ARTICLE IN PRESS
Journal of Constructional Steel Research ▮ (▮▮▮▮) ▮▮▮–▮▮▮
Contents lists available at ScienceDirect
Journal of Constructional Steel Research
journal homepage: www.elsevier.com/locate/jcsr

# A new prediction model for the load capacity of castellated steel beams

Amir Hossein Gandomi [a,*], Seyed Morteza Tabatabaei [b], Mohammad Hossein Moradian [c], Ata Radfar [d], Amir Hossein Alavi [e]

[a] *College of Civil Engineering, Tafresh University, Tafresh, Iran*
[b] *Department of Civil Engineering, The Institute of Higher Education of Eqbal Lahoori, Mashhad, Iran*
[c] *Department of Civil Engineering, Islamic Azad University, Science and Research Branch, Tehran, Iran*
[d] *Department of Civil Engineering, Ferdowsi University of Mashhad, Mashhad, Iran*
[e] *School of Civil Engineering, Iran University of Science and Technology, Tehran, Iran*

ARTICLE INFO

*Article history:*
Received 29 July 2010
Accepted 21 January 2011

*Keywords:*
Castellated beam
Failure load
Gene expression programming

ABSTRACT

In this study, a robust variant of genetic programming, namely gene expression programming (GEP), is utilized to build a prediction model for the load capacity of castellated steel beams. The proposed model relates the load capacity to the geometrical and mechanical properties of the castellated beams. The model is developed based on a reliable database obtained from the literature. To verify the applicability of the derived model, it is employed to estimate the load capacity of parts of the 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. A multiple least squares regression analysis is performed to benchmark the GEP-based model. A sensitivity analysis is further carried out to determine the contributions of the parameters affecting the load capacity. The results indicate that the proposed model is effectively capable of evaluating the failure load of the castellated beams. The GEP-based design equation is remarkably straightforward and useful for pre-design applications.
© 2011 Elsevier Ltd. All rights reserved.

## 1. Introduction

With the development of electrical welding technology in steel construction, the castellated steel beam (CSB) became available to structural engineers. The CSB, initially known as the "Boyd beam", was first used in 1910 [1], and then designed and manufactured in the early 1930s as roof beams in Czech Republic. The initial inspiration for the elastic and plastic calculation methods were respectively introduced in 1942 and the early 1970s. Castellation is the procedure of cutting the web of a rolled section in a zigzag pattern [2]. One of the halves is turned round and welded to the other half. This procedure increases the depth of the original beam ($h$) by the depth of the cut ($d$). As shown in Fig. 1, CSBs are often made from I sections by the castellation process. This shape fits the dictionary definition of castellated as "castle-like". According to Zirakian and Showkati [3], the basic reasons for the fabrication of the castellated beams are as follows.

I. Augmentation of section height, resulting in the enhancement of moment of inertia, section modulus, stiffness, and flexural resistance of sections;
II. Decreasing the weight of structures;
III. Optimum use of existing profiles;
IV. By-passing the used plate girders; and
V. The passage of services through the web openings.

The seven potential failure modes generally associated with the castellated beams are as follows [4].

i. Flexure mechanism formation;
ii. Overall beam lateral–torsional buckling;
iii. Vierendeel mechanism formation;
iv. Welded joint rupture in the web;
v. Web post-shear bucking;
vi. Web post-compression buckling; and
vii. Tee compression buckling.

These modes are of two different categories. The first two modes are similar to the corresponding modes for solid-web beams and may be analyzed in almost identical fashions. Modes iii to vii are specific to castellated beams, since they are associated with the tees and web posts that bound openings. There is a clear correlation between mode iv and shear failure, and between mode vi and the buckling solid web. However, it is necessary to develop new analytical models for modes iii to vii. Detailed explanations for these failure modes can be found in [5].

Knowles [6] showed that the failure load predicted using a column in compression approach adopting Blogett's force

* Corresponding author.
*E-mail addresses:* a.h.gandomi@gmail.com (A.H. Gandomi), ah_alavi@hotmail.com (A.H. Alavi).

0143-974X/$ – see front matter © 2011 Elsevier Ltd. All rights reserved.
doi:10.1016/j.jcsr.2011.01.014

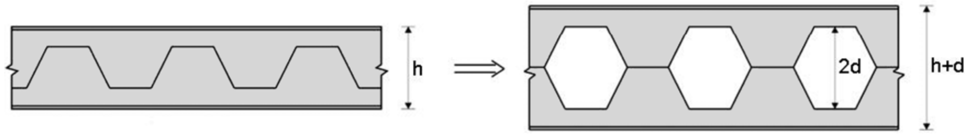


**Fig. 1.** Castellation process in a CSB.

distribution model and an effective length factor would provide a good agreement with the experimental results. In this approach, the web post is treated as a column having a width equal to the narrowest width of the web, a length equal to the clear height of the castellation, and a thickness equal to the web thickness. The model recommended in BS 449, Clause 28a, is for 45° dispersion angle. Applying BS 449 to CSBs requires the consideration of the width to thickness ratio in order that the section can be correctly classified for local buckling. The rules apply only to sections castellated to the profile adopted in the UK. Although many of the rules may well apply to other types of section, they have not been fully approved. This is particularly true with unrestrained beams [7].

Developments in computational software and hardware have enabled several alternative computer-aided data mining approaches to emerge. As an example, pattern recognition systems learn adaptively from experience and extract various discriminators. Artificial neural networks (ANNs) are the most widely used pattern recognition procedures. ANNs have been applied to assess different characteristics of steel beams [8–10]. Despite the acceptable performance of ANNs, they do not commonly give a particular function to estimate the outcome using the input values.

Genetic programming (GP) [11] is an alternative approach for behavior modeling of structural engineering problems. GP may generally be defined as a supervised machine learning technique that searches a program space instead of a data space [12,13]. Many researchers have employed GP and its variants to discover complex relationships among experimental data [13–16]. Gene expression programming (GEP) [17] is a recent extension of GP. GEP evolves computer programs of different sizes and shapes encoded in linear chromosomes of fixed length. The GEP chromosomes are composed of multiple genes, each gene encoding a smaller subprogram. The GEP approach has been shown to be an efficient alternative to traditional GP [17,18]. There have been some scientific efforts directed at applying GEP to structural engineering tasks [19–21].

The main purpose of this paper is to utilize the GEP technique to build a predictive model for the failure load of CSBs. A reliable database of previously published CSB test results is utilized to develop the model. The performance of the derived model is subsequently compared with that of a regression-based model.

## 2. Genetic programming

GP is a subset of genetic algorithms (GAs) with a great ability to automatically evolve computer programs. The evolutionary process followed by GP is inspired by the principle of Darwinian natural selection. GP was introduced by Koza [11] after experiments on symbolic regression. The main difference between GP and a GA is the representation of the solution. The GP solutions are computer programs that are represented as tree structures and expressed in a functional programming language (like LISP) [11]. A GA creates a string of numbers that represent the solution. In other words, in GP, the evolving programs (individuals) are parse trees than can vary in length throughout the run rather than fixed-length binary strings. Traditional optimization techniques, such as the GA, are generally used in parameter optimization to evolve the best values for a given set of model parameters. GP, on the other hand, gives the basic structure of the approximation model together with

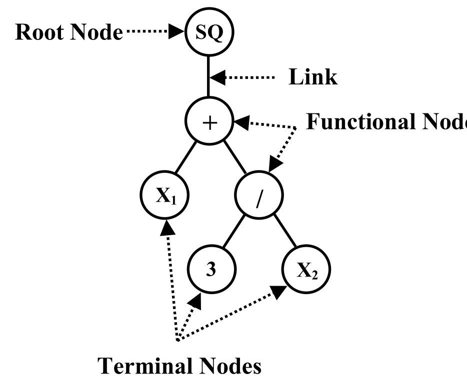


**Fig. 2.** Tree representation of a GP model $(X1 + 3/X_2)^2$.

the values of its parameters [22,23]. GP optimizes a population of computer programs according to a fitness landscape determined by a program's ability to perform a given computational task. The fitness of each program in the population is evaluated using a fitness function. Thus, the fitness function is the objective function GP aims to optimize [23,24].

This classical GP approach is referred to as tree-based GP. A population member in tree-based GP is a hierarchically structured tree comprising functions and terminals. The functions and terminals are selected from a set of functions and a set of terminals. For example, the function set $F$ can contain the basic arithmetic operations $(+, -, \times, /,$ etc.), Boolean logic functions (AND, OR, NOT, etc.), or any other mathematical functions. The terminal set $T$ contains the arguments for the functions and can consist of numerical constants, logical constants, variables, etc. The functions and terminals are chosen at random and constructed together to form a computer model in a tree-like structure with a root point with branches extending from each function and ending in a terminal. An example of a simple tree representation of a GP model is illustrated in Fig. 2 [23].

GEP is a linear variant of GP. The linear variants of GP make a clear distinction between the genotype and the phenotype of an individual. Thus, the individuals are represented as linear strings [13,18].

### 2.1. Gene expression programming

GEP is a natural development of GP first invented by Ferreira [17]. Most of the genetic operators used in GAs can also be implemented in GEP with minor changes. GEP consists of five main components: function set, terminal set, fitness function, control parameters, and termination condition [13]. Unlike the parse-tree representation in conventional GP, GEP uses a fixed length of character strings to represent solutions to the problems, which are afterwards expressed as parse trees of different sizes and shapes. These trees are called GEP expression trees (ETs). One advantage of the GEP technique is that the creation of genetic diversity is extremely simplified, as genetic operators work at the chromosome level. Another strength of GEP is its unique multigenic nature, which allows the evolution of more complex programs composed of several subprograms [13]. Each GEP gene contains a list of symbols with a fixed length that can be any element from a function set

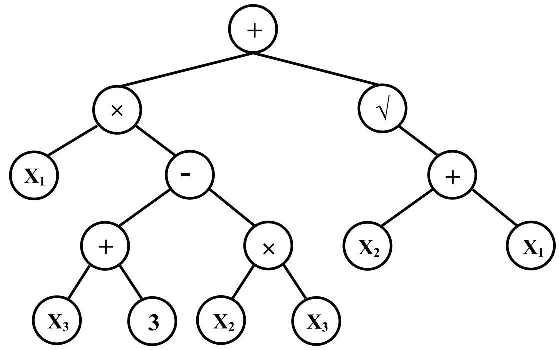


**Fig. 3.** Example of expression trees (ETs).

such as $\{+, -, \times, /, \sqrt{}\}$ and a terminal set such as $\{X_1, X_2, X_3, 3\}$. The function set and the terminal set must have the closure property: each function must be able to take any value of data type which can be returned by a function or assumed by a terminal. A typical GEP gene with the given function and terminal sets can be as follows [13]:

$$\underline{+.\times.\sqrt{}.X_1.-.}+.+.\times.X_2.X_1.X_3.3.X_2.X_3, \quad (1)$$

where $X_1, X_2$, and $X_3$ are variables and 3 is a constant; "." is element separator for easy reading. The above expression is termed Karva notation or a K-expression [17,25]. A K-expression can be represented by a diagram which is an ET. For example, the above sample gene can be expressed as in Fig. 3.

The conversion starts from the first position in the K-expression, which corresponds to the root of the ET, and reads through the string one by one [13] The above GEP gene can also be expressed in a mathematical form as

$$X_1((X_1+3)-(X_2\times X_3))+\sqrt{(X_2+X_1)}. \quad (2)$$

An ET can inversely be converted into a K-expression by recording the nodes from left to right in each layer of the ET, from the root layer down to the deepest one to form the string. As previously mentioned, GEP genes have fixed length, which is predetermined for a given problem. Thus, what varies in GEP is not the length of the genes but the size of the corresponding ETs [13]. This means that there exist a certain number of redundant elements, which are not useful for genome mapping. Hence, the valid length of a K-expression may be equal to or less than the length of the GEP gene. To guarantee the validity of a randomly selected genome, GEP employs a head–tail method. Each GEP gene is composed of a head and a tail. The head may contain both function and terminal symbols, whereas the tail may contain only terminal symbols [13,17,25].

A basic representation of the GEP algorithm is presented in Fig. 4 [26]. In GEP, the individuals are selected and copied into the next generation according to the fitness by roulette wheel sampling with elitism. This guarantees the survival and cloning of the best individual to the next generation. Variation in the population is introduced by conducting single or several genetic operators on selected chromosomes, which include crossover, mutation, and rotation. The rotation operator is used to rotate two subparts of the element sequence in a genome with respect to a randomly chosen point. It can also drastically reshape the ETs. As an example, the following gene [13]

$$+.+.\times.X_2.X_1.X_3.3.X_2.X_3.\underline{+.\times.\sqrt{}.X_1.-} \quad (3)$$

rotates the first five elements of gene (1) to the end. Only the first seven elements are used to construct the solution function $(X_2+X_1)+(X_3\times 3)$, with the corresponding expression shown in Fig. 5 [13].

## 3. Modeling of a CSB

The enhanced performance characteristics of steel beams are generally achieved by many processes such as castellation. The use

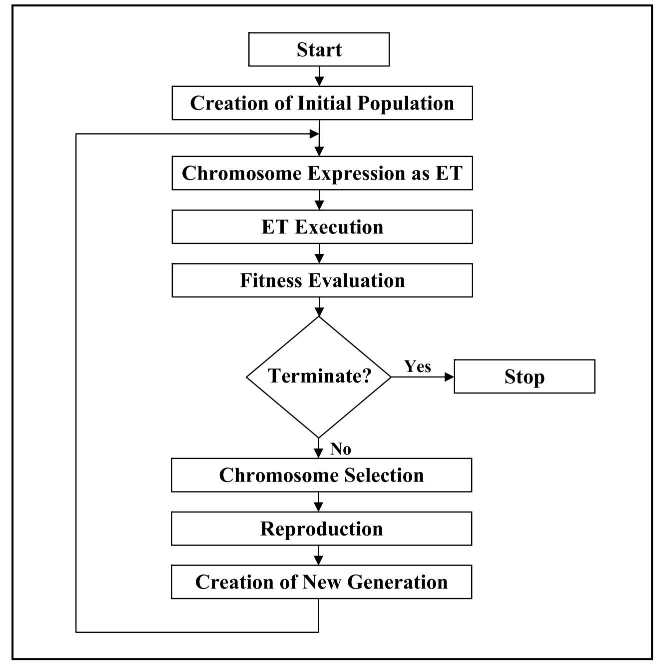


**Fig. 4.** A basic representation of the GEP algorithm.

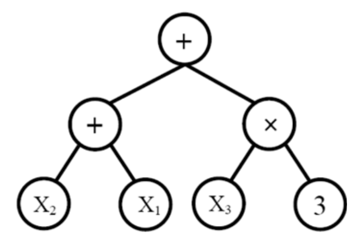


**Fig. 5.** Example of expression trees (ETs).

of a CSB plays an important role in the structural performance of steel beams [3]. In its current state, behavior modeling of a CSB is more difficult than that of normal I-section beams. In order to provide an accurate assessment of the performance characteristics of CSBs, the effects of all influencing parameters should be incorporated into the model development. In this study, the GEP approach was utilized to obtain a meaningful relationship between the load capacity of a CSB ($P$) and the influencing variables, as follows:

$$P = f(F_{yw}, h_c, B, t_w, t_f, S, L, LC), \quad (4)$$

where

$F_{yw}$ (MPa): Minimum web yield stress
$h_c$ (mm): Overall depth
$B$ (mm): Width of flange
$t_w$ (mm): Web thickness
$t_f$ (mm): Flange thickness
$S$ (mm): Minimum width of the web post
$L$ (m): Span of castellated beam
$LC$: Indicator variable representing different loading conditions
1: One-point load
2: Two-point load
3: Distributed load.

The above variables were chosen as the predictor variables on the basis of a literature review [4] and a trial study.

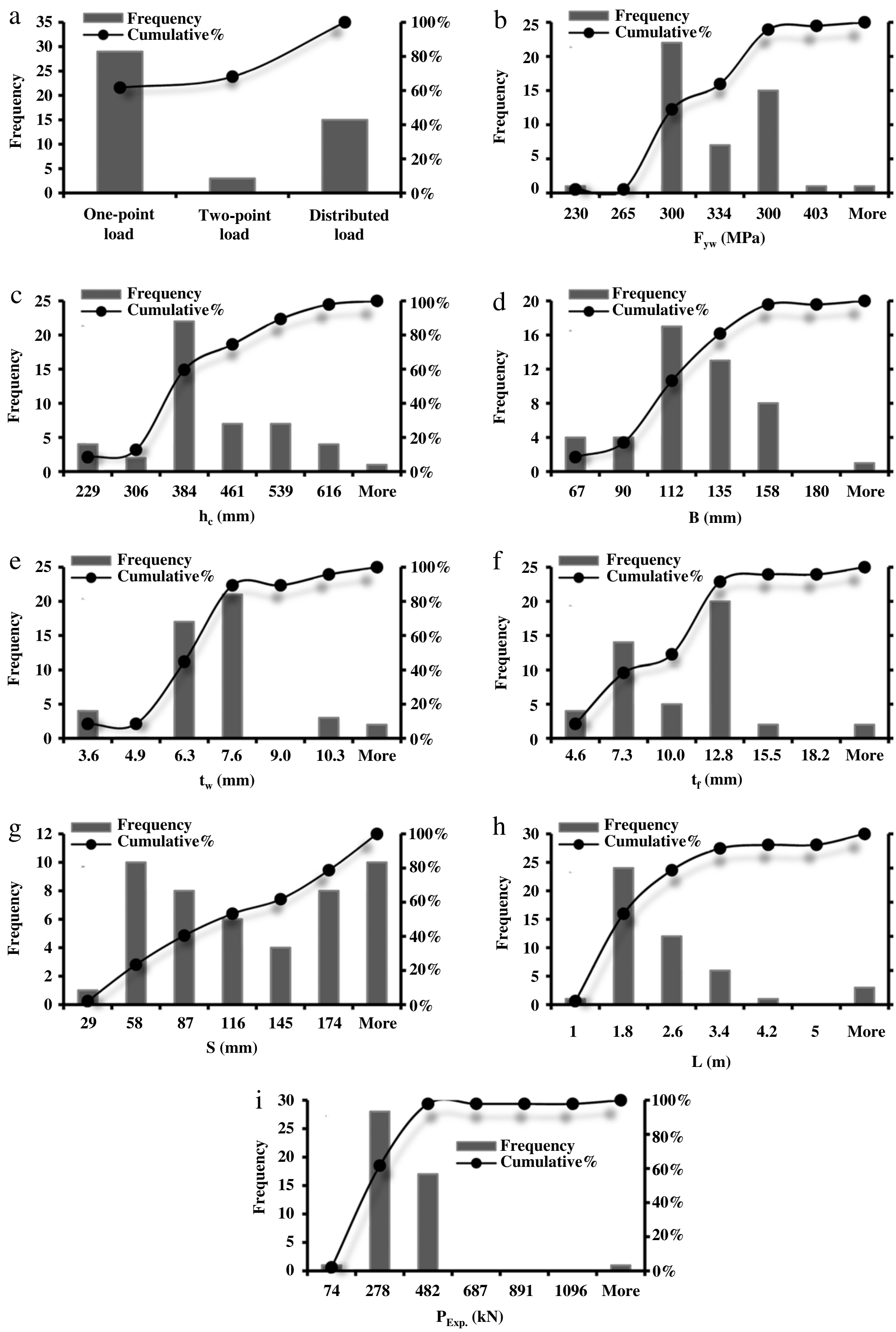


**Fig. 6.** Histograms of the variables used for model development.

*3.1. Experimental database*

A reliable database was obtained from the literature to develop the models. The database contains 47 test results for the load capacity of CSBs carried out by several researchers [1,27–32] and presented by Amayreh and Saka [4]. To visualize the distribution of the samples, the data are presented by frequency histograms (Fig. 6). The descriptive statistics of the database used in this study are given in Table 1. The complete list of the data is given in Table 2.

Overfitting is one of the major problems in machine learning generalization. An efficient approach to prevent overfitting is to test the derived models on a validation set to find a better generalization [12]. This strategy was considered in this study for improving the generalization of the models. For this aim, the

**Table 1**
Descriptive statistics of the variables used for model development.

| | *LC* | $F_{yw}$ (MPa) | $h_c$ (mm) | $B$ (mm) | $t_w$ (mm) | $t_f$ (mm) | $S$ (mm) | $L$ (m) | $P_{\mathrm{Exp}}$ (kN) |
|---|---|---|---|---|---|---|---|---|---|
| Mean | – | 314.894 | 419.862 | 112.481 | 6.741 | 9.497 | 110.604 | 2.095 | 250.417 |
| Standard error | – | 4.959 | 15.500 | 4.048 | 0.246 | 0.488 | 8.317 | 0.171 | 25.658 |
| Standard deviation | – | 34.000 | 106.260 | 27.750 | 1.685 | 3.345 | 57.017 | 1.171 | 175.905 |
| Sample variance | – | 1156.014 | 11 291.089 | 770.065 | 2.840 | 11.187 | 3250.943 | 1.370 | 30 942.431 |
| Kurtosis | – | 3.177 | 0.185 | 1.197 | 1.728 | 3.278 | −1.497 | 4.496 | 28.351 |
| Skewness | – | 1.043 | 0.339 | 0.567 | 0.690 | 1.242 | 0.095 | 2.095 | 4.675 |
| Range | 2 | 208 | 464.7 | 136.3 | 8.14 | 16.37 | 174.42 | 4.8 | 1226.5 |
| Minimum | 1 | 230 | 229 | 66.9 | 3.56 | 4.59 | 28.58 | 1 | 73.5 |
| Maximum | 3 | 438 | 693.7 | 203.2 | 11.7 | 20.96 | 203 | 5.8 | 1300 |
| Sum | – | 14800 | 19 733.5 | 5286.6 | 316.81 | 446.35 | 5198.41 | 98.46 | 11 769.6 |

**Table 2**
The experimental database used for the model construction.

| No. | *LC* | $F_{yw}$ (MPa) | $h_c$ (mm) | B (mm) | $t_w$ (mm) | $t_f$ (mm) | $S$ (mm) | $L$ (m) | $P_{\mathrm{Exp}}$ (kN) | $P_{\mathrm{GEP}}$ (kN) | $P_{\mathrm{LSR}}$ (kN) |
|---|---|---|---|---|---|---|---|---|---|---|---|
| 1 | One-point | 320 | 380 | 150 | 7.1 | 10.7 | 40 | 5.8 | 176.5 | 148.3 | 74.6 |
| 2 | One-point | 335 | 500 | 150 | 7.1 | 10.7 | 100 | 5.8 | 73.5 | 132.0 | 221.8 |
| 3 | One-point | 290.3 | 451.8 | 123.4 | 7.55 | 10.72 | 150 | 2.14 | 285 | 269.2 | 221.0 |
| 4 | One-point | 293.2 | 606.2 | 145.6 | 7.31 | 11.4 | 200 | 1.5 | 280 | 297.0 | 337.7 |
| 5 | One-point | 335 | 440 | 150 | 7.1 | 10.7 | 70 | 5.8 | 145 | 134.0 | 177.7 |
| 6 | One-point | 294.7 | 524.8 | 124.4 | 7.03 | 10.79 | 180 | 2.6 | 280 | 239.1 | 212.3 |
| 7 | Two-point | 355 | 693.7 | 153.5 | 11.7 | 11.7 | 114.3 | 2.5 | 1300 | 1226.8 | 1149.0 |
| 8 | Distributed | 335 | 381 | 101.6 | 5.84 | 6.83 | 89 | 1.3 | 295 | 226.4 | 330.8 |
| 9 | Distributed | 320 | 229 | 76.2 | 5.84 | 9.58 | 38.1 | 1.32 | 147 | 141.6 | 105.9 |
| 10 | One-point | 290.8 | 451.4 | 123.9 | 7.51 | 10.7 | 150 | 1.15 | 275 | 282.5 | 266.5 |
| 11 | One-point | 352.9 | 380.5 | 66.9 | 3.56 | 4.59 | 66.55 | 1.83 | 94.8 | 158.9 | 150.2 |
| 12 | Distributed | 292.8 | 381 | 101.6 | 5.84 | 6.83 | 127 | 1.8 | 186 | 189.7 | 125.5 |
| 13 | Distributed | 292.8 | 381 | 127 | 9.14 | 14.02 | 63.5 | 1.8 | 249.1 | 312.6 | 365.5 |
| 14 | One-point | 352.9 | 380.5 | 66.9 | 3.56 | 4.59 | 66.55 | 1.22 | 100.9 | 156.2 | 179.8 |
| 15 | Two-point | 292.8 | 381 | 114.3 | 7.62 | 12.83 | 63.5 | 1.6 | 299 | 200.0 | 215.2 |
| 16 | One-point | 295.5 | 526.7 | 124.7 | 7.08 | 10.68 | 176 | 2 | 232 | 261.9 | 255.7 |
| 17 | One-point | 277 | 605.5 | 145.4 | 7.3 | 11.34 | 203 | 1.6 | 288 | 269.0 | 267.5 |
| 18 | One-point | 277 | 605.5 | 143.3 | 7.27 | 11.25 | 199 | 1.5 | 253 | 268.8 | 273.4 |
| 19 | One-point | 293.2 | 603.6 | 143.7 | 7.28 | 11.27 | 196 | 1 | 226 | 301.5 | 361.3 |
| 20 | One-point | 320 | 381 | 101.6 | 5.84 | 6.83 | 44.45 | 3 | 265.5 | 218.6 | 168.0 |
| 21 | One-point | 290 | 524.3 | 124.6 | 7.04 | 10.7 | 179 | 2.6 | 275 | 234.3 | 196.6 |
| 22 | Distributed | 335 | 229 | 76.2 | 5.84 | 9.58 | 38.1 | 1.32 | 137.2 | 157.4 | 164.8 |
| 23 | One-point | 320 | 381 | 101.6 | 5.84 | 6.83 | 34.93 | 3 | 269 | 237.7 | 171.9 |
| 24 | One-point | 304 | 460.4 | 103.1 | 7.21 | 10.73 | 143 | 1.1 | 267 | 274.0 | 307.3 |
| 25 | Distributed | 292.8 | 342.9 | 117.8 | 10.16 | 20.96 | 57.15 | 1.3 | 279.6 | 234.9 | 337.0 |
| 26 | One-point | 395 | 381 | 101.6 | 5.84 | 6.83 | 28.58 | 3 | 352 | 365.0 | 468.9 |
| 27 | Distributed | 292.8 | 266.7 | 101.6 | 6.35 | 9.83 | 44.5 | 1.5 | 113.7 | 131.4 | 63.6 |
| 28 | Distributed | 320 | 229 | 76.2 | 5.84 | 9.58 | 38.1 | 1.15 | 196 | 137.7 | 114.1 |
| 29 | One-point | 290.8 | 451.6 | 124.1 | 7.62 | 10.66 | 146 | 2 | 280 | 280.2 | 239.9 |
| 30 | Distributed | 230 | 280 | 100 | 10.8 | 5.7 | 40 | 3.33 | 194.1 | 314.4 | 322.5 |
| 31 | One-point | 335 | 381 | 101.6 | 5.84 | 6.83 | 127 | 1.6 | 290 | 237.9 | 261.8 |
| 32 | Distributed | 292.8 | 381 | 101.6 | 5.84 | 6.83 | 101.6 | 2.4 | 89 | 174.1 | 106.5 |
| 33 | One-point | 335 | 381 | 101.6 | 5.84 | 6.83 | 127 | 2.4 | 303.4 | 227.8 | 222.9 |
| 34 | One-point | 352.9 | 380.5 | 66.9 | 3.56 | 4.59 | 66.55 | 2.44 | 84.4 | 158.9 | 120.5 |
| 35 | One-point | 438 | 381 | 101.6 | 5.84 | 6.83 | 165.1 | 3.5 | 277.2 | 298.2 | 558.4 |
| 36 | Distributed | 292.8 | 381 | 101.6 | 5.84 | 6.83 | 88.9 | 2.5 | 89 | 170.5 | 106.8 |
| 37 | Distributed | 335 | 381 | 101.6 | 5.84 | 6.83 | 101.6 | 1.37 | 300 | 229.5 | 322.3 |
| 38 | Distributed | 335 | 381 | 101.6 | 5.84 | 6.83 | 165.1 | 1.6 | 358.6 | 249.0 | 285.6 |
| 39 | One-point | 304 | 459.6 | 103 | 7.15 | 10.7 | 157 | 2.3 | 273 | 254.8 | 236.5 |
| 40 | One-point | 290 | 524.3 | 124.4 | 7.02 | 10.73 | 176 | 1.35 | 280 | 257.8 | 255.9 |
| 41 | One-point | 295.5 | 526.1 | 124.7 | 7.08 | 10.77 | 176 | 1.35 | 260 | 271.0 | 284.9 |
| 42 | One-point | 335 | 229 | 76.2 | 5.84 | 9.58 | 38.1 | 1.32 | 117.6 | 148.6 | 125.6 |
| 43 | Distributed | 335 | 381 | 101.6 | 5.84 | 6.83 | 165.1 | 1.75 | 310.3 | 247.3 | 278.3 |
| 44 | One-point | 294.7 | 524.1 | 124.3 | 7.07 | 10.73 | 177 | 1.35 | 240 | 269.5 | 279.4 |
| 45 | Two-point | 352.9 | 380.5 | 66.9 | 3.56 | 4.59 | 66.55 | 1.22 | 92.7 | 163.9 | 199.4 |
| 46 | Distributed | 292.8 | 381 | 203.2 | 10.16 | 19.89 | 63.5 | 1.3 | 310.5 | 282.1 | 352.5 |
| 47 | One-point | 297.3 | 450.3 | 123.4 | 7.54 | 10.68 | 150 | 1.15 | 279 | 295.5 | 294.9 |

available data sets were randomly divided into learning, validation, and testing subsets. The learning data were taken for training (genetic evolution). The validation data were used to specify the generalization capability of the models on data they did not train on (model selection). Thus, both the learning and the validation data were involved in the modeling process and were categorized into one group referred to as "training data". The model with the best performance on both the learning and the validation data sets was finally selected as the outcome of the runs. The testing data were further employed to measure the performance of the optimal model obtained by GEP on data that played no role in building the models. To obtain consistent data division, several combinations of the training and testing sets were considered. The selection was in a way that the maximum, minimum, mean, and standard deviation of parameters were consistent in the training and testing data sets. Of the 47 data sets, 38 data vectors were taken for the training process (30 sets for learning and 8 sets for validation). The remaining 9 sets were used for the testing of the derived model.

Please cite this article in press as: Gandomi AH, et al. A new prediction model for the load capacity of castellated steel beams. Journal of Constructional Steel Research (2011), doi:10.1016/j.jcsr.2011.01.014

### 3.2. Performance measures

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 learning set of data.
iii. Providing the best fitness value on a validation set of data.

The first objective was controlled by the user through the parameter settings (e.g., number of genes or head size). For the other objectives, the following objective function (OBJ) was constructed as a measure of how well the model predicted output agrees with the experimental results. The best GEP model was then deduced by the minimization of the following function:

$$\mathrm{OBJ} = \left(\frac{\mathrm{No.}_{\mathrm{Learning}} - \mathrm{No.}_{\mathrm{Validation}}}{\mathrm{No.}_{\mathrm{Training}}}\right)\frac{\mathrm{MAE}_{\mathrm{Learning}}}{R^2_{\mathrm{Learning}}} + \frac{2\mathrm{No.}_{\mathrm{Validation}}}{\mathrm{No.}_{\mathrm{Training}}}\frac{\mathrm{MAE}_{\mathrm{Validation}}}{R^2_{\mathrm{Validation}}}, \tag{5}$$

where $\mathrm{No.}_{\mathrm{Train}}$, $\mathrm{No.}_{\mathrm{Learning}}$, and $\mathrm{No.}_{\mathrm{Validation}}$ are respectively the number of training, learning, and validation data. $R$ and MAE are respectively the correlation coefficient and mean absolute error given in the form of the following relationships:

$$R = \frac{\sum_{i=1}^{n}(h_i - \overline{h_i})(t_i - \overline{t_i})}{\sqrt{\sum_{i=1}^{n}(h_i - \overline{h_i})^2 \sum_{i=1}^{n}(t_i - \overline{t_i})^2}} \tag{6}$$

$$\mathrm{MAE} = \frac{\sum_{i=1}^{n}|h_i - t_i|}{n}, \tag{7}$$

in which $h_i$ and $t_i$ are respectively the actual and calculated outputs for the $i$th output; $\overline{h_i}$ and $\overline{t_i}$ are the average of the actual and calculated outputs, and $n$ is the number of samples. It is well known that $R$ alone is not a good indicator for predicting the accuracy of a model. This is because, on equal shifting of the output values of a model, the $R$ value does not change. The constructed objective function takes into account the changes of $R$ and MAE simultaneously. Higher $R$ and lower MAE values result in lowering OBJ, and hence indicate a more precise model. In addition, the above function takes the effects of different data divisions into account for the learning and validation data.

### 3.3. Development of an empirical model using GEP

Eight input parameters, $F_{yw}$, $h_c$, $B$, $t_w$, $t_f$, $S$, $L$, and $LC$, were used to create the GEP model. Various parameters involved in the GEP predictive algorithm are shown in Table 3. The parameter selection will affect the model generalization capability of GEP. Several runs were conducted to come up with a parameterization of GEP that provided enough robustness and generalization to solve the problem. The number of programs in the population that GEP will evolve is set by the population size (number of chromosomes). A run will take longer with a larger population size. The proper population size depends on the number of possible solutions and the complexity of the problem. Three levels were set for the population size. The chromosome architectures of the models evolved by GEP include the head size and the number of genes. The head size determines the complexity of each term in the evolved model. The number of terms in the model is determined by the number of genes per chromosome. Each gene codes for a different sub-expression tree or sub-ET. Three optimal levels were considered for the head size and the number of genes. For the number of genes greater than 1, the addition linking function was used to link the mathematical terms encoded in each gene. One level was considered for the other parameters based on some previously suggested values [13,19–21] and also after making several preliminary runs and observing the performance behavior. There are $3 \times 3 \times 3 = 27$ different combinations of the parameters. All of these combinations were tested, and 10 replications for each combination were carried out. Therefore, the overall number of runs was equal to $27 \times 10 = 270$. The period of time acceptable for evolution to occur without improvement in best fitness is set via the generations without change parameter. After 2000 generations considered herein, a mass extinction or a neutral gene was automatically added to the model. In this study, basic arithmetic operators and mathematical functions were utilized to get the optimum GEP model. The mean absolute error function was used to calculate the overall fitness of the evolved programs. The program was run until there was no longer any significant improvement in the performance of the models. The GEP algorithm was implemented using GeneXproTools [33].

**Table 3**
Parameter settings for the GEP algorithm.

| Parameter settings | | |
|---|---|---|
| General | | |
| | Chromosomes | 50, 150, 300 |
| | Genes | 1, 2, 3 |
| | Head size | 3, 5, 8 |
| | Tail size | 9 |
| | Dc size | 9 |
| | Gene size | 26 |
| | Linking function | Addition |
| Complexity increase | | |
| | Generations without change | 2000 |
| | Number of tries | 3 |
| | Maximum complexity | 5 |
| Genetic operators | | |
| | Mutation rate | 0.044 |
| | Inversion rate | 0.1 |
| | IS transposition rate | 0.1 |
| | RIS transposition rate | 0.1 |
| | One-point recombination rate | 0.3 |
| | Two-point recombination rate | 0.3 |
| | Gene recombination rate | 0.1 |
| | Gene transposition rate | 0.1 |
| Numerical constants | | |
| | Constants per gene | 2 |
| | Data type | Integer |
| | Lower bound | −10 |
| | Upper bound | 10 |

#### 3.3.1. GEP-based formulation for the load capacity of a CSB

The GEP-based formulation of the failure load ($P$) in terms of $F_{yw}$, $h_c$, $B$, $t_w$, $t_f$, $S$, $L$, and $LC$ is as given below:

$$P_{\mathrm{GEP}}(kN) = t_w\left(L\left(\frac{F_{yw} - B}{S} - L\right) + (t_w - 4)(t_w - LC) + \sqrt[3]{(h_c + F_{yw})(F_{yw} - t_w - 216)}\right) - (t_f - LC)^2 + (\sqrt[3]{t_f + h_c} - 10)\left(t_w^2 - \frac{S}{6}\right). \tag{8}$$

The expression tree of the above formulation is given in Fig. 7. A comparison of the predicted against experimental failure load values is shown in Table 2 and Fig. 8.

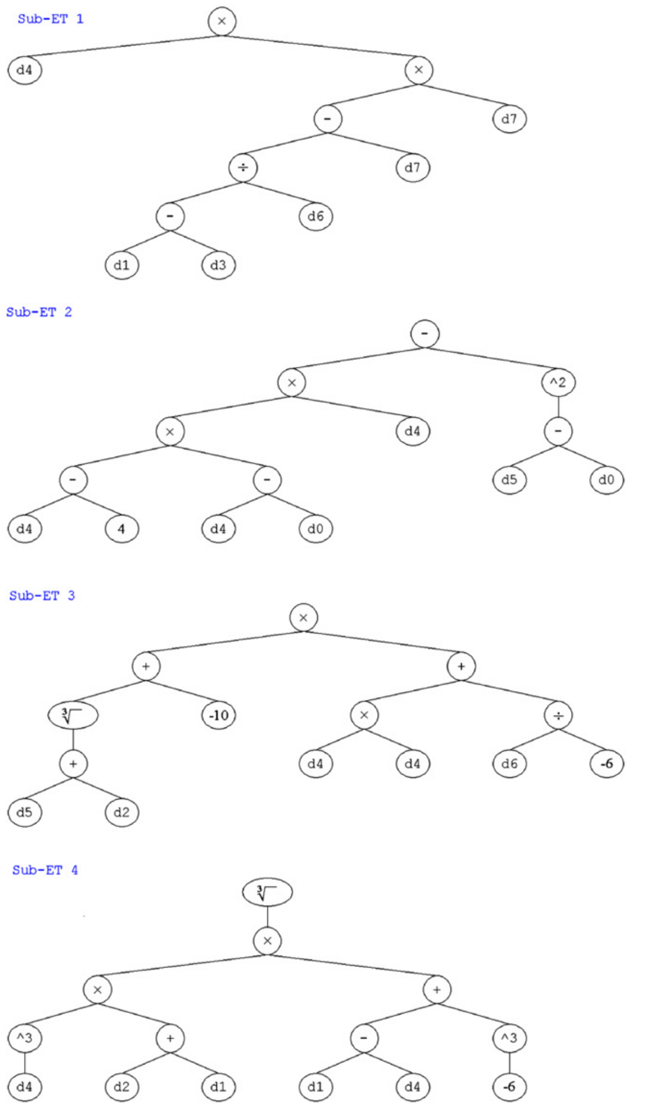


**Fig. 7.** Expression tree for the load capacity of a CSB; where $d_1 = LC, d_2 = F_{yw}, d_3 = h_c, d_4 = B, d_5 = t_w, d_6 = t_f, d_7 = S, d_8 = L$.

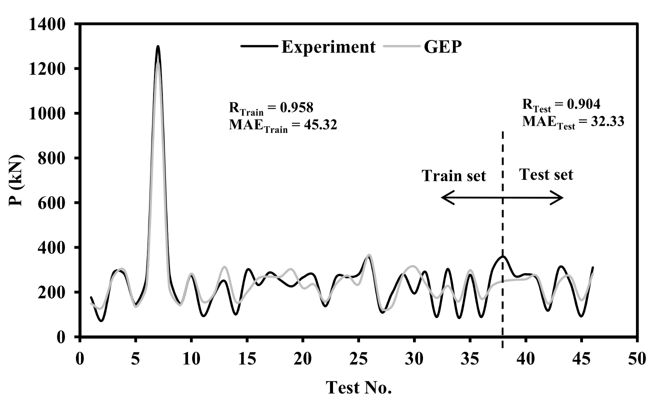


**Fig. 8.** Experimental versus predicted failure load values using the GEP model.

### 3.4. Development of an empirical model using regression analysis

In the conventional material modeling process, regression analysis is an important tool in building models. In the present study, a multivariable least squares regression (LSR) analysis [34] was initially performed to get an idea of the predictive power of the GEP technique in comparison with a classic statistical approach. The LSR method is extensively used in primary regression analysis due to its interesting nature. LSR minimizes the sum-of-squared residuals for each equation, accounting for any cross-equation restrictions on the parameters of the system. If no restrictions exist, this technique becomes identical to estimating each equation by single-equation ordinary least squares. The EViews software package [35] was used to perform the regression analysis. The formulation of the failure load ($P$), for the best results using the LSR analysis, is as given below:

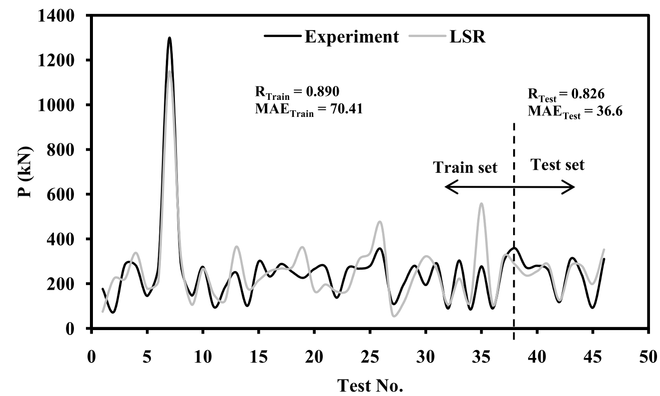


**Fig. 9.** Experimental versus predicted failure load values using the LSR model.

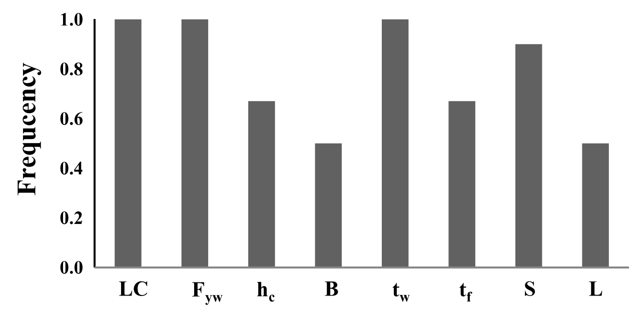


**Fig. 10.** Contributions of the predictor variables in the GEP model.

$$P_{LSR.}(kN) = 19.564LC + 3.926F_{yw} + 0.937h_c - 0.455B + 111.794t_w - 19.869t_f - 0.403S - 48.641L - 1772.193. \quad (9)$$

A comparison of the experimental against predicted failure load values is shown in Table 2 and Fig. 9.

## 4. Sensitivity analysis

Sensitivity analysis is of utmost concern for selecting the important input variables. The contribution of each predictor variable in the GEP model was evaluated through a sensitivity analysis. For this aim, frequency values of the input variables were obtained. A frequency value equal to 1.00 for an input indicates that this variable has appeared in 100% of the best 30 programs evolved by GEP. This methodology is a common approach in GP-based analyses [13–15]. The frequency values of the predictor variables are presented in Fig. 10. According to these figures, the failure load is more sensitive to $LC$, $F_{yw}$, and $t_w$. This can be regarded as an expected case from a structural engineering point of view.

## 5. Performance analysis

A new model was developed for the estimation of the failure load upon a reliable database. Based on a logical hypothesis [36], if a model gives $R > 0.8$, and the MAE values are at the minimum,

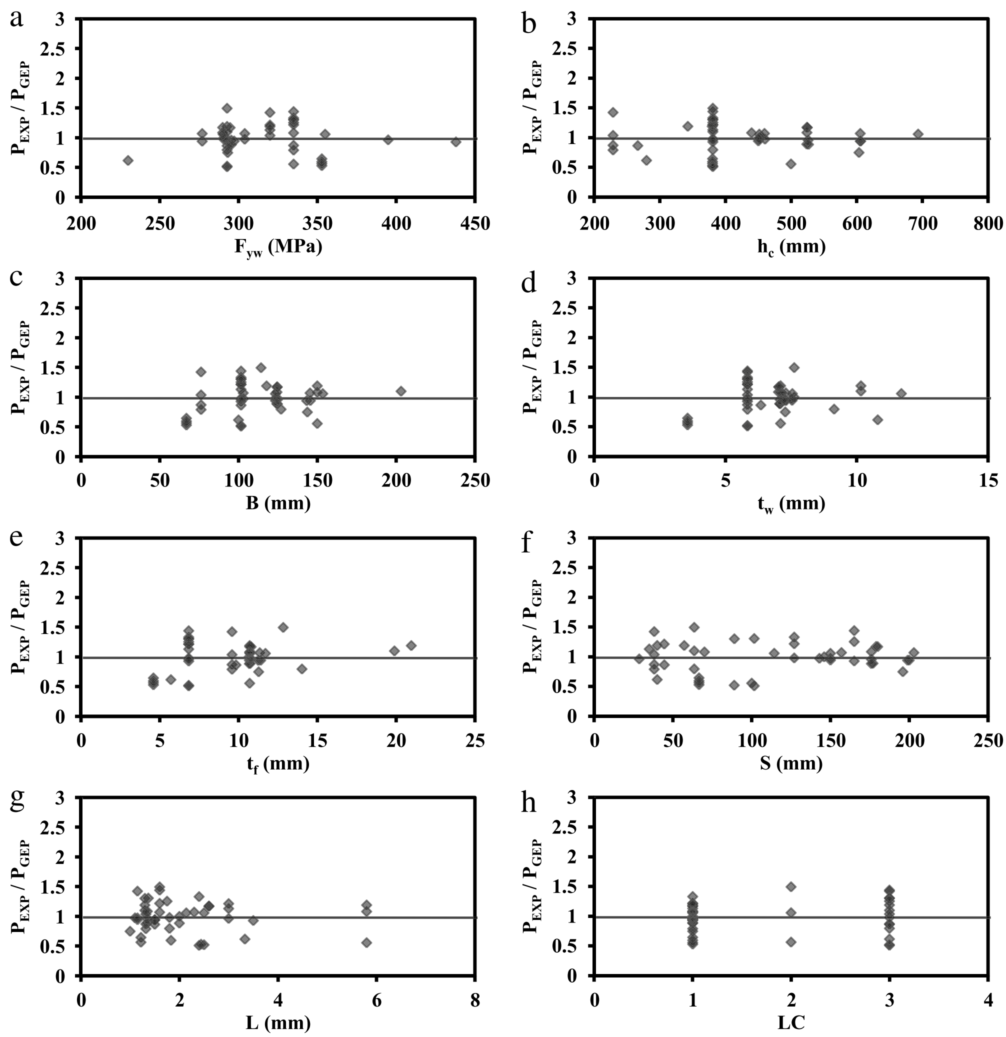


**Fig. 11.** The ratio between the experimental and predicted failure load values with respect to the design parameters (average = 0.988).

there is a strong correlation between the predicted and measured values [23]. The model can therefore be judged as very good. Based on the results, the proposed GEP model with high $R$ and low MAE values is able to predict the target values to an acceptable degree of accuracy. The performance of the model on the training and testing data suggests that it has both good predictive ability and generalization performance.

The models derived using soft computing techniques (e.g., neural networks or GP-based approaches) mostly have a predictive capability within the data range used for their development. The amount of data used for the training process of these techniques is an important issue, as it bears heavily on the reliability of the final models. To cope with this limitation, Frank and Todeschini [37] argue that the minimum ratio of the number of objects over the number of selected variables for model acceptability is 3. They also suggest that considering a ratio equal to 5 is more reasonable. In the present study, this ratio is higher, and is equal to $47/8 = 5.9$. The above facts ensure that the final GEP model has prediction power and is not a chance correlation. Furthermore, new criteria recommended by Golbraikh and Tropsha [38] were checked for the external validation of the models on the validation data sets. It is suggested that at least one of the slopes of the regression lines ($k$ or $k'$) through the origin should be close to 1. The performance indices of $m$ and $n$ should be lower than 0.1. Also, either the squared correlation coefficient (through the origin) between the predicted and experimental values ($Ro^2$) or the coefficient between the experimental and predicted values ($Ro'^2$) should be close to 1 [23]. The considered validation criteria and the relevant results obtained by the models are presented in Table 4. The results of the model validity indicate that the derived GEP model is strongly valid.

**Table 4**
Statistical parameters of the GEP model for external validation.

| Item | Formula | Condition | GEP |
|---|---|---|---|
| 1 | $R$ | $R > 0.8$ | 0.904 |
| 2 | $k = \frac{\sum_{i=1}^{n}(h_i \times t_i)}{h_i^2}$ | $0.85 < K < 1.15$ | 1.003 |
| 3 | $k' = \frac{\sum_{i=1}^{n}(h_i \times t_i)}{t_i^2}$ | $0.85 < K' < 1.15$ | 0.974 |
| 4 | $m = \frac{R^2 - Ro^2}{R^2}$ | $m < 0.1$ | −0.225 |
| 5 | $n = \frac{R^2 - Ro'^2}{R'^2}$, | $n < 0.1$ | −0.203 |
| where | $Ro^2 = 1 - \frac{\sum_{i=1}^{n}(t_i - h_i^o)^2}{\sum_{i=1}^{n}(t_i - \bar{t}_i)^2}$, $h_i^o = k \times t_i$ | | 1.000 |
| | $Ro'^2 = 1 - \frac{\sum_{i=1}^{n}(h_i - t_i^o)^2}{\sum_{i=1}^{n}(h_i - \bar{h}_i)^2}$, $t_i^o = k' \times h_i$ | | 0.983 |

In addition, Fig. 11 shows the ratio of the experimental to the GEP predicted failure against different parameters. As the scattering increases in these figures, the accuracy of the model decreases. It can be observed that, with the exception of $LC$, the scattering slightly decreases with increasing different parameters. In the case of $LC$, the results do not exhibit any noticeable trend. It is worth noting that the observed deviation between the experimental and predicted failure load is not only due to the deficiency of the proposed model. It can partly be attributed to uncertainties, errors, and inconsistencies in the data used for the training and testing of the model.

The results presented in Figs. 8 and 9 indicate that the GEP-based formula significantly outperforms the LSR model on both the training and the testing sets. However, no rational model has been developed for the prediction of failure load including the variables considered in this study. Thus, it is not possible to conduct a comparative study between the results of this research and those in hand.

Note that one of the major advantages of the GEP approach over traditional regression analysis is its ability to derive explicit relationships for failure load without assuming prior forms of the existing relationships. The best solution (equation) evolved by this technique is determined after controlling numerous preliminary models, even millions of linear and nonlinear models. However, one of the 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. In the initial steps of design, information about the features and properties of targeted output or process are often imprecise and incomplete [39]. Nevertheless, it is idealistic to have some initial estimates of the outcome before performing any extensive laboratory work. The GEP approach employed in this research is based on the data alone to determine the structure and parameters of the models. Thus, the derived constitutive models can be particularly valuable in the preliminary design stages. For more reliability, the results of the GEP-based analyses are suggested to be treated as a complement to conventional computing techniques (such as the finite element method). In any case, the importance of engineering judgment in the interpretation of the results obtained should not be underestimated. In order to develop a sophisticated prediction tool, GEP can be combined with advanced deterministic models. Assuming that the deterministic model captures the key physical mechanisms, it needs appropriate initial conditions and carefully calibrated parameters to make accurate predictions. An idea could be to calibrate the parameters by the use of GEP which takes into account historic data sets as well as the laboratory test results. This allows integrating the uncertainties related to testing conditions which the conventional constitutive models do not explicitly account for [23].

## 6. Conclusion

A robust variant of GP, namely GEP, was utilized to formulate the load capacity of castellated steel beams. An accurate empirical model was derived for the prediction of the failure load. A reliable database from previously published failure load test results was used to develop the model. The following conclusions are drawn based on the results presented.

- The proposed GEP-based model is capable of predicting the failure loads of CSBs with high accuracy. The validity of the model was tested for a part of test results beyond the training data domain. Furthermore, the GEP 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 load capacity estimation of CSBs.
- Due to the nonlinearity in collapse behavior of CSBs, the GEP model produces considerably better outcomes than the multivariable linear regression-based model.
- The proposed model simultaneously takes into account the role of several important factors representing the failure load of CSB behavior.
- They derived equation is very simple and can readily be used for practical pre-planning and pre-design purposes via hand calculations.
- An observation from the results of the sensitivity analysis is that the most important parameters governing the behavior of the load capacity of CSBs are the loading condition, minimum web yield stress, and web thickness.
- The constitutive equation derived using GEP is basically different from the conventional constitutive models based on first principles (e.g., elasticity and plasticity theories) [23]. One of the distinctive features of the GEP-based model is it is based on the experimental data rather than on assumptions made in developing the conventional models. Consequently, as more data become available, this model can be retrained and improved without repeating the development procedures from the beginning.

## Acknowledgement

The authors are thankful to Professor Mohammad Mehdi Alinia (Amirkabir University of Technology, Tehran, Iran) for his support and stimulating discussions.

## References

[1] Aglan AA, Redwood RG. Web buckling in castellated beams. Proceedings of the Institution of Civil Engineers, Part 2: Research and Theory 1974;57:307–20.
[2] Dougherty BK. Castellated beams: a state of the art report. Technical report. Journal of the South African Institution of Civil Engineering. 35 (2). 1993. p. 12–20.
[3] Zirakian T, Showkati H. Distortional buckling of castellated beams. Journal of Constructional Steel Research 2006;62:863–71.
[4] Amayreh L, Saka MP. Failure load prediction of castellated beams using artificial neural networks. Asian Journal of Civil Engineering (Building and Housing) 2005;6(1–2):35–54.
[5] Kerdal D, Nethercot DA. Failure modes for castellated beams. Journal of Constructional Steel Research 1984;4:295–315.
[6] Knowles PR. Castellated beams. Proceedings of the Institution of Civil Engineers, Part 1: Design and Construction 1991;90:521–36.
[7] Knowles PR. Design of castellated beams. (for use with BS 5950 and BS 449). Croydon (UK): Constrado; 1985.
[8] Pala M, Caglar N. A parametric study for distortional buckling stress on cold-formed steel using neural networks. Journal of Constructional Steel Research 2007;63:686–91.
[9] Caglar N, Pala M, Elmas M, Eryilmaz DM. A new approach to determine the base shear of steel frame structures. Journal of Constructional Steel Research 2009;65:188–95.
[10] Guzelbey IH, Cevik A, Erklig A. Prediction of web crippling strength of cold-formed steel sheetings using neural networks. Journal of Constructional Steel Research 2006;62(10):962–73.
[11] Koza JR. Genetic programming: on the programming of computers by means of natural selection. Cambridge (MA): MIT Press; 1992.
[12] Banzhaf W, Nordin P, Keller R, Francone F. Genetic programming—an introduction. In: On the automatic evolution of computer programs and its application. Heidelberg, San Francisco: Dpunkt, Morgan Kaufmann; 1998.
[13] Gandomi AH, Alavi AH, Mirzahosseini MR, Moghadas Nejad F. Non-linear genetic-based models for prediction of flow number of asphalt mixtures. Journal of Materials in Civil Engineering, ASCE 2011;23(3): 1–18.
[14] Alavi AH, Gandomi AH, Sahab MG, Gandomi M. Multi expression programming: a new approach to formulation of soil classification. Engineering with Computers 2010;26(2):111–8.
[15] Gandomi AH, Alavi AH, Sahab MG. New formulation for compressive strength of CFRP confined concrete cylinders using linear genetic programming. Materials and Structures 2010;43(7):963–83.
[16] Gandomi AH, Alavi AH, Kazemi S, Alinia MM. Behavior appraisal of steel semi-rigid joints using linear genetic programming. Journal of Constructional Steel Research 2009;65(8–9):1738–50.
[17] Ferreira C. Gene expression programming: a new adaptive algorithm for solving problems. Complex Systems 2001;13(2):87–129.
[18] Oltean M, Grosan C. A comparison of several linear genetic programming techniques. Complex Systems 2003;14(4):1–29.
[19] Pala M. Genetic programming-based formulation for distortional buckling stress of cold-formed steel members. Journal of Constructional Steel Research 2008;64:1495–504.
[20] Cevik A. A new formulation for web crippling strength of cold-formed steel sheeting using genetic programming. Journal of Constructional Steel Research 2007;63:867–83.
[21] Cevik A. Genetic programming based formulation of rotation capacity of wide flange beams. Journal of Constructional Steel Research 2007;63:884–93.

[22] Javadi AA, Rezania M. Applications of artificial intelligence and data mining techniques in soil modeling. Geomechanics and Engineering 2009;1: 53–74.
[23] Alavi AH, Ameri M, Gandomi AH, Mirzahosseini MR. Formulation of flow number of asphalt mixes using a hybrid computational method. Construction and Building Materials 2011;25:1338–55.
[24] Torres RS, Falcão AX, Gonçalves MA, Papa JP, Zhang B, Fan W, et al. A genetic programming framework for content-based image retrieval. Pattern Recognition 2009;42:283–92.
[25] Ferreira C. Gene expression programming: mathematical modeling by an artificial intelligence. 2nd ed. Germany: Springer-Verlag; 2006.
[26] Sadat Hosseini SS, Gandomi AH. Short-term load forecasting of power systems by gene expression programming. Neural Computing and Applications, (in press) doi:10.1007/s00521-010-0444-y.
[27] Redwood RG, Demirdjian S. Castellated beam web buckling in shear. Journal of Structural Engineering 1998;124(10):1202–7.
[28] Hosain MU, Speirs WG. Experiments on castellated steel beams. Journal of American Welding Society, Welding Research Supplement 1973;8(52): 329S–42.
[29] Okubo T, Nethercot DA. Web post strength in castellated beams. Proceedings of the Institution of Civil Engineers, Part 2: Research and Theory 1985;79: 533–57.
[30] Van Oostrom J, Sherbourne AN. Plastic analysis of castellated beams—part II: analysis and tests. Computers and Structures 1972;2:111–40.
[31] Maalek S, Burdekin FM. Weld quality requirements for castellated beams. Structural Engineer 1991;13(69):243–54.
[32] Sherbourne AN, Van Oostrom J. Plastic analysis of castellated beams—part I: interaction of moment, shear and axial force. Computers and Structures 1972; 2:79–109.
[33] GEPSOFT. GeneXpro Tools. Version 4.0. 2006. Available at: http://www.gepsoft.com.
[34] Ryan TP. Modern regression methods. New York: Wiley; 1997.
[35] Maravall A, Gomez V. EViews software, ver. 5. Irvine (CA): Quantitative Micro Software, LLC; 2004.
[36] Smith GN. Probability and statistics in civil engineering. London: Collins; 1986.
[37] Frank IE, Todeschini R. The data analysis handbook. Amsterdam: Elsevier; 1994.
[38] Golbraikh A, Tropsha A. Beware of $q^2$. Journal of Molecular Graphics and Modelling 2002;20:269–76.
[39] Kraslawski A, Pedrycz W, Nyström L. Fuzzy neural network as instance generator for case-based reasoning system: an example of selection of heat exchange equipment in mixing. Neural Computing and Applications 1999;8: 106–13.