Genetic programming to improvement FIB model: bond and anchorage of reinforcing steel in structural concrete

Juan Luis Pérez[1], Ismael Vieito[2], Juan Rabuñal[3], and Fernando Martínez-Abella[2]

[1] *School of Building Engineering and Technical Architecture, University of A Coruña, Spain*
[2] *Department of Construction Technology, University of A Coruña, Spain*
[3] *Department of Information and Communication Technologies, University of A Coruña, Spain*
{jlperez,camivr01,juanra,fmartinez}@udc.es

Abstract
Starting from the FIB database, this work is aimed to analyze the current equations which predict the main datum that can be provided by bond tests: the ultimate bar stress when the failure is reached. Furthermore, Genetic Programming (GP) techniques are also applied in order to enhance the expression of the FIB, which achieves the best adjustment so far, giving rise to the new Model Code 2010. The final result shown is a highly predictive equation. The results are compared with those included in the Model Code and it is showed the influence of the main variables on the phenomenon (concrete strength, yield strength of steel, concrete cover, transverse reinforcement and diameter of the bar).

1 Introduction

Since the dawn of the 20th century, when Abrams's tests were performed, the bond between concrete and steel led to numerous scientific papers, complex laboratory tests and many approaches of structural codes. There are very few expressions that are so different when comparing the various rules, such as those aimed at predicting the anchorage length of reinforcing bars in structural concrete. Two main lines were created starting from the works carried out by Orangun, Jirsa & Breen [1], precursors of the ACI code equation, and the studies performed by Tepfers [2], which inspired the guidelines of the Model Code, leading to the Eurocode. The tests carried out in Spain also were of great importance, as they gave rise to a specific formulation, extremely conservative for large-diameter reinforcement.

Despite the varied approaches, the three lines have a common nexus: the proposals are developed from the experimental evidence. From a basic expression of bond stress, dependent on a main variable are incorporated as a multiplicative factor the effect of other variables When a pull-out test is performed, a state of radial tension is generated around the bar that can cause damage to the surrounding concrete. The damage can be mitigated by the placement of transverse reinforcement, and having adequate cover bar. Figure 1 shows graphically the phenomenon of a bar anchorage [3].

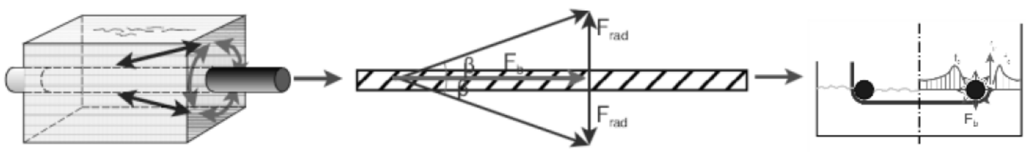

Fig. 1. Phenomenon of a bar anchorage

Various factors affect the bond capacity. They are generally associated with different origins: the materials used, the geometric conditions and, finally, the applied loads.

The design equations for the anchorage length determination are based in the basic straight anchorage length necessary to reach the break of the bar. On this equation, of experimental base, are added the effects of other variables, also obtained experimentally [1],[4].

2 FIB Model

The workgroup TG4.5 of the *Fédération Internationale du Beton* (FIB) [5] has been working for a long time in the analysis of the anchorage and the bond characteristics of reinforcement steel bars. Based in the works of Canbay and Frosch [6], the group has proposed two equations: the version 2006 (1) and the current version employed in the development of the Model Code 2010 (2), that provides the work stress that can be reached on an anchored bar. This equation depends on the parameters seen until now and is protected by a contrast with a strong experimental database.

The variables described in both formulas correspond to the bond stresses (σ_{su}), compressive strength in cylindrical specimen (f_c), diameter of the bar (d_b), length of anchored bar (l_s), maximum and minimum coatings of the bar (c_{min}, c_{max}) and the contribution of the transverse reinforcement (K_{tr}).

$$\sigma_{su} = 25 * f_c^{0.25} * \left(\frac{20}{d_b}\right)^{0.2} * \left(\frac{l_s}{d_b}\right)^{0.55} * \left(\frac{c_{min}}{d_b}\right)^{0.33} * \left(\frac{c_{max}}{c_{min}}\right)^{0.1} * (1 + 10 * K_{tr}) \quad (1)$$

$$\sigma_{su} = 54 * \left(\frac{f_c}{20}\right)^{0.25} * \left(\frac{20}{d_b}\right)^{0.2} * \left(\frac{l_s}{d_b}\right)^{0.55} * \left[\left(\frac{c_{min}}{d_b}\right)^{0.33} * \left(\frac{c_{max}}{c_{min}}\right)^{0.1} + 8K_{tr}\right] \quad (2)$$

2.1 Dataset

The database currently contains data (variables and results measured or calculated) corresponding to 813 trials. As will be applied GP techniques for analysis, so that the database range is consistent and frequencies of each of the data must be analyzed. Considering the frequency histogram data, several filters are applied and are accepted those recommended by the FIB [5]. One of the most important is related to the output data σ_{su}, the stress reached by the bar during the test. Since any bar limits its maximum stress f_y, the maximum value that can take the relationship σ_{su}/f_y is 1.05. In other variables their relative values are also limited, for example those related to the concrete cover. Thus, filters over c_{min}/c_{max} and c_{min}/d_b are applied. Table 1 shows the parameters used for filtering and the filter finally applied.

Table 1. Filters applied to the dataset

Variables	Filter
f_c	>15 and <115 MPa
d_b	< 37 mm
l_s	< 2100 mm
C_b	<136 mm
c_{min}/c_{max}	≥ 1.0 and ≤ 5.0
c_{min}/d_b	≥ 0.5 and ≤ 3.0
σ_{su}/f_y	≤ 1.05

After applying the filter, the BD is reduced to a total of 628 trials, of which 77.5% (487 trials), by random selection, are used for training, dedicating the remaining 22.5% (141 trials) to check. Table 2 shows the final range of the data in each of the subsets set (training and verification).

On the BD filtered FIB expressions produce results whose accuracy is presented in the following sections.

Table 2. Distribution of data in subsets defined over the BD filtered

	Training (#487)		Test (#141)	
	Min	Max	Min	Max
l_s (mm)	50	2095	120	2032
d_b (mm)	8	35.81	10	35.81
c_{min} (mm)	8	76	8	76
c_{max} (mm)	20	140	20	140
f_c (MPa)	15	114	20	110
K_{tr}	0	0.114	0	0.106
σ_{su}	126	788	182	814

3 Method

The method followed is oriented to improving the FIB equations developed for predicting the stress of bar anchored. The method used follows the same guidelines referred to in the paper developed by Pérez et al [7]. In summary, the method is based on GP techniques, imposing some restrictions based on knowledge of the problem provided by an expert. Symbolic regression data is one of the capabilities provided by the GP.

Having a data set (input-output), the GP is able to relate these data algebraically by an equation. Its complexity may vary, and dimensional integrity is not guaranteed. This technique, applied in many cases in civil engineering, is one of those followed by Ashour al [8], for example, to predict shear strength in concrete beams. Naturally, the form of the equations obtained is very different from the ones in the common codes. The presented method improves the terms accepted by the scientific community, getting a better fit when the results are applied to a database.

It starts from the expression FIB-2006, because it shows better results over the database. The search expression will determine the bar stress predicted (spred) to be compared with the real stress test (σ_{test}). Firstly, it is necessary to define how individuals will be evaluated in the fitness function (equation 4). In this equation, σ_{test} is the bar stress at failure, α is the parsimony coefficient, s_i is the number of nodes in the expression and n is the number of cases of the database. It should set the parameters p_i and l_{bias} defined in equation 3.

After performing several tests, was adopted l_{bias} = 1.0, and equation 4 shows the value of pi (DP). This equation is based on the use of the technique of "demerit points".

$$fitness(i) = \frac{\sum * \left| l_{bias} - \frac{\sigma_{test}}{\sigma_{pred}} \right|}{n} + \alpha * s_i \qquad (3)$$

$$DP = \sum_{i=1}^{n} p_i,\ p_i = \begin{cases} 16, & \frac{\sigma_{test}}{\sigma_{pred}} < 0.5 \\ 8, & 0.5 \le \frac{\sigma_{test}}{\sigma_{pred}} < 0.67 \\ 4, & 0.67 \le \frac{\sigma_{test}}{\sigma_{pred}} < 0.85 \\ 2, & 0.85 \le \frac{\sigma_{test}}{\sigma_{pred}} < 1.0 \\ 0, & 1.0 \le \frac{\sigma_{test}}{\sigma_{pred}} < 1.3 \\ 3, & 1.3 \le \frac{\sigma_{test}}{\sigma_{pred}} < 2 \\ 6, & \frac{\sigma_{test}}{\sigma_{pred}} \ge 2 \end{cases} \qquad (4)$$

The technique was adapted for “oriented” searches were possible, with different purposes.

The orientation was introduced through impositions or restrictions, which include:

- restriction on the type of functions that link the variables
- preferred selection of individuals with the highest ratios $\sigma_{test}/\sigma_{pred}$. From the structural point of view, is much more appropriate this option for safety reasons

The method used starts with the establishment of a “framework” from wich genetic programming will make the evolutive process, taking into account the restrictions and impositions.

The framework is defined directly from the equation FIB-2006, which is divided into subexpressions. Also, each subexpression is written indicating which factor (branch) may change in the search process. The working lines can find:

- The optimization of the numerical coefficients of the equation. The branches will be Real values
- The introduction of a new subexpression. This can be a Real number or a function (new branch) linked to a variable

As mentioned, in this type of model is very important that the predicted stress is equal to or greater than the actual value.

In general, if an individual differs from the real value is penalized during training. From a mathematical point of view, *S* values equal to 0.5 or 1.5 should be penalized equally. To take into account the structural safety, the individual 0.5 should be penalized more than the individual 1.5, as it causes structural insecurity (collapse).

This is achieved through the technique of demerit points, whereby the error of the expression is weighted according to the ranges defined by Pérez [9]. The fitness function (3) shows how the p_i factor weights the prediction error, according to the intervals and values of the equation (4).

The method used starts with the establishment of a “framework” about whom genetic programming will make the evolutive process, taking into account the restrictions and impositions. Such “framework” is based on the 2006 FIB formulation, about which it will be introduced new variables or its coefficients will be modified.

In the searching process it has been proposed three basic equations (5)(6)(7). Each *branch* is designated as Bi. Table 3 shows the default settings implemented, based on the initial tests. The input data have not been standardized, so expressions can be used directly.

Table 3. Parameters used

Parameter	Default value	Other values
Population size	1000	
Crossover rate	80%	
No-terminal selection rate:	90%	
Mutation rate:	20%	
Algorithms:	Selection: Tournament Initialization: Ramped Half & Half Mutation& Crossover: Subtree	
Elitist strategy	Yes	
Parsimony	0	0.0001, $1*10^{-6}$ ó $1*10^{-9}$
Initial tree depth	4	5
Maximum tree depth	6	7
Maximum mutation depth	4	5

$$\sigma_{su} = B_1^{B_2} * fc^{B_3} * \left(\frac{l_s}{d_b}\right)^{B_4} * \left(\frac{c_{min}}{d_b}\right)^{B_5} * \left(\frac{c_{max}}{c_{min}}\right)^{B_6} * (1 + B_7 * K_{tr})^{B_9} \quad (5)$$

$$\sigma_{su} = ls^{B_1} * B_2 * (1 + B_3 * K_{tr})^{B_4} \quad (6)$$

$$\sigma_{su} = B_1 * fc^{B_2} * \left(\frac{B_3}{d_b}\right)^{B_4} * \left(\frac{l_s}{d_b}\right)^{B_5} * \left(\frac{c_{min}}{d_b}\right)^{B_6} * \left(\frac{c_{max}}{c_{min}}\right)^{B_7} * (1 + B_8 * K_{tr})^{B_9} \quad (7)$$

By default, addition, subtraction, multiplication and protected division were chosen as operators or non-terminal nodes. Variables from the data set (l_s, d_b, c_{min}, c_{max} and f_c), and integers in the range [-10, 10] were adopted as terminal nodes.

Constraints over the equations are showed in table 4. Equation (5) have three types of constraints ("A", "B" and "C"), the constraint "D" is imposed to equation (6) and finally the constraint "E" is imposed to equation (7).

Table 4. Constraints

Eq.	Const.	B_1	B_2	B_3	B_4	B_5	B_6	B_7	B_8	B_9
(5)	A	l_s d_b c_{min} c_{max} f_c	Const. 2 dec.	Const. 2 dec.	Const. 2 dec.	Const. 2 dec.	Const. 2 dec.	Const. 2 dec.	Const. 2 dec.	-
(5)	B	d_b c_{min} c_{max} f_c	Const. 2 dec.	Const. 2 dec.	Const. 2 dec.	Const. 2 dec.	Const. 2 dec.	Const. 2 dec.	Const. 2 dec.	-
(5)	C	d_b c_{min} c_{max} f_c	Const. 2 dec.	Const. 2 dec	0.5	Const. 2 dec	Const. Ent.	Const. 2 dec	Const. 2 dec	-
(6)	D	Const. 2 dec.	d_b c_{min} c_{max} f_c	Const. 2 dec	Const. 2 dec	-	-	-	-	-
(7)	E	Const. 2 dec.	Const. 2 dec.	Const. 2 dec.	Const. 2 dec.	Const. 2 dec.	Const. 2 dec.	Const. 2 dec.	Const. 2 dec.	Const. 2 dec.

4 Results

In total, more than 4,500 executions were carried out. The results are analyzed essentially through the following indicators: COV (variation coefficient), $\sigma_{test}/\sigma_{pred}$, R^2 (square root of Pearson product-moment correlation coefficient), MSE (mean square root error), ME (mean error), and finally demerit points calculated according to equation (4).

According to the best results, a select group of equations was chosen. If the denominator could be negative, expressions containing function "protected division" were rejected. Also too complex equations were also discarded. PG_9RSC4 (8), PG_8v2R5 (9), PG_7v3F2 (10), PGcc6 (11) were more accurate equations. Since not provide substantial improvements, these equations do not contain the derivatives of the classic GP. The results are shown in Table 5. The significant improvement achieved is evident by comparing the results of the equations FIB.

Table 5. Results

	FIB (2006)	FIB CM2010	PG_7v3F2	PG_8v2R5	PG_9RSC4	PGcc6
COV	15.683	16.010	14,404	14,900	15,239	15,442
$\sigma_{test}/\sigma_{pred}$	0.9712	0.9748	1,0254	1,0082	0,9994	1,0079
Max ($\sigma_{test}/\sigma_{pred}$)	1.5091	1.5367	1,5228	1,5119	1,4815	1,5684
Min ($\sigma_{test}/\sigma_{pred}$)	0.4990	0.4885	0,5702	0,5137	0,5016	0,4873
R^2	0.7095	0.6938	0,7545	0,7400	0,7271	0,7193
MSE	4215	4343	3551	3608	3740	3847
ME	51.26	52.19	46.08	47.24	48.10	48.40
DP	2642	2646	2398	2492	2508	2486

Some of the expressions stand out by different appearances. PG_9RSC4 is a simple improvement of the FIB equation, achieved with better adjusts of the exponents and constants. To clear the value of the length, it is necessary to impose conditions to the search, proposing a first free function, not dependent on the length, and a adjust coefficient for the rest of parameters: the equation PG_8v2R5 arises this way. The marked tendency that exhibits the exponent (l_s/d_b) to the value 0.5, induces a new group of executions in which this constant is fixed. With this procedure, the PG_7v3F2 equation is obtained, achieving a very noticeable distribution.

$$PG9RSC4\text{:}\ \sigma_{su} = 27 * fc^{0{,}27} * \left(\frac{29}{d_b}\right)^{0.23} * \left(\frac{l_s}{d_b}\right)^{0.47} * \left(\frac{c_{min}}{d_b}\right)^{0.16} * \left(\frac{c_{max}}{c_{min}}\right)^{0.13} * (1 + 286 * Ktr)^{0.11} \tag{8}$$

$$PG8v2R5\text{:}\ \sigma_{su} = \left(\frac{c_{max}}{36} - \frac{f_c}{36} + \frac{150}{c_{min}} + 19.294\right) * fc^{0.29} * \left(\frac{ls}{d_b}\right)^{0.49} * \left(\frac{c_{min}}{d_b}\right)^{0.29} * \left(\frac{c_{max}}{c_{min}}\right)^{0.04} * (1 + 239.29 * Ktr)^{0.12} \tag{9}$$

$$PG7v3F2\text{:}\ \sigma_{su} = \left(\frac{48}{c_{min}} - \frac{fc - 44}{\frac{c_{min}}{48} + 20} + 9\right) * fc^{0.5} * \left(\frac{ls}{d_b}\right)^{0.5} * \left(\frac{c_{min}}{d_b}\right)^{0.34} * \left(\frac{c_{max}}{c_{min}}\right)^{0.01} * (1 + 173 * Ktr)^{0.14} \tag{10}$$

$$PGcc6\text{:}\ \sigma_{su} = ls^{0.5} * \left(\left(\frac{75}{d_b} + \frac{19}{7}\right) * \left(\frac{fc}{5 * d_b} + \frac{c_{max} * fc}{360 * d_b} + 2\right)\right) * (1 + 165 * Ktr)^{0.15} \tag{11}$$

In the last remarkable groups, it is allowed the apparition of a free function (without l_s) that multiplies l_s with constant exponent and the classical term of transversal reinforcement contribution, improved with constants. This is the PGcc6 expression, which exhibit a strong concentration around the unit.

Next, the stresses that can be developed for some specific variables are compared in two of the equations found against the FIB deduced expressions. It can be observed the similarity of the approach, even for equations that are not born from the structure of the FIB.

As a result of the previously exposed, it can be recommended to adopt the expression PG_7v3F2 as a good equation to get the bond behavior of the passive reinforcement in a concrete element.

5 Conclusions

FIB equation to determine rebar tension stress was improved with the application of heuristic techniques.

In the applied method, structural safety was taken into account, through the weighting provided by demerit points

As a final conclusion and summary it should be noted that it has managed to implement a novel method based on genetic programming to extract knowledge from experimental data based on the experience. This experience is implemented through constraints that are induced in the algorithm.

Acknowledgements. This work was partially supported by the Spanish Ministry of Science and Innovation (Ref. BIA2010-21551) and grants from the Ministry of Economy and Industry (Consellería de Economía e Industria) of the Xunta de Galicia (Ref. 10MDS014CT, Ref. 10TMT042E, Ref. 10TMT118004PR and Ref. 10TMT034E).

References

1. Orangun, C.O., Jirsa, J.O., Breen, J.E.: A Reevaluation of Test Data on Development Length and Splices. ACI Journal 74(11), 114–122 (1977)
2. Tepfers, R.A.: Theory of bond applied to overlapped tensile reinforcement splices for deformed bars Ph.D. Thesis, Division of concrete structures, Chalmers University of Technology. Gothenburg, Sweden (1973)
3. Cairns, J.: Model for strength of lapped joints and anchorages. Paper presented at the meeting of the TG group 4.5 of the International Federation for Structural Concrete, Stuttgart, Germany (2006)
4. Abrams, D.A.: Tests of Bond Between Concrete and Steel. University of Illinois Bulletin Nº 71. University of Illinois, Urbana (1913)
5. FIB Task Group 4.5 “Bond models” (s.f) (2007), http://fibtg45.dii.unile.it/ (January 25, 2013)
6. Canbay, E., Frosch, R.: Bond strength of lap-spliced bars. ACI Structural Journal 102(4), 605–614 (2005)
7. Pérez, J.L., Cladera, A., Rabuñal, J.R., Martínez-Abella, F.: Optimization of existing equations using a new genetic programming algorithm: Application to the shear strength of reinforced concrete beams. Advances in Engineering Software 50(1), 82–96 (2012)
8. Ashour, A.F., Alvarez, L.F.: Toropov VV.: Empirical modeling of shear strength of RC deep beams by genetic programming. Computers & Structures 81(5), 331–338 (2003)
9. Pérez, J.L.: Metodología para orientar procesos de extracción de conocimiento basados en Computación Evolutiva. Aplicación al desarrollo de modelos y formulaciones en el ámbito del hormigón estructural, Ph.D. Thesis, Department of Information and Communication Technologies, University of A Coruña (2010)