takes into account the changes of R and MAE together. Higher R values and lower MAE values result in lowering OBJ and, consequently, indicate a more precise model. In addition, the above function considers the effects of different data divisions for the training and testing sets.

Table 6 summarizes the optimal developed LGP models with their selected input parameters and the corresponding OBJ values. As can be seen in this table, the best LGP model for predicting the shear strength of SFRCB was built using a/d, f_c', ρ, and v_b (Model Type 6). These four parameters seem to be sufficient representatives of the geometrical parameters and mechanical properties of SFRCB. The best LGP single program (Model 6) obtained at the end of training in C++ is given in Appendix A. This program can be run in C++ environment. The resulting code may be linked to the optimizer and compiled or it may be called from the optimization routines (Deschaine 2000). To facilitate the use of the derived model, it was converted into a functional representation by successive replacements of variables starting with the last effective instruction (Oltean and Grossan 2003). The optimal LGP-based formulation of the shear strength of SFRCB, v_u is as follows

$$v_u \text{ (MPa)} = \frac{2d}{a}(\rho f_c' + v_b) + \frac{d}{2a}\frac{\rho}{(288\rho - 11)^4} + 2 \tag{17}$$

The population size, crossover rate and maximum number of tournaments for the optimal run were respectively equal to 500, 95% and 3460800. This run took 7 min and 13 s on a Pentium 4 personal computer with 3.00 GHz of processor speed and 1 Gb of memory. The number of the computer programs evolved and evaluated by the LGP algorithm during the conducted run was equal to 4312860. A comparison of the experimental and predicted shear strength using the LGP model is shown in Fig. 8.

Table 6 Summary of the LGP optimal solutions developed with various input parameters

Combin-ation #	1	2	3	4	5	**6**	7	8	9	10	11	12	13	14	15	16
OBJ	0.980	0.903	1.001	0.986	0.932	**0.884**	0.930	0.889	1.021	0.971	0.939	0.911	0.996	0.994	0.943	0.941
Rank	11	3	15	12	6	**1**	5	2	16	10	7	4	14	13	9	8

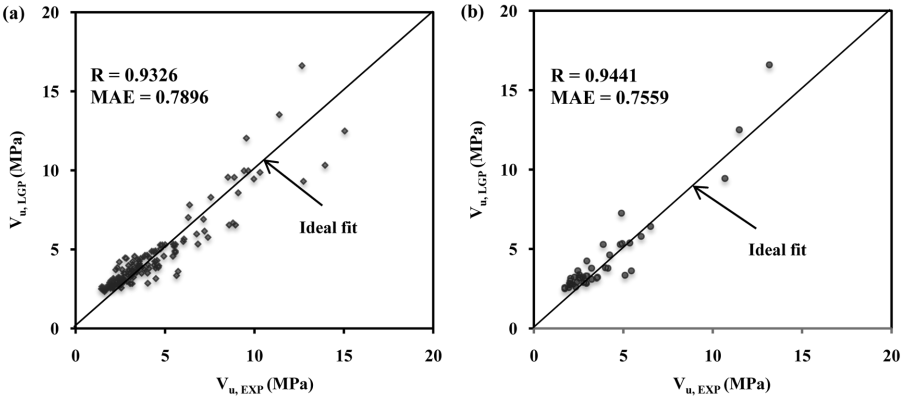


Fig. 8 Experimental versus predicted shear strength using the LGP model (a) training data, (b) testing data

4. Performance analysis, model validity, and comparative study

A precise formula for the shear strength of SFRCB was obtained by means of LGP. Performance statistics of the LGP model and available codes on the entire database are summarized in Table 7. Based on a logical hypothesis (Smith 1986), if a model gives R > 0.8, and the error value (e.g., MAE) is at the minimum, there is a strong correlation between the predicted and measured values. The model can therefore be judged as very good. It can be observed from Table 6 and Fig. 8 that the LGP model with high R and low MAE, and therefore low OBJ, values predicts the target values to an acceptable degree of accuracy. Meanwhile, it is noteworthy that the MAE values are not only low but also as similar as possible for the training and validation sets, which suggests that the proposed model has both predictive ability (low values) and generalization performance (similar values) (Pan *et al.* 2009).

It is known that the models derived using the ANNs, GP, or other soft computing tools, in most cases, have a predictive capability within the data range used for their development. This is because of the nature of these techniques which distinguishes them from the other conventional techniques. Thus, the amount of data used for the modeling process is an important issue, as it bears heavily on the reliability of the final models. To cope with this limitation, Frank and Todeschini (1994) argue that the minimum ratio of the number of objects over the number of selected variables for model acceptability is 3, but often a safer value of 5 is more reasonable. In the present study, this ratio is

Table 7 Statistical parameters of the LGP model for external validation

Item	Formula	Condition	LGP
1	Eq. (15)	$R > 0.8$	0.944
2	$k = \frac{\sum_{i=1}^{n}(h_i \times t_i)}{h_i^2}$	$0.85 < K < 1.15$	0.890
3	$k' = \frac{\sum_{i=1}^{n}(h_i \times t_i)}{t_i^2}$	$0.85 < K' < 1.15$	1.087
4	$m = \frac{R^2 - Ro^2}{R^2}$	$m < 0.1$	-0.076
5	$n = \frac{R^2 - Ro'^2}{R^2}$	$n < 0.1$	-0.092
6	$R_m = R^2 \times (1 - \sqrt{\lvert R^2 - Ro^2 \rvert})$	$R_m > 0.5$	0.659
where	$Ro^2 = 1 - \frac{\sum_{i=1}^{n}(t_i - h_i^o)^2}{\sum_{i=1}^{n}(t_i - \bar{t}_i)^2}$, $h_i^o = k \times t_i$		0.959
	$Ro'^2 = 1 - \frac{\sum_{i=1}^{n}(h_i - t_i^o)^2}{\sum_{i=1}^{n}(h_i - \bar{h}_i)^2}$, $t_i^o = k' \times h_i$		0.973

much higher and is at least equal to 213/5 = 42.6. Furthermore, new criteria recommended by Golbraikh and Tropsha (2002) were checked for external verification of the LGP model on the testing data sets. It is suggested that at least one slope of regression lines (k or k') through the origin should be close to 1. Also, the performance indexes of m and n should be lower than 0.1. Recently, Roy and Roy (2008) introduced a confirm indicator of the external predictability of models (R_m). For $R_m > 0.5$, the condition is satisfied. Either the squared correlation coefficient (through the origin) between predicted and experimental values (Ro^2), or the coefficient between experimental and predicted values (Ro'^2) should be close to 1. The considered validation criteria and the relevant results obtained by the models are presented in Table 7. As it is seen, the derived model satisfies the required conditions. The validation phase ensures the proposed model is strongly valid, has the prediction power and is not a chance correlation.

Besides, comparisons of the shear strength predictions made by the LGP model and various empirical models found in the literature (Swamy *et al.* 1993, Sharma 1986, Narayanan *et al.* 1987, Ashour *et al.* 1992, Kwak *et al.* 2002, Zsutty 1971, Li *et al.* 1992, Khuntia *et al.* 1999, Shin *et al.* 1994, Mansur *et al.* 1986) for the entire database are presented in Table 8. Fig. 9 visualizes the histogram plots of the ratio of the experimental to predicted shear strength values. The considered performance measures were the errors in the predictions (R and MAE), as well as the values of standard deviation (SD), coefficient of variation (COV) and mean (Mean) of the ratio of the observed to predicted values. As it is seen, the results provided by the proposed model (R = 0.932, MAE = 0.783, SD = 0.208, COV = 0.235) are significant improvement over the other existing models. Considering the Mean values, the best results are obtained by Eq. (6) proposed by Kwak *et al.* (2002).

In addition to its reasonable accuracy, the LGP model is remarkably simple and can be used for design practice via hand calculations. In the conventional statistical analyses, a linear relationship is often assumed between the outcome and the predictor variables, which is not always true. In most cases, the best models developed using the commonly used statistical approaches are obtained after controlling just some equations established in advance. Thus, such models cannot efficiently

Table 8 Overall performances of the shear strength prediction models

Model	$V_{u,\,Exp.}$ vs $V_{u,\,Pre.}$		$V_{u,\,Exp.} / V_{u,\,Pre.}$	
	R	MAE	SD	COV
Swamy *et al.* (1993), Eq. (1)	0.756	1.669	0.802	0.476
Sharma (1986), Eq. (2)	0.686	1.369	0.664	0.504
Narayanan and Darwish (1987), Eq. (3)	0.725	1.482	0.365	0.310
Ashour *et al.* (1992), Eq. (4)	0.789	1.377	0.419	0.341
Ashour *et al.* (1992), Eq. (5)	0.820	1.078	0.384	0.354
Kwak *et al.* (2002), Eq. (6)	0.823	0.935	0.316	0.309
Kwak *et al.* (2002), Eq. (7)	0.812	1.082	0.313	0.303
Li *et al.* (1992), Eq. (8)	0.684	1.333	0.635	0.508
Khuntia *et al.* (1999), Eq. (9)	0.795	1.708	0.741	0.425
Shin *et al.* (1994), Eq. (10)	0.782	1.064	0.304	0.273
Mansur *et al.* (1986), Eq. (11)	0.524	1.746	0.963	0.550
LGP (This work), Eq. (17)	0.932	0.783	0.208	0.235

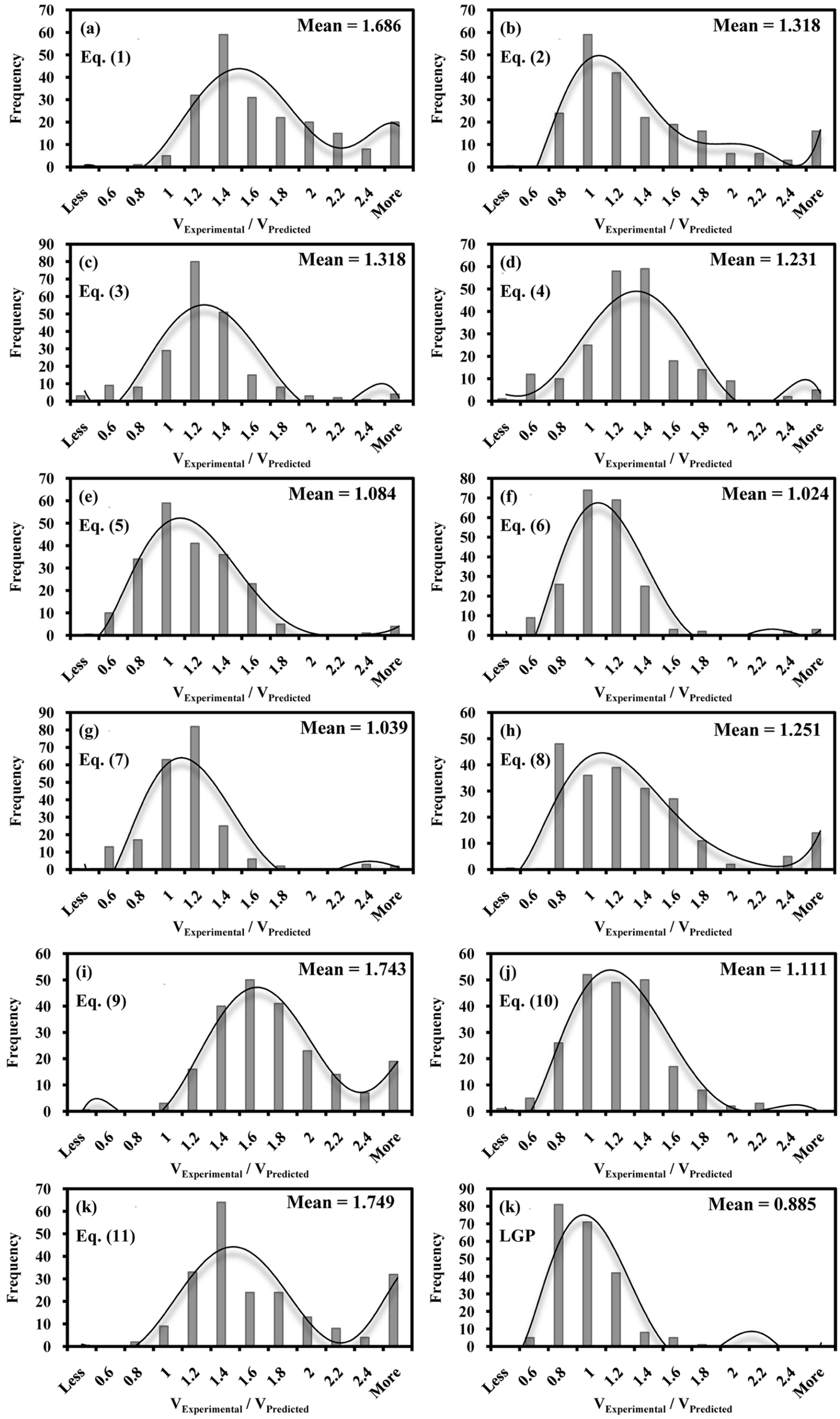


Fig. 9 Histogram of the experimental/predicted shear strength values using different models

consider the interactions between the dependent and independent variables. On the other hand, LGP introduces completely new characteristics and traits. One of the major advantages of the LGP approach over the traditional analyses is its ability to derive explicit relationships for the shear strength without assuming prior forms of the existing relationships. The best solutions (equations) evolved by this technique are determined after controlling numerous preliminary models, even billions of linear and nonlinear models. For instance, the proposed model for the estimation of the shear strength of SFRCB was selected among a total of 1260346068 programs evolved and evaluated by the LGP method during the 960 conducted runs. The results indicate that the optimal LGP model has efficiently captured the interactions between the input and output variables. It should be noted that the obtained equation is in its final form and no other simplifications can be applied to it. As more data become available, including those for other types of SFRC beams, the proposed model can be improved to make more accurate predictions for a wider range. The user physical insight and the shape of the classical models can be regarded in making propositions on the elements in the evolved functions and on their structures.

However, one of the main goals of introducing expert systems, such as the GP-based approaches, into the design processes is better handling of the information in the pre-design phase. The initial steps of design are based on imprecise and incomplete information about the features and properties of targeted output or process (Kraslawski *et al.* 1999). Nevertheless, it is idealistic to have some initial estimates of the outcome before performing any extensive laboratory or field work. The LGP approach employed in this research is based on the data alone to determine the structure and parameters of the model. Thus, the derived models are considered to be mostly valid for use in preliminary design stages. For more reliability, the results of the LGP-based analyses are suggested to be verified by those obtained using deterministic methods.

5. Sensitivity and parametric analyses

Sensitivity analysis is of utmost concern for selecting the important input variables. The contribution of each input parameter in the best LGP model (Model 6) was evaluated through a sensitivity analysis. In order to evaluate the importance of the input parameters, their frequency values (Francone 2001) were obtained. A frequency value equal to 100% for an input indicates that this input variable has been appeared in 100% of the best thirty programs evolved by LGP. This is a common approach in the GP-based analyses (e.g., Gandomi *et al.* 2010c). The frequency values of the predictor variables are presented in Fig. 10. According to this figure, the shear strength is less sensitive to ρ than the other variables. In addition to frequencies, Fig. 10 presents the average impact of removing all instances of each input from the best thirty programs of the project. A value of 100% represents the largest impact value possible. The greater the value, the more impact removal had. As it is seen, the shear strength is more influenced by a/d and f_c' considering their higher removal impacts in comparison with ρ and v_b.

For further verification of the LGP-based prediction model, a parametric analysis was performed in this study. The parametric analysis investigates the response of the predicted shear strength from the LGP model to a set of hypothetical input data generated over the ranges of the minimum and maximum data used for the model training. The methodology is based on the change of only one input variable at a time while the other variables are kept constant at the average values of their entire data sets. A set of synthetic data for the single varied parameter is generated by increasing the

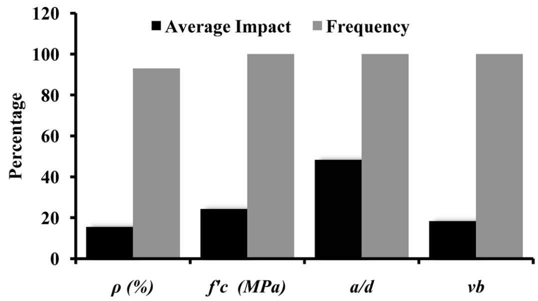


Fig. 10 Contributions of the predictor variables in the LGP model

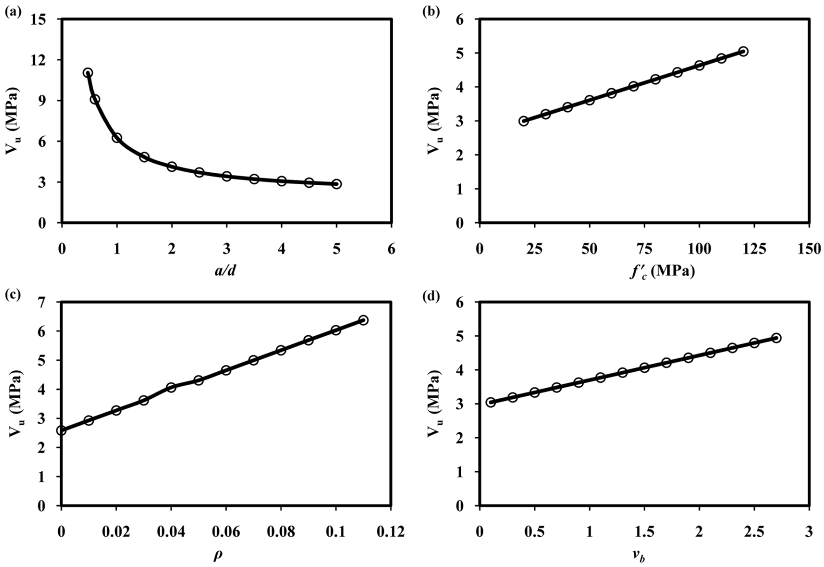


Fig. 11 Shear strength parametric analysis in the LGP model

value of this in increments. These inputs are presented to the prediction equation and the shear strength is calculated. This procedure is repeated using another variable until the model response is tested for all input variables. The robustness of the design equations is determined by examining how well the predicted values agree with the underlying physical behavior of the investigated system (Kuo *et al.* 2009). Fig. 11 presents the tendency of the shear strength predictions to the variations of the design parameters, *a/d*, f_c', ρ, and v_b. As it is seen, the shear strength continuously decreases with increasing *a/d and*increases due to increasing f_c', ρ, and v_b. The parametric analysis results are soundly expected cases from structural engineering viewpoint (Kwak *et al.* 2002, Ashour *et al.* 1992, Imam *et al.* 1994, Narayanan and Darwish 1987, Noghabai 2000).

6. Conclusions

A constitutive model was developed to assess the shear resistance of SFRC beams without stirrups using the LGP paradigm. The following conclusions can be derived from the results presented in this research:

- The LGP-based model accurately predicts the shear strength of SFRCB. The validity of the model was tested for a part of test results beyond the training data domain. Furthermore, the LGP prediction model efficiently satisfies the conditions of different criteria considered for its external validation. The validation phases confirm the efficiency of the model for its general application to the shear strength estimation of SFRCB.
- The proposed model simultaneously takes into account the role of several important factors representing the behavior of the shear strength. The final explanatory variables (a/d, f_c', ρ, v_b) were selected after developing different models with different combinations of the input parameters.
- The model can be used for practical pre-planning and design purposes since it was derived from tests on beams with a wide range of geometrical and mechanical properties, including HSC and NSC beams.
- The proposed model produces considerably better outcomes than several prediction equations found in the literature.
- An observation from the results of the sensitivity analysis is that the shear strength is more affected by a/d and f_c' than ρ and v_b.
- A general criticism about the GP-based models is that they are only randomly formed functions which are not based on physical processes. This ambiguity was illuminated by the parametric analysis. Based on the obtained results, the derived model is a meaningful combination of the predictor variables as it efficaciously incorporates the underlying physical relations governing the system.
- The models derived using LGP are substantially different from the conventional models developed based on first principles (e.g., elasticity and plasticity theories). A distinctive feature of the LGP-based models is that they are based on experimental data rather than on assumptions made in developing the conventional models.

Acknowledgements

This research is partly supported by the new faculty start-up fund. Authors are grateful of their support. The authors are also thankful to Prof. Marc O. Eberhard (University of Washington) for providing a part of the experimental database.

References

ACI Committee 544 (1988), "Design considerations for steel fiber-reinforced concrete", *ACI Struct. J.,* **85**(5), 563-580.

Adebar, P., Mindess, S.S., Pierre, D. and Olund, B. (1997), "Shear tests of fiber concrete beams without stirrups", *ACI Struct. J.,* **94**(1), 68-76.

Adhikary, B.B. and Mutsuyoshi, H. (2006), "Prediction of shear strength of steel fiber RC beams using neural networks", *Constr. Build. Mater.,* **20**, 801-811.

Alavi, A.H., Gandomi, A.H. and Heshmati, A.A.R. (2010b), "Discussion on soft computing approach for real-time estimation of missing wave heights", *Ocean Eng.,* **37**(13), 1239-1240.

Alavi, A.H. and Gandomi, A.H. (2011), "A robust data mining approach for formulation of geotechnical

engineering systems", *Int. J. Comput. Aid. Meth. Eng.-Eng. Computat.*, **28**(3).
Alavi, A.H., Gandomi, A.H., Sahab, M.G. and Gandomi, M. (2010a), "Multi expression programming: a new approach to formulation of soil classification", *Eng. Comput-Germany,* **26**(2), 111-118.
Ashour, S.A., Hasanain, G.S. and Wafa, F.F. (1992), "Shear behavior of high-strength fiber reinforced concrete beams", *ACI Struct. J.,* **89**(2), 176-184.
Bäck, T. (1996), *Evolutionary Algorithms in Theory and Practice: Evolution Strategies, Evolutionary Programming, Genetic Algorithms*, Oxford University Press, USA.
Banzhaf, W., Nordin, P., Keller, R. and Francone, F. (1998), *Genetic Programming-an Introduction, On the Automatic Evolution of Computer Programs and Its Application*, Dpunkt/Morgan Kaufmann, Heidelberg/San Francisco.
Batson, G., Jenkins, E. and Spatney, R. (1972), "Steel fibers as shear reinforcement in beams", *ACI J.,* **69**(10), 640-644.
Baykasoglu, A., Gullub, H., Canakcı, H. and Ozbakır, L. (2008), "Prediction of compressive and tensile strength of limestone via genetic programming", *Expert Syst. Appl.,* **35**(1-2), 111-123.
Brameier, M. and Banzhaf, W. (2001), "A comparison of linear genetic programming and neural networks in medical data mining", *IEEE T. Evolut. Comput.,* **5**(1), 17-26.
Brameier, M. and Banzhaf, W. (2007), *Linear Genetic Programming*, Springer Science + Business Media, New York.
Casanova, P. and Rossi, P. (1999), "High-strength concrete beams submitted to shear: steel fibers versus stirrups", *Structural Applications of Fiber Reinforced Concrete, SP-182*, (Eds. N. Banthia, C. MacDonald and P. Tatnall), American Concrete Institute, Farmington Hills, Mich., 53-67.
Chanh, V.N. (2004), "Steel fiber-reinforced concrete", Faculty of Civil Engineering Ho chiminh City University of Technology, Seminar Material, 108-116.
Conrads, M., Dolezal, O., Francone, F.D. and Nordin, P. (2004), "Discipulus Lite^{TM}-fast genetic programming based on AIM learning technology", Register Machine Learning Technologies Inc., Littleton, CO.
Cucchiara, C., Mendola, L.L. and Papia, M. (2004), "Effectiveness of stirrups and steel fibres as shear reinforcement", *Cement Concrete Comp.,* **26**, 777-786.
Deschaine, L.M. (2000), "Using genetic programming to develop a C/C++ simulation model of a waste incinerator science applications", Draft Technical Report, International Corp.
Francone, F. (2000), *Discipulus^{TM} Owner's Manual, Version 3.0. Register Machine Learning Technologies*, Littleton, CO.
Francone, F.D. and Deschaine, L.M. (2004), "Extending the boundaries of design optimization by integrating fast optimization techniques with machine-code-based, linear genetic programming", *Inform. Sciences,* **161**, 99-120.
Frank, I.E. and Todeschini, R. (1994), *The Data Analysis Handbook*, Elsevier, Amsterdam, The Nederland.
Gandomi, A.H., Alavi, A.H. and Sahab, M.G. (2010b), "New formulation for compressive strength of CFRP confined concrete cylinders using linear genetic programming", *Mater. Struct.,* **43**(7), 963-983.
Gandomi, A.H., Alavi, A.H., Sahab, M.G. and Arjmandi, P. (2010c), "Formulation of elastic modulus of concrete using linear genetic programming", *J. Mech. Sci. Technol.,* **24**(6), 1273-1278.
Gandomi, A.H., Alavi, A.H. and Yun, G. (2010a), "Formulation of uplift capacity of suction caissons using multi expression programming", *KSCE J. Civil Eng.*, **15**(2), 363-373.
Gencoglu, M. (2007), "The effects of stirrups and the extents of regions used sfrc in exterior beam-column joints", *Struct. Eng. Mech.,* **27**(2), 223-241.
Golbraikh, A. and Tropsha, A. (2002), "Beware of q2", *J. Mol. Graph. Model.,* **20**(4), 269-276.
Guven, A., Azamathullab, H.M. and Zakaria, N.A. (2009), "Linear genetic programming for prediction of circular pile scour", *Ocean Eng.,* **36**(12-13), 985-991.
Imam, M., Vandewalle, L. and Mortelmans, F. (1994), "Shear capacity of steel fiber high-strength concrete beams", *High-Performance Concrete, SP-149,* (Ed. V.M. Malhotra), American Concrete Institute, Farmington Hills, Mich., 227-241.
Jindal, R.L. (1984), "Shear and moment capacities of steel fiber reinforced concrete beams", *Fiber Reinforced Concrete, SP-81,* (Ed. G.C. Hoff), American Concrete Institute, Farmington Hills, Mich., 1-16.
Johari, A., Habibagahi, G. and Ghahramani, A. (2006), "Prediction of soil-water characteristic curve using

genetic programming", *J. Geotech. Geoenviron.,* **132**(5), 661-665.
Kadir, M.R.A. and Saeed, J.A. (1986), "Shear strength of fibre reinforced concrete beams", *J. Eng. Technol.,* **4**(3), 98-112.
Kaushik, S.K., Gupta, V.K. and Tarafdar, N.K. (1987), "Behavior of fiber reinforced concrete beams in shear", *Proceedings of the International Symposium on Fiber Reinforced Concrete*, Madras, India.
Khuntia, M., Stojadinovic, B. and Goel, S.C. (1999), "Shear strength of normal and high-strength fiber reinforced concrete beams without stirrups", *ACI Struct. J.,* **96**(2), 282-289.
Koza, J. (1992), "Genetic programming, on the programming of computers by means of natural selection", MIT Press, Cambridge, MA.
Kraslawski, A., Pedrycz, W. and Nyström, L. (1999), "Fuzzy neural network as instance generator for case-based reasoning system: an example of selection of heat exchange equipment in mixing", *Neural Comput. Appl.,* **8**(2), 106-113.
Kuo, Y.L., Jaksa, M.B., Lyamin, A.V. and Kaggwa, W.S. (2009), "ANN-based model for predicting the bearing capacity of strip footing on multi-layered cohesive soil", *Comput. Geotech.,* **36**, 503-516.
Kwak, Y.K., Eberhard, M.O., Kim, W.S. and Kim, J. (2002), "Shear strength of steel fiber-reinforced concrete beams without stirrups", *ACI Struct. J.,* **99**(4), 530-538.
Li, V.C., Ward, R. and Hamza, A.M. (1992), "Steel and synthetic fibers as shear reinforcement", *ACI Mater. J.,* **89**(5), 499-508.
Lim, T.Y., Paramasivam, P. and Lee, S.L. (1987), "Shear and moment capacity of reinforced steel-fiber-concrete beams", *Mag. Concrete Res.,* **39**(140), 148-160.
Mansur, M.A., Ong, K.C.G. and Paramsivam, P. (1986), "Shear strength of fibrous concrete beams without stirrups", *J. Struct. Eng.-ASCE*, **112**(9), 2066-2079.
Mousavi, A.H., Gandomi, A.H., Alavi, A.H. and Vesalimahmood, M. (2010), "Modeling of compressive strength of HPC mixes using a combined algorithm of genetic programming and orthogonal least squares", *Struct. Eng. Mech.,* **36**(2).
Murty, D.S.R. and Venkatacharyulu, T. (1987), "Fiber reinforced concrete beams subjected to shear force", *Proceedings of the International Symposium on Fiber Reinforced Concrete*, Madras, India.
Naaman, A.E. and Reinhardt, H.W. (2003), "High performance fiber reinforced cement composites: *HPFRCC* 4", *RILEM Proceedings Pro 30*, RILEM Publications S.A.R.L., 95-113.
Narayanan, R. and Darwish, I.Y.S. (1987), "Use of steel fibers as shear reinforcement", *ACI Struct. J.,* **84**(3), 216-227.
Narayanan, R. and Darwish, I.Y.S. (1988), "Fiber concrete deep beams in shear", *ACI Struct. J.,* **85**(2), 141-149.
Noghabai, K. (2000), "Beams of fibrous concrete in shear and bending: experiment and model", *J. Struct. Eng-ASCE,* **126**(2), 243-251.
Oh, B.H, Lim, D.H., Yoo, S.W. and Kim, E.S. (1998), "Shear behavior and shear analysis of reinforced concrete beams containing steel fibers", *Mag. Concrete Res.,* **50**(4), 283-291.
Oltean, M. and Grosan, C. (2003), "A comparison of several linear genetic programming techniques", *Complex Syst.,* **14**(4), 1-29.
Oltean, M. and Dumitrescu, D. (2002), "Multi expression programming", Technical Report, UBB-01-2002, Babes-Bolyai University, Cluj-Napoca, Romania.
Pan, Y., Jiang, J., Wang, R., Cao, H. and Cui, Y. (2009), "A novel QSPR model for prediction of lower flammability limits of organic compounds based on support vector machine", *J. Haz. Mater.,* **168**(2-3), 962-969.
Poli, R., Langdon, W.B., McPhee, N.F. and Koza, J.R. (2007), "Genetic programming: an introductory tutorial and a survey of techniques and applications", Technical Report CES-475, ISSN:1744-8050.
Regmi, S., Deschaine, L.M. and Regmi, S.R. (2004), "High fidelity approximation of slow simulators using machine learning for real-time simulation/optimization", 2004 Business and Industry Symposium, Washington, DC, USA.
RILEM TC-162-TDF (2000), "Test and design methods for steel fiber reinforced concrete", *Mater. Struct.,* **33**(226), 75-81.
Roberts, T.M. and Ho, N.L. (1982), "Shear failure of deep fiber reinforced concrete beams", *Int. J. Cement Compo. Lightweight Concrete,* **4**(3), 145-152.

Roy, P.P. and Roy, K. (2008), “On some aspects of variable selection for partial least squares regression models”, *QSAR Comb. Sci.,* **27**, 302-313.

Schwefel, H.P., Wegener, I. and Weinert, K. (2002), *Advances in Computational Intelligence -Theory and Practic*, Springer-Verlag, Berlin.

Sharma, A.K. (1986), “Shear strength of steel fiber reinforced concrete beams”, *ACI J.,* **83**(4), 624-628.

Shin, S.W., Oh, J. and Ghosh, S.K. (1994), “Shear behavior of laboratory-sized high-strength concrete beams reinforced with bars and steel fibers”, Fiber Reinforced Concrete Developments and Innovations, SP-142, American Concrete Institute, Farmington Hills, 181-200.

Smith, G.N. (1986), *Probability and Statistics in Civil Engineering*, Collins, London.

Swamy, R.N. and Bahia, H.M. (1985), “The effectiveness of steel fibers as shear reinforcement”, *Concrete Int.,* **7**(3), 35-40.

Swamy, R.N., Jones, R. and Chiam, A.T.P. (1993), “Influence of steel fibers on the shear resistance of lightweight concrete I-beams”, *ACI Struct. J.,* **90**(1), 103-114.

Swamy, R.N., Mangat, P.S. and Rao, C.V.S.K. (1974), “The mechanics of fiber reinforcement of cement matrices”, *Fiber Reinforced Concrete*, SP-44, American Concrete Institute, Farmington Hills, Mich., 1-28.

Tan, K.H., Murugappan, K. and Paramasivam, P. (1992), “Shear behavior of steel fiber reinforced concrete beams”, *ACI Struct. J.,* **89**(6), 3-11.

Uomoto, T., Weerarathe, R.K., Furukoshi, H. and Fujino, H. (1986), “Shear strength of reinforced concrete beams with fibre reinforcement”, *Proceedings, Third International RILEM Symposium on Developments in Fibre Reinforced Cement and Concrete*, Sheffield, 1986. RILEM Technical Committee 49-TFR, Sheffield University Press Unit, Sheffield.

Wang, C. (2006), “Experimental investigation on behavior of steel fiber-reinforced concrete”, M.Sc. Dissertation, University of Canterbury, New Zealand.

Zsutty, T. (1971), “Shear strength prediction for separate categories of simple beam tests”, *ACI J.,* **68**(2), 138-143.

Appendix A.

The following optimum LGP programme can be run in the Discipulus interactive evaluator mode or can be compiled in C++ environment. (Note: v[0], v[1], v[2], and v[3] respectively represent ρ, f_c', a/d, and v_b.)

```
float Discipulus C Function(float v[0], v[1], v[2],
 v[3])
{
double f[8];
doubletmp = 0;
f[1]=f[2]=f[3]=f[4]=f[5]=f[6]=f[7]=0;
f[0]=v[0];
l0: f[0]-=f[0];
l1: f[0]-=-3;
l2: f[0]-=v[1];
l3: f[1]-=f[0];
l4: f[1]-=f[0];
f[1]/=f[0];
f[0]/=f[1];
l16: f[0]/=f[1];
f[0]*=f[0];
l17: f[0]+=v[1];
l18: f[0]*=v[0];
l19: f[0]+=v[3];
l5: f[0]+=v[1];
l6: f[0]+=f[0];
l7: f[0]*=f[0];
l8: f[0]*=v[0];
l9: f[0]+=f[1];
l10: f[0]+=f[0];
f[0]+=f[0];
l11: l12: f[0]-=-3;
l13: f[0]*=f[1];
f[1]-=f[0];
l14: f[1]-=f[0];
f[0]/=f[0];
l15: f[0]+=f[0];
l20: f[0]+=v[2];
l21: f[0]+=f[0];
l22: f[0]/=v[2];
l23:
l24:
return f[0];}
```