

Please cite this paper as:

Naser M.Z., Thai S., Thai, T. (2021). "Evaluating Structural Response of Concrete-Filled Steel Tubular Columns through Machine Learning." *Journal of Building Engineering*. <https://doi.org/10.1016/j.jobbe.2020.101888>.

## Evaluating Structural Response of Concrete-Filled Steel Tubular Columns through Machine Learning

M.Z. Naser<sup>1,\*</sup>, Son Thai<sup>2,3</sup>, Huu-Tai Thai<sup>2</sup>

<sup>1</sup>Glenn Department of Civil Engineering, Clemson University, Clemson, SC, 29634, USA  
E-mail: [mznaser@clemson.edu](mailto:mznaser@clemson.edu), [m@mznaser.com](mailto:m@mznaser.com), Website: [www.mznaser.com](http://www.mznaser.com)

<sup>2</sup>Department of Infrastructure Engineering, The University of Melbourne, Parkville VIC 3010, Australia

<sup>3</sup>Faculty of Civil Engineering, HCMC University of Technology, VNU-HCM, 268 Ly Thuong Kiet Street, District 10, Ho Chi Minh City, Vietnam

**Abstract:** Concrete-filled steel tubular (CFST) columns are unique structural members that capitalize on the synergy between steel and concrete materials. Due to complexities arising from the interaction between steel tube and concrete filling, the analysis and design of CFST columns are both intricate and tedious. A closer examination to the provisions of American, European and Australian/New Zealand design guidelines shows how these building codes seem to diverge on a proper methodology to design CFST columns. This leverages naturally inspired machine learning (NIML) algorithms (namely genetic algorithms and gene expression programming) to derive compact and one-stepped predictive expressions that can accurately predict the structural response of CFST columns. These expressions were developed and validated using the results of 3,103 available tests carried out on CFST columns over the past few years. The outcome of this work shows that the NIML-derived expressions have superior prediction capabilities than those in currently used design codes.

**Keywords:** Machine Learning, Genetic Algorithms; Gene Programming; CFST columns; Structural response.

## INTRODUCTION

Concrete-filled steel tubular (CFST) columns are attractive solutions for field applications of special demands (e.g., high strength and ductility, excellent energy dissipation and among others (Thai et al., 2019; Yuan et al., 2018)). As such, CFST columns are often used in high-rise buildings as well as in industrial and large-sized structures. In general, CFST columns can offer structural engineers a number of benefits. For example, CFST columns have high strength-to-weight ratio when compared to equivalent columns made of steel or reinforced concrete only. Moreover, CFST columns do not require sophisticated construction and fabrication as the steel tube act as a permanent formwork for concrete casting. This leads to further saving on material and labor costs and accelerating construction (Fike and Kodur, 2011). In fact, the confinement generated by the steel tube enhances the strength of in-filled concrete, while this filling prevents the inward buckling of the steel tube, and thus increasing the overall stability and strength of CFST columns. On a separate note, CFST columns present economical solutions against unique loading conditions as they often satisfy fire resistance requirements without the need for external proofing (Kodur, 1999).

Please cite this paper as:

Naser M.Z., Thai S., Thai, T. (2021). "Evaluating Structural Response of Concrete-Filled Steel Tubular Columns through Machine Learning." *Journal of Building Engineering*. <https://doi.org/10.1016/j.jobbe.2020.101888>.

As a result, CFST columns have captured architects' attention and promoted the integration of aesthetically pleasing exposed steelwork.

An extensive testing program was carried out by Knowles and Park (1969) in the late 1960s to investigate the response of CFST columns under concentric and eccentric loading. This was followed by Liu and Goel (2008) who examined the response of CFST columns under cyclic loading in the 1980s. During the 1980-1990s, a number of testing programs ran in parallel, e.g. Lie and Stringer (1994) explored the fire resistance, Kilpatrick and Rangan (1999) studied effect of high strength in-filled concrete. Then, in early 2000s, Sakino et al. (2004) examined the influence of steel tube shape and strength, tube diameter-to-thickness ratio and concrete strength on 114 CFST columns, and proposed design formulas to estimate their ultimate axial load capacity. The outcomes of the above investigations as well as many others were compiled and documented into databases. Some of these databases were developed by a number of researchers such as Thai et al. (2019), Goode (2008), Leon et al. (2011), Hajjar (2002), Tao et al. (2008), Liew et al. (2016) and Denavit (2019a). The aim of compiling these databases is to provide a permanent record of tests on various CFST columns to enable development and verification of codal provisions and design expressions.

While it is true that there are a number of codal provisions (e.g., American code AISC 360-16 (2016a), Eurocode 4 (2004), Chinese code GB 50936 (2014), Australian/New Zealand code AS/NZS 2327 (2017)) that can be used to design CFST columns, these provisions are only applicable for CFST columns with a certain section slenderness ratio and material grade. For instance, AISC 360 permits the use of steel tubes with a yield stress up to 525 MPa with concrete strength reaching 69 MPa. Eurocode 4 also limits the yield stress of steel tubes and compressive strength of concrete filling to 460 MPa and 50 MPa, respectively. The Chinese building code further limits the yield stress of steel tubes beyond that in the American and European codes to 420 MPa, while allows the use of comparatively high strength concrete (up to 67 MPa). It is worth noting that the recently released AS/NZS 2327 is the most accommodating between all other codes in which it allows higher strength materials (steel with a yield stress up to 690 MPa and concrete with compressive strength up to 100 MPa.) to be used in CFST columns. It should be noted that the nature of the material strength and section slenderness limitations of modern design codes are due to the lack of the experimental data carried out on the specimens beyond the code limits. For example, only less than 1% of over 3,100 test specimens recently collected by Thai et al., (2019) was carried out on CFST sections with the slenderness ratio beyond the code limit.

From the perspective of this work, the continuous improvement in materials science has led to developing high strength and ultra-high performance concrete and high strength steel (concrete compressive strength over 90 MPa and steel yield stress over 690 MPa) which possess much improved characteristics and properties than those traditionally used in construction (Naser, 2019a). Thus, their incorporation into civil applications not only enables realizing durable and resilient constructions, but also optimally designs structures by allowing the use of more efficient (slender) cross sections. Unfortunately, and as discussed above, current design provisions for

Please cite this paper as:

Naser M.Z., Thai S., Thai, T. (2021). "Evaluating Structural Response of Concrete-Filled Steel Tubular Columns through Machine Learning." *Journal of Building Engineering*. <https://doi.org/10.1016/j.jobbe.2020.101888>.

CFST columns may not be properly extended to allow the use of these high strength materials nor CFST columns with slender sections due to limited test data.

The above code limitations can be overcome in this work by adopting artificial intelligence (AI) as a tool to analyze and process database collected from actual tests on CFST columns. In order to pursue such goal, this paper explores utilizing naturally inspired machine learning (NIML) algorithms to comprehend hidden relations between influencing geometric and material factors to derive design expressions which are capable of accurately predicting the load-carrying capacity of CFST columns made of traditional or modern construction materials. These algorithms were trained and validated against a comprehensive database of 3,103 tests covering a wide range of material and geometric properties as well as loading configurations. The outcome of this work shows the practicality of machine learning tools in predicting the load-carrying capacity of CFST columns as well as in paving the way towards developing AI-driven, unified and modern design approaches.

## DESCRIPTION OF DATABASE

The successful application of NIML algorithms requires compiling a comprehensive database. As such, a literature survey was carried out to identify well-documented tests on CFST columns as well as previously developed databases. In pursuit of compiling a large database, the results of 3,103 tests were collected from 173 studies. This database comprises of columns of various configurations (short, slender, circular, square/rectangular, built-up and roll-formed sections) as well as columns under concentric and eccentric loading. It is worth noting that the compiled database incorporated the data collected by Goode (2008) and Denavit (2019). While these two researchers also collected the test data on preloaded columns and columns made of stainless steel and, these columns were not of interest to this work and hence were not added to the developed database. It should be noted that the complete list of these columns can be found elsewhere (Thai et al., 2019).

Various geometric and material properties were collected for each CFST specimen. The geometric features include physical dimensions of CFST columns, i.e., effective length ( $L_e$ ), tube thickness ( $t$ ) and tube diameter ( $D$ ) for circular columns, eccentricities at end supports ( $e_t$ ,  $e_b$ ), bending axis ( $X$ ) as well as height ( $H$ ) and width ( $B$ ) for square/rectangular columns. The material properties include elastic modulus ( $E_s$ ), yield stress ( $f_y$ ) and ultimate stress ( $f_u$ ) of steel as well as compressive strength ( $f_c$ ) and modulus ( $E_c$ ) of in-filled concrete. Table 1 summarizes main attributes of the collected database in terms of material and geometric features. It should be noted that the concrete compressive strength obtained from the tests was based on both available cylinder and cube specimens, and the cube strength will be converted into cylinder strength to be used in design equations.

Please cite this paper as:

Naser M.Z., Thai S., Thai, T. (2021). "Evaluating Structural Response of Concrete-Filled Steel Tubular Columns through Machine Learning." *Journal of Building Engineering*. <https://doi.org/10.1016/j.jobbe.2020.101888>.

**Table 1** Key statistics from compiled database.

Section	No. of tests	Features	D or H (mm)	t (mm)	D/t or H/t	L <sub>e</sub> (mm)	f <sub>y</sub> (MPa)	f <sub>c</sub> (MPa)
Circular (concentric loading)	1245	Min	44.45	0.52	7.42	152.35	178.28	9.17
		Max	1020.00	16.54	220.93	5560.00	853.00	193.30
		Average	158.52	4.31	44.28	1060.53	336.35	50.21
		St. deviation	105.42	2.45	32.37	1005.28	90.89	31.57
		Median	127.3	4.00	33.33	662.00	325.00	41.00
		Skewness	3.71	1.58	2.86	1.98	2.18	2.06
Square/ Rectangular (concentric loading)	979	Min	40.00	0.70	10.49	60.00	115.00	8.52
		Max	400.00	12.50	285.71	4500.00	835.00	164.10
		Average	154.19	4.49	41.49	965.31	400.55	54.40
		St. deviation	54.85	2.12	27.41	869.23	169.45	32.79
		Median	150	4.31	34.13	595	340.1	44.0
		Skewness	1.14	1.24	3.12	1.61	1.24	1.22
Circular (eccentric loading)	485	Min	76.00	0.86	13.69	284.5	185.7	18.4
		Max	600.00	16.00	220.93	4956.00	517.00	184.00
		Average	143.28	4.22	39.75	1758.46	327.66	51.45
		St. deviation	55.83	1.972	27.07	1042.2	58.99	27.300
		Median	133	4.5	33.33	1700	322	42.2
		Skewness	2.87	2.17	4.36	0.81	0.5	2.28
Square/ Rectangular (eccentric loading)	394	Min	76.20	1.90	15.00	360	242	183
		Max	323.00	12.50	82.00	4910	761	18.76
		Average	154.47	4.50	37.71	1801.78	384.36	57.65
		St. deviation	49.80	1.67	16.02	1112.72	117.94	30.74
		Median	150	4.18	33.33	1814.5	340	47.1
		Skewness	1.13	1.96	0.9	0.48	1.51	1.16

Table 1 shows that the collected columns cover the full spectrum of practical scenarios. For instance, the minimum and maximum diameters of circular columns and the width of square/rectangular columns are 44.45 mm, 1020.00 mm, 40.00 mm and 400.00 mm, respectively. Both circular and square/rectangular columns have thickness varying between 0.52-16.54 mm and 0.70-12.50 mm, respectively. The average yield strength of steel tubes and compressive strength of concrete filling in concentrically loaded circular and square/rectangular columns are 336.35 MPa and 50.21 MPa, and 400.55 MPa and 54.40 MPa, respectively. However, the ranges of these materials are from 9.17 MPa to 193.30 MPa for concrete, and from 115.00 MPa to 853.00 MPa for steel. Table 1 also shows that these ranges are slightly lower in the case of eccentrically loaded columns. The reader is encouraged to remember that these ranges exceed that adopted in currently

Please cite this paper as:

Naser M.Z., Thai S., Thai, T. (2021). “Evaluating Structural Response of Concrete-Filled Steel Tubular Columns through Machine Learning.” *Journal of Building Engineering*. <https://doi.org/10.1016/j.jobbe.2020.101888>.

used design codes as mentioned in the previous section. As such, this database can come in handy in developing NIML algorithms with improved prediction capabilities that could be more inclusive than that in codal provisions.

Once the database was properly set-up, a sensitivity/correlation analysis was carried out to identify key geometric features and material properties that have significant influence on the strength of CFST columns. The outcome of this analysis shows that all identified geometric features ( $L_e$ ,  $t$ ,  $D$ ,  $H$ ,  $B$ ) seem to be high importance to CFST columns. Similarly, both yield stress of steel and compressive strength of concrete also have high relevance. It should be noted that influencing parameters for eccentrically loaded CFST columns include the magnitude and direction of eccentricity. The correlation matrices for the cases of circular and rectangular columns are shown in Table 2. These matrices indicate that cross sectional size, thickness of steel tube and material properties of CFST columns hold the highest correlation, while the length of columns holds the lowest correlation.

Please cite this paper as:

Naser M.Z., Thai S., Thai, T. (2021). "Evaluating Structural Response of Concrete-Filled Steel Tubular Columns through Machine Learning." *Journal of Building Engineering*. <https://doi.org/10.1016/j.jobbe.2020.101888>.

**Table 2** Correlation matrices for influencing parameters.

Circular CFST columns (concentrically loaded)											
	$D$	$f_c$	$f_y$	$L_e$	$t$	$N$	Importance				
$D$	1.000	-0.003	0.072	0.201	0.478	0.911	0.911				
$f_c$	-0.003	1.000	0.030	-0.154	-0.022	0.126	0.146				
$f_y$	0.072	0.030	1.000	0.080	0.238	0.145	0.144				
$L_e$	0.201	-0.154	0.080	1.000	0.216	0.109	0.108				
$t$	0.478	-0.022	0.238	0.216	1.000	0.549	0.544				
$N$	0.911	0.126	0.145	0.109	0.549	1.000	-				
Square/Rectangular CFST columns (concentrically loaded)											
	$B$	$f_c$	$f_y$	$H$	$L_e$	$t$	$N$	Importance			
$B$	1.000	-0.011	0.084	0.878	0.001	0.165	0.655	0.655			
$f_c$	-0.011	1.000	0.462	- 0.072	-0.018	0.376	0.502	0.502			
$f_y$	0.084	0.462	1.000	0.046	0.087	0.445	0.542	0.542			
$H$	0.878	-0.072	0.046	1.000	0.012	0.136	0.598	0.598			
$L_e$	0.001	-0.018	0.087	0.012	1.000	-0.095	-0.117	0.117			
$t$	0.165	0.376	0.445	0.136	-0.095	1.000	0.621	0.621			
$N$	0.655	0.502	0.542	0.598	-0.117	0.621	1.000	-			
Circular CFST columns (eccentrically loaded)											
	$D$	$e_b$	$e_t$	$f_c$	$f_y$	$L_e$	$t$	$N$	Importance		
$D$	1.000	0.598	0.635	0.138	0.239	0.106	0.469	0.785	0.785		
$e_b$	0.598	1.000	0.855	- 0.105	0.273	0.093	0.318	0.225	0.225		
$e_t$	0.635	0.855	1.000	- 0.105	0.320	0.106	0.324	0.217	0.217		
$f_c$	0.138	-0.105	0.105	1.000	0.248	0.194	0.191	0.486	0.486		
$f_y$	0.239	0.273	0.320	0.248	1.000	0.309	0.226	0.199	0.199		
$L_e$	0.106	0.093	0.106	0.194	0.309	1.000	0.312	0.033	0.033		
$t$	0.785	0.225	0.217	0.486	0.199	0.033	1.000	0.535	0.535		
$N$	0.469	0.318	0.324	0.191	0.226	0.312	0.535	1.000	-		
Square/Rectangular CFST columns (eccentrically loaded)											
	$B$	$X$	$e_b$	$e_t$	$f_c$	$f_y$	$H$	$L_e$	$t$	$N$	Importance
$B$	1.000	-0.083	0.464	0.452	0.006	0.153	0.709	-0.256	0.152	0.551	0.551
$X$	-0.083	1.000	- 0.136	- 0.111	- 0.095	- 0.183	- 0.121	0.246	-0.082	-0.220	0.220

Please cite this paper as:

Naser M.Z., Thai S., Thai, T. (2021). "Evaluating Structural Response of Concrete-Filled Steel Tubular Columns through Machine Learning." *Journal of Building Engineering*. <https://doi.org/10.1016/j.jobbe.2020.101888>.

$e_b$	0.464	-0.136	1.000	0.976	0.188	0.192	0.460	-0.301	0.148	-0.014	0.014
$e_t$	0.452	-0.111	0.976	1.000	0.145	0.170	0.402	-0.254	0.146	-0.063	0.063
$f_c$	0.006	-0.095	0.188	0.145	1.000	0.010	0.143	0.188	0.200	0.342	0.342
$f_y$	0.153	-0.183	0.192	0.170	0.010	1.000	0.192	-0.033	0.463	0.466	0.466
$H$	0.709	-0.121	0.460	0.402	0.143	0.192	1.000	-0.299	0.185	0.596	0.596
$L_e$	-0.256	0.246	0.301	0.254	0.188	0.033	0.299	1.000	0.216	-0.166	0.596
$t$	0.152	-0.082	0.148	0.146	0.200	0.463	0.185	0.216	1.000	0.503	0.503
$N$	0.551	-0.220	0.014	0.063	0.342	0.466	0.596	-0.166	0.503	1.000	-

$N$ : axial capacity (kN),  $D$ : tube diameter (mm),  $t$ : tube thickness (mm),  $f_y$ : yield stress of steel tube (MPa),  $f_c$ : compressive strength of concrete (MPa),  $B$ : width of tube (mm),  $H$ : height of tube (mm),  $L_e$ : effective length (mm),  $X$ : bending axis (around y-y) = 1.0, (around z-z) = 2.0,  $e_t$  and  $e_b$ : eccentricities at both ends of the column (mm).

Other material properties such as moduli of steel and concrete and ultimate stress of steel were deemed to be low importance. For example, the modulus of steel can be taken as 200 GPa, and this value is common for all grades of steel used in all columns considered herein. Hence, this property will have a minimal effect of the NIML-derived expressions. In a similar manner, the modulus of concrete is directly influenced by concrete compressive strength. Since values for compressive strength is available for all specimens, the interdependency between the modulus property and compressive strength of concrete can be avoided; effectively only incorporating the strength property as an influencing factor. A similar approach was also used to negate the need for ultimate stress of steel. Thus, to maintain a homogenous database that would lead to optimal expressions, these properties are not further discussed herein.

In any case, the NIML algorithms utilized herein are flexible enough to incorporate any parameters, e.g. material properties, geometric features, loading conditions. All that is required is to obtain data on specific parameters and add them to the developed database. For example, the elastic modulus of concrete filling can be estimated through equations adopted in relative design codes and then the calculated value can be input as a new parameter to the developed database.

## DESCRIPTION OF NIML ALGORITHMS

This section summarizes both mathematical and computational backgrounds to the NIML algorithms used in this study. Both genetic algorithms (GA) and gene expression programming (GEP) were used to develop predictive models that can accurately capture the behavior of CFST columns. Detailed descriptions of these algorithms are provided herein while details on other



Please cite this paper as:

Naser M.Z., Thai S., Thai, T. (2021). "Evaluating Structural Response of Concrete-Filled Steel Tubular Columns through Machine Learning." *Journal of Building Engineering*. <https://doi.org/10.1016/j.jobbe.2020.101888>.

algorithms can be found elsewhere (Gharsellaoui et al., 2020; Huang and Burton, 2019; Kaveh et al., 2019; M. Z. Naser, 2019a; Sarir et al., 2019; Yucel et al., 2019).

GA is an umbrella that encompasses evolutionary regression techniques which include genetic programming (GP) and a newly developed subset GEP. Genetic regression is performed through GA or GEP to reach a population of candidate solutions or computer programs. While GA uses space search to arrive at actions and values using a fixed complexity, GEP on the other hand, uses an explicit structure of smaller computer programs (or codes). This method of regression is particularly advantageous in scenarios where statistical methods become complicated and require the use of specialized software with high computational capacity, or when the exact physics of a problem is well established in detail. Both GA and GEP have been applied into various fields such as materials science (Cheu et al., 2004; Naser, 2019), fire engineering (Naser, 2019b, 2019c), geotechnical engineering (Alavi et al., 2010), earthquake engineering and design (El Ansary et al., 2010; Hejazi et al., 2013).

All of the above are evolutionary techniques which were initially developed by Holland (1988), Koza (1992) and Ferreira (2001). These algorithms incorporate a supervised learning process that mimics the natural selection process (i.e. Darwinian evolution) to express hidden relations between a number of factors. These hidden relations are often tied up to a physical phenomenon which is the capacity of CFST columns in this study. The main advantage of these approaches over traditional soft computing techniques is their capability to produce predictive expressions without relying on past formula or relationship. In these techniques, predefined strings of expressions and/or computer programs strive to realize mathematical formulation of the phenomenon on hand. The key variance between GA and GEP lies in their depiction of the final relation between the selected inputs (those listed in Table 2) and output (capacity of CFST columns). While GA creates a binary string that lists actions and values (i.e. equation/expression) that represent the solution, GEP develops computer codes which is often in a tree structure of actions and values that is expressed in functional programming languages such as C++, Matlab and Fortran. In other words, GA searches a data space, while GEP searches a program space. Once developed, these programs can be run in respective software to solve a phenomenon. For a given problem, GEP can develop a macro that can be run using Matlab software, while GA can derive a formula that can be substituted into by hand calculation using Excel spreadsheet. Interested readers are encouraged to review the following references for comprehensive understanding of GA and GEP (Ferreira, 2001; Goldberg and Holland, 1988; Koza, 1992).

In both techniques, a random population of individuals often referred to as "tree" is created to initiate the search for possible solutions. As such, a possible solution in GA/GEP is a ranked tree consisting of functions and terminals. For instance, a function ( $F$ ) may contains basic mathematical operations (+, −, ×, ÷), power functions (^, log, exp), conditional and logic functions (<, ≥, AND, OR), among others. Conversely, the terminal ( $T$ ) consists of arguments as well as numerical constants and/or variables. Both functions and terminals are first randomly generated and then joined together to develop a model in the form of an expression/equation or computer code. Hence,

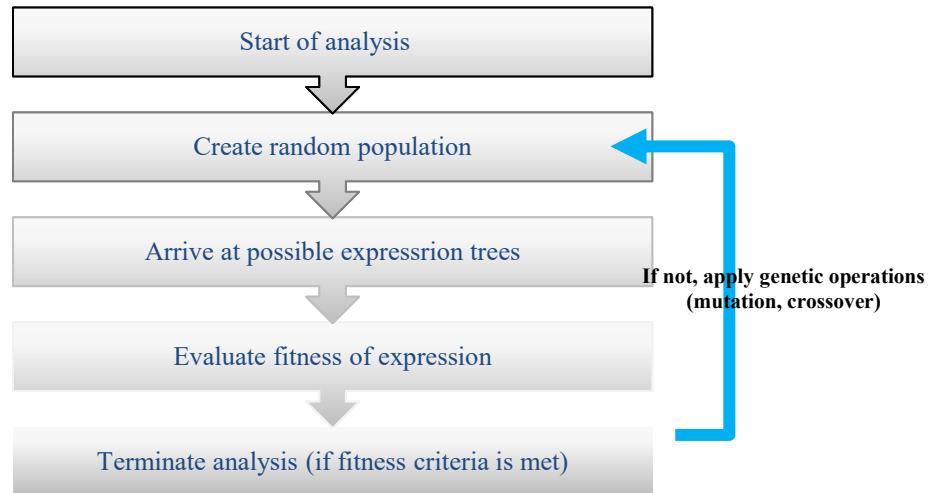


Please cite this paper as:

Naser M.Z., Thai S., Thai, T. (2021). "Evaluating Structural Response of Concrete-Filled Steel Tubular Columns through Machine Learning." *Journal of Building Engineering*. <https://doi.org/10.1016/j.jobbe.2020.101888>.

a developed model has a tree-like configuration in which branches can extend from a function and end in a terminal.

The NIML analysis starts by assessing the fitness of candidates in the solution population. This fitness or ranking of candidate solutions is then applied to evaluate the probability that a particular solution will be nominated for further processing (elitism), or undergo one (or series) of genetic operations. Such operations transform the candidate model by manipulation via reproduction, crossover or mutation (Alavi et al., 2010; Koza, 1992). The reproduction operation assigns a larger selection probability to this successful model. On the other hand, the crossover operation facilitates replacing genetic material code between this particular model and other candidate models within the solution population. Another operation is mutation. In this operation, the algorithm selects a random node in the tree expression of an individual solution to be deleted and then to be substituted with a randomly generated node. The used values for these operations were 0.2%, 0.1% and 0.04%, respectively as suggested by Ferreira (2001). Finally, the fitness for all processed models is calculated and terminated once a convergence condition is met. The fitness of a model is defined as a value that best reflects how far the model's predicted results are from that observed in real life. A flowchart representing the process of deriving a suitable solution is shown in Figure 1.



**Figure 1** A flowchart of analysis procedure

## DEVELOPMENT AND VALIDATION OF NIML-DERIVED EXPRESSIONS

The database was randomly arranged to eliminate any biasness arising from a parameter or experimental program. The database was split into a model development set (for training and validation purposes ~70%) and a testing set (for evaluating performance of applied algorithms after completing training ~30%). The database was then analyzed using GA and GEP, and the outcome is listed in Table 3. This table shows the accuracy of NIML-derived expressions by means of two fitness metrics, i.e., coefficient of determination ( $R^2$ ) and correlation coefficient ( $R$ ). These fitness metrics are close to unity indicating high accuracy of expressions.

Please cite this paper as:

Naser M.Z., Thai S., Thai, T. (2021). "Evaluating Structural Response of Concrete-Filled Steel Tubular Columns through Machine Learning." *Journal of Building Engineering*. <https://doi.org/10.1016/j.jobbe.2020.101888>.

**Table 3** NIML-derived expressions (see appendix for GEP expressions).

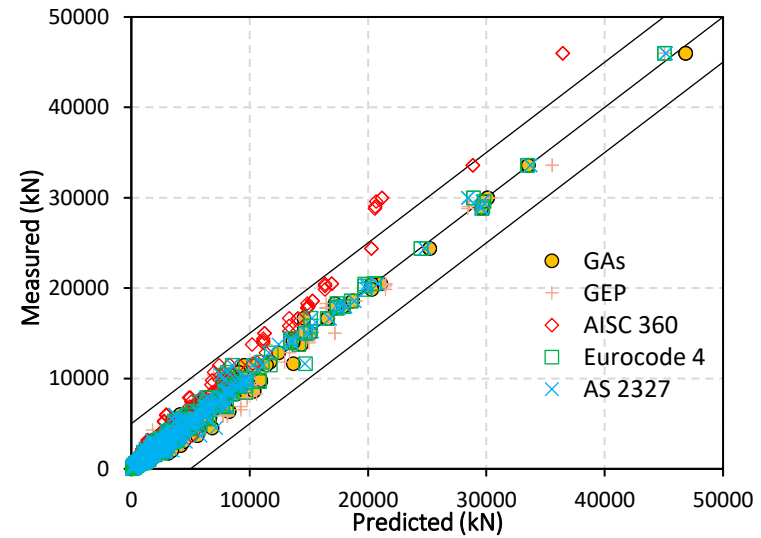
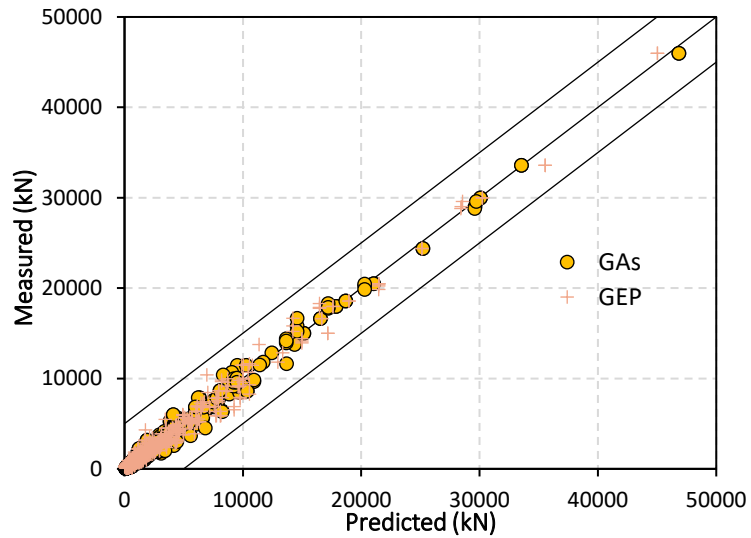
<i>Type</i>			<i>Expression</i>	<i>R</i> <sup>2</sup> (%)
Concentric loading	Circular	GA	$N = \text{abs}(0.00439Dt f_y + 0.000727tD^2 + 0.000727f_c D^2 - 1.38 \times 10^{-5}DL_e f_c - 3.71 \times 10^{-7}DtL_e f_y)$	99.1
		GEP	See appendix	95.8
	Square/ Rectangular	GA	$N = H + 0.00415B^2 + 0.00436Bt f_y + 0.0008f_c H^2 - 1.34 \times 10^{-7}L_e^2 f_y - 3.17 \times 10^{-5}f_y H^2$	97.7
		GEP	See appendix	96.1
Eccentric loading	Circular	GA	$N = \text{abs}(2.012D + 0.2609f_y + Dt + e_t \cos(D) + 0.1403Df_c + 0.001766L_e e_t + \tan(1.233L_e) + \tan(0.1403Df_c) - 0.0004983L_e - 2.3e_t - 2.543e_b - 9.581f_c - 19.01t - 0.04602tL_e - 0.1659f_c e_t)$	97.6
		GEP	See appendix	95.2
	Square/ Rectangular	GA	$N = \text{abs}\left(115.2t + 3.024X + 1.356B + 1.308e_t + 0.1339H + 0.1775Hf_c + 0.01306Bf_y + \frac{0.01306Bf_y}{Xe_t} - 417.5 - 0.0002848f_y - 0.1637L - 3.961e_b - 12.2f_c - 0.1505f_c e_t\right)$	95.4
		GEP	See appendix	93.9

*N*: axial capacity (kN), *D*: tube diameter (mm) , *t*: tube thickness (mm) , *f<sub>y</sub>*: yield strength of tube (MPa), *f<sub>c</sub>*: compressive strength of concrete (MPa) , *B*: width of tube (mm), *H*: height of tube (mm), *L<sub>e</sub>*: effective length (mm) , *X*: bending axis (around y-y) = 1.0, (around z-z) = 2.0, *e<sub>t</sub>* and *e<sub>b</sub>*: eccentricities at both ends of the column (mm).

Figure 2 also shows that the predictions from both GA and GEP algorithms outperform those obtained from current design codes (i.e., American code AISC 360-16, Eurocode 4 and Australian/New Zealand code AS/NZS 2327). In fact, a closer look into Figure 2 also shows the high predictive capability of these expressions in which the majority of data points lie within the bounding error of 10% (at which +10% or -10% of the exact measured values lie). This is not the case for current design codes which tend to severely underperform CFST columns under eccentric loading. It is worth noting that the NIML-derived expressions account a wide range of geometric and material properties including those beyond the aforementioned design codes.

Please cite this paper as:

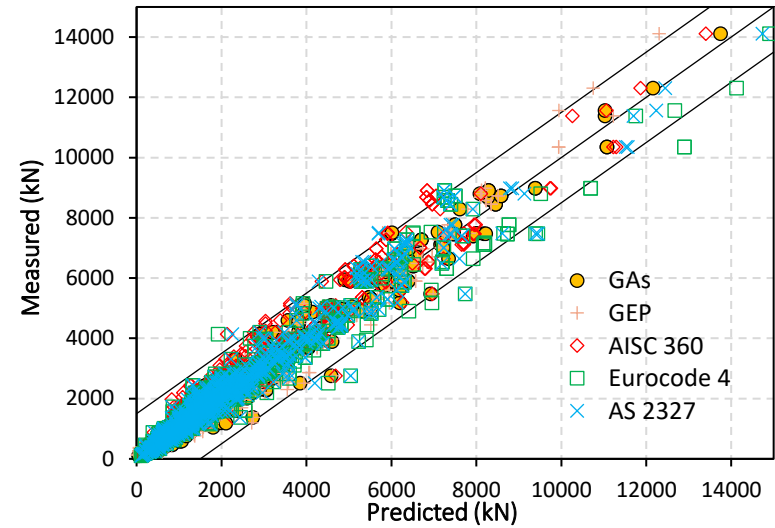
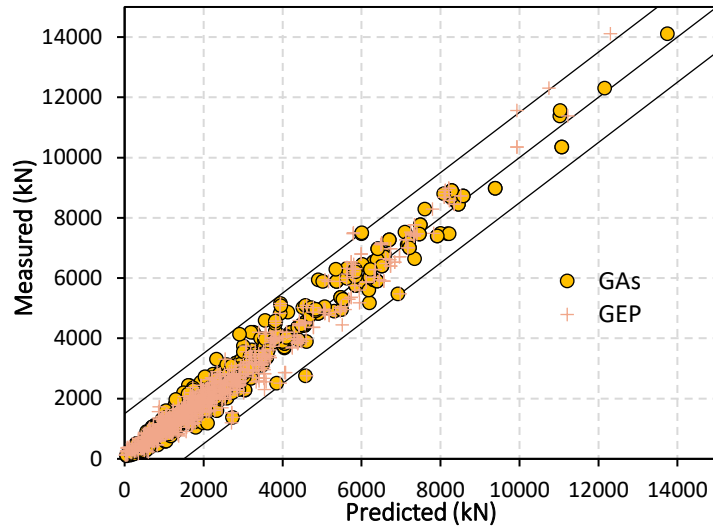
Naser M.Z., Thai S., Thai, T. (2021). “Evaluating Structural Response of Concrete-Filled Steel Tubular Columns through Machine Learning.” *Journal of Building Engineering*. <https://doi.org/10.1016/j.jobe.2020.101888>.



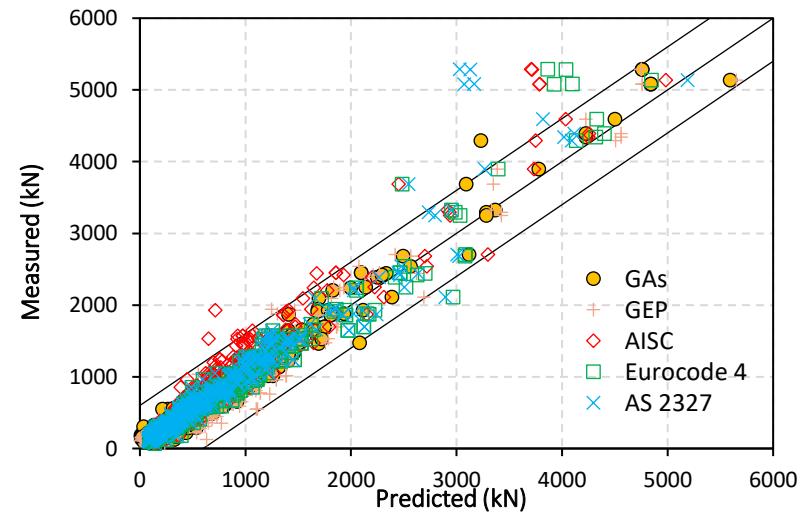
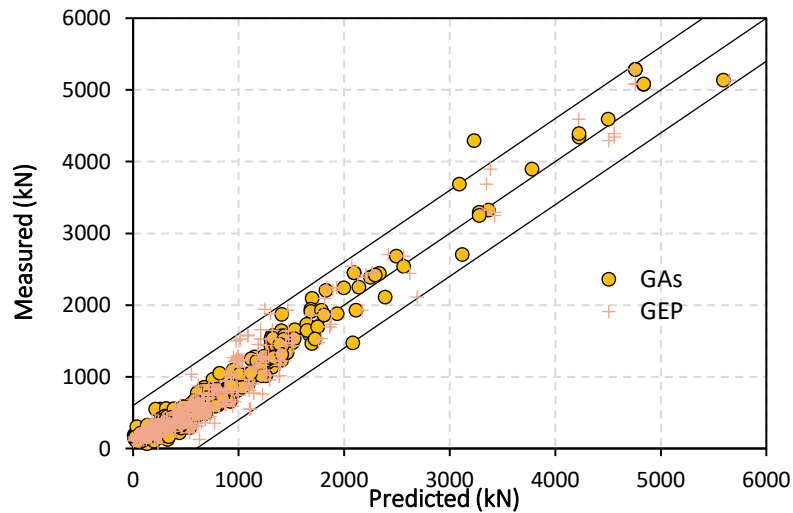
(a) Circular CFST columns (under concentric loading)

Please cite this paper as:

Naser M.Z., Thai S., Thai, T. (2021). "Evaluating Structural Response of Concrete-Filled Steel Tubular Columns through Machine Learning." *Journal of Building Engineering*. <https://doi.org/10.1016/j.jobe.2020.101888>.

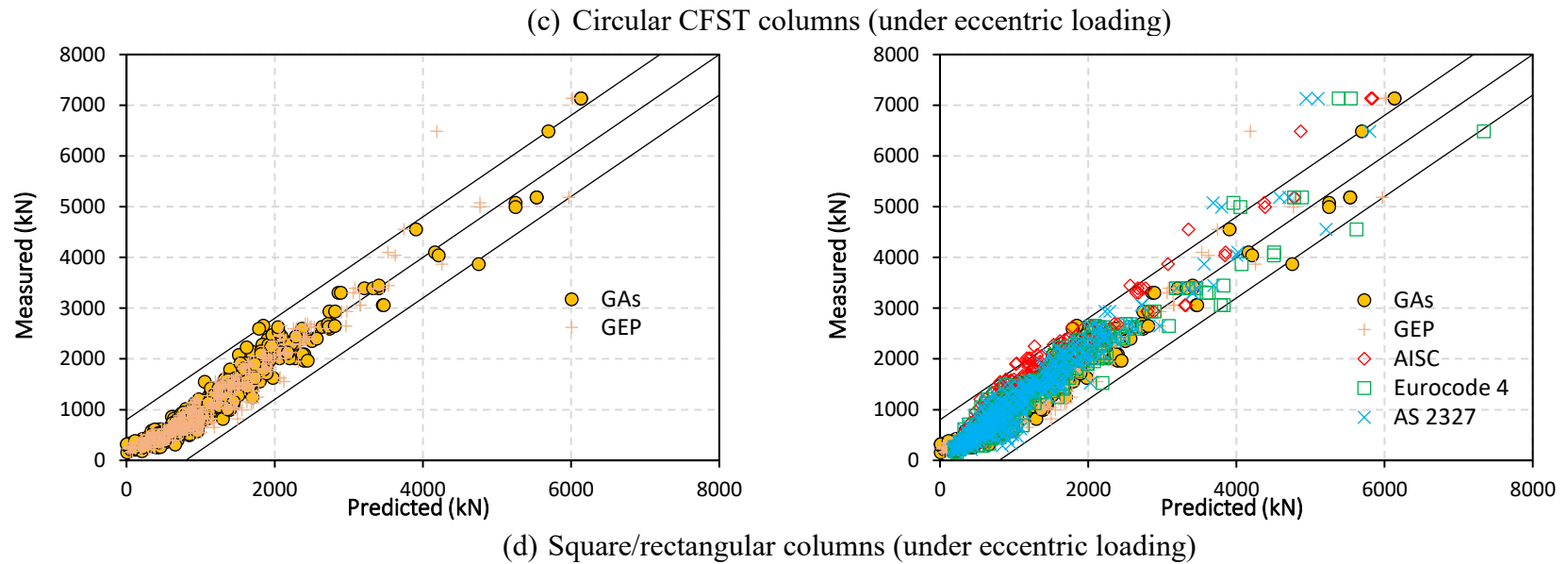


(b) Square/rectangular columns (under concentric loading)



Please cite this paper as:

Naser M.Z., Thai S., Thai, T. (2021). "Evaluating Structural Response of Concrete-Filled Steel Tubular Columns through Machine Learning." *Journal of Building Engineering*. <https://doi.org/10.1016/j.jobbe.2020.101888>.



**Figure 2** Predictions from newly derived expressions against those obtained from codal provisions

Please cite this paper as:

Naser M.Z., Thai S., Thai, T. (2021). "Evaluating Structural Response of Concrete-Filled Steel Tubular Columns through Machine Learning." *Journal of Building Engineering*. <https://doi.org/10.1016/j.jobbe.2020.101888>.

Other key statistics obtained from newly derived expressions can be found in Table 4. This table lays out statistics from predictions obtained using the NIML-derived expressions as well as codal provisions. This table shows how GA and GEP managed to achieve similar or at least very comparable performance to all other approaches in terms of average difference between measured and predicted strengths for both circular and square/rectangular columns. Predictions from GA and GEP also scored third and second in terms of predictions within 5% of that observed in the examined tests. Comparing error metrics (MAE: Mean absolute error and RMSE: Root-mean-square deviation) also shows well performance of GA and GEP models. Overall, NIML-based expressions seem to properly capture the behavior of concentrically and eccentrically loaded CFST columns with a better overall performance of GA as opposed to GEP. It is worth repeating that the predictions from NIML-based expressions come from a step substitution into the formula presented in Table 4 instead of carrying out a lengthy and multi-stage procedure.

**Table 4** Key statistics from predicted results

<i>Section</i>	<i>No. of specimens</i>	<i>Difference: measured/predicted</i>	<i>GA</i>	<i>GEP</i>	<i>AISC 360</i>	<i>Eurocode 4</i>	<i>AS 2327</i>
Circular (under concentric loading)	1245	Mean	0.98	0.97	1.27	1.09	1.09
		CoV	0.16	0.20	0.44	0.15	0.16
		MAE	232.4	282.1	493.79	193.66	202.63
			4	6			
		RMSE	384.7	464.7	1032.68	347.54	380.27
			3	0			
		No. of predictions within 5% of the true value	335	268	121	382	398
		No. of predictions > 5% of the true value	416	486	1101	698	680
		No. of predictions < 5% of the true value	509	506	38	180	182
		<i>k</i>	0.991	0.988	—	—	—
		<i>R<sub>m</sub></i>	0.897	0.875	—	—	—
Square/Rectangular (under concentric loading)	979	Mean	1.02	1.06	1.18	1.06	1.07
		CoV	0.13	0.15	0.22	0.20	0.18
		MAE	202.4	237.5	330.08	270.81	255.10
				1			
		RMSE	295.3	340.2	481.47	442.44	403.15
				4			
		No. of predictions within 5% of the true value	296	262	169	294	299



Please cite this paper as:

Naser M.Z., Thai S., Thai, T. (2021). "Evaluating Structural Response of Concrete-Filled Steel Tubular Columns through Machine Learning." *Journal of Building Engineering*. <https://doi.org/10.1016/j.jobbe.2020.101888>.

Circular (under eccentric loading)	485	No. of predictions > 5% of the true value	353	368	722	451	475
		No. of predictions < 5% of the true value	330	349	88	234	205
		$k$	1.007	1.004	—	—	—
		$R_m$	0.829	0.800	—	—	—
		Mean	1.20	1.09	1.22	1.08	1.16
		CoV	0.17	0.23	0.33	0.23	0.31
		MAE	91.56	129.3 9	152.58	92.39	113.22
		RMSE	126.8 6	176.6 2	252.03	173.37	239.40
		No. of predictions within 5% of the true value	110	73	86	125	77
		No. of predictions > 5% of the true value	176	216	327	255	357
		No. of predictions < 5% of the true value	199	196	72	105	51
		$k$	1.009	1.006	—	—	—
		$R_m$	0.826	0.746	—	—	—
		Mean	1.26	0.96	1.22	1.05	1.10
		CoV	0.18	0.20	0.30	0.21	0.24
		MAE	168.4 2	167.5 6	264.82	163.41	188.05
		RMSE	218.7 6	250.5 9	365.76	257.96	292.91
		No. of predictions within 5% of the true value	75	90	45	105	87
		No. of predictions > 5% of the true value	185	144	289	187	234
		No. of predictions < 5% of the true value	134	160	60	102	73
		$k$	1.005	1.006	—	—	—
Square/Rectangular (under eccentric loading)	394						

Please cite this paper as:

Naser M.Z., Thai S., Thai, T. (2021). “Evaluating Structural Response of Concrete-Filled Steel Tubular Columns through Machine Learning.” *Journal of Building Engineering*. <https://doi.org/10.1016/j.jobbe.2020.101888>.

$R_m$	0.752	0.708	—	—	—
-------	-------	-------	---	---	---

Coefficient of variation,  $CoV = \frac{RMSE}{mean}$  (Botchkarev, 2019); Mean absolute error,  $MAE = \frac{\sum_{i=1}^n |measured_i - predicted_i|}{\text{number of samples}}$ ; Root-mean-square,  $RMSE = \sqrt{\frac{\sum_{i=1}^n (measured_i^2 - predicted_i^2)}{\text{number of samples}}}$

In lieu of the above metrics and statistical comparisons, a number of researchers proposed other measures to critique the performance of a given model. For example, Smith (1986) suggested that when a model yields a correlation ( $R$ ) > 0.8, a strong correlation between the predictions and actual measurements exists. Table 3 shows that all NIML-based expressions have a correlation exceeding 96%, and thus satisfying this criterion. Another criterion was also proposed by Frank and Todeschini (1994) who recommended data scientists to maintain a ratio of between 3 and 5 between the number of observations and input parameters. In this work, these ratios were 252, 163.7, 26.2, and 43.7 for concentrically loaded circular and rectangular CFST columns as well as eccentrically loaded circular and rectangular CFST columns, respectively. It is clear that this criterion is also satisfied.

In lieu of the above criteria, another criterion was proposed by Golbraikh and Tropsha (2003). This criterion emphasizes the need for external verification by suggesting that at least one slope of regression lines ( $k$  or  $k'$ ) between the regressions of actual ( $A_i$ ) against predicted output ( $P_i$ ) or  $P_i$  against  $A_i$  through the origin, i.e.  $A_i = k \times P_i$  and  $t_i = k' A_i$ , respectively, passing through the origin needs to be close to 1.0 or at least within the range of 0.85 and 1.15. A look into Figure 2 and Table 4 shows that all NIML-models had a slope in the vicinity of 0.98-1.01, and thus satisfying the Golbraikh and Tropsha's criterion.

For the criterion proposed by Roy and Roy (2008) who developed a confirm indicator ( $R_m$ ) to assess the external predictability of predictive models, a model is said to have a good prediction capability when its indicator  $R_m > 0.5$ . The indicator ( $R_m$ ) is calculated as:

$$R_m = R^2 \times (1 - \sqrt{|R^2 - Ro^2|}) \quad (1)$$

where

$$Ro^2 = 1 - \frac{\sum_{i=1}^n (predicted_i - updated_i^o)^2}{\sum_{i=1}^n (predicted_i - \text{mean of predictions})^2}, \quad updated_i^o = k \times predicted_i$$

Table 4 shows that this criterion is also valid for all NIML-based models. The above discussion shows how the derived models are properly validated. It can be inferred that the use of the developed expressions and tools can provide designers with a quick and one-stepped assessment of the capacity of CFST columns. It should be noted that the proposed expressions are valid within the limit ranges described in Table 1. Any extensions to these expressions beyond these ranges may need to be applied with caution.

A note worthy of mentioning is that the number of predictions that exceeded 5% the value recorded in compiled tests shows how codal provisions seem to generally overestimate axial capacity of

Please cite this paper as:

Naser M.Z., Thai S., Thai, T. (2021). "Evaluating Structural Response of Concrete-Filled Steel Tubular Columns through Machine Learning." *Journal of Building Engineering*. <https://doi.org/10.1016/j.jobbe.2020.101888>.

CFST columns by a large margin. However, the NIML-based expressions seem to avoid that. Another note is that NIML attempt to optimize their prediction capabilities as to minimize over-prediction of capacities, and thus higher number of specimens of < 5% of the true value as compared to codal provisions. This shows the need for fine-tuning these algorithms in order to better optimize their predictive capability without compromising their accuracy. This will be further discussed in Section 5 and will be pursued in future works.

To further illustrate the application of the NIML-derived expressions, one example is carried out herein. This example covers a short circular column loaded under concentric loading. In this example, the axial capacity of the CFST column is calculated. This column is identical to that tested by Zhong and Zhang (1999) (column no. 3). This column has the following geometric and material properties.

Effective length ( $L_e$ ) = 310 mm

Tube thickness ( $t$ ) = 2.44 mm

Tube diameter ( $D$ ) = 104.2 mm

Yield strength of steel ( $f_y$ ) = 264.9 MPa

Compressive strength ( $f_c$ ) of concrete = 21.9 MPa

$$\text{The axial capacity } (N) = \text{abs}(0.00439Dt f_y + 0.000727tD^2 + 0.000727f_c D^2 - 1.38 \times 10^{-5}DL_e f_c - 3.71 \times 10^{-7}DtL_e f_y) \quad (2)$$

$$\begin{aligned} \text{The axial capacity } (N) &= \text{abs}(0.00439 \times 104.2 \times 2.44 \times 264.9 + 0.000727 \times 21.9 \times 104.2^2 + 0.000727 \times 21.9 \times 104.2^2 \\ &- 1.38 \times 10^{-5} \times 104.2 \times 310 \times 21.9 - 3.71 \times 10^{-7} \times 104.2 \times 2.44 \times 310 \times 264.9) = \\ &470.11 \text{ kN (within 3\% of the measured value at 450.0 kN)} \end{aligned} \quad (3)$$

The axial capacity of this column using other expressions is listed in Table 5. The reader can be deduced that GA gives better prediction of the column capacity.

**Table 5** Comparison between predictions from design expressions.

Expression	Predicted value (kN)	Within
GA	470.11	4%
GEP	399.00	11%
AISC 360	365.88	19%
Eurocode 4	481.16	7%
AS 2327	485.58	8%

## CHALLENGES AND FUTURE RESEARCH DIRECTIONS

In order for machine learning tools to comprehend a given phenomenon, these tools heavily rely on the availability of actual observations (preferably obtained from experimental tests). With 3,103

Please cite this paper as:

Naser M.Z., Thai S., Thai, T. (2021). "Evaluating Structural Response of Concrete-Filled Steel Tubular Columns through Machine Learning." *Journal of Building Engineering*. <https://doi.org/10.1016/j.jobbe.2020.101888>.

specimens used herein, there is still a lack on few fronts: (1) do not contain tests on specimens of exact (replicated) features, (2) presents outcome of tests carried out under similar (but not exact) conditions (e.g. with slightly different loading rate, testing equipment), (3) may lack aspects with regard to presentations on a complete range of geometric and material properties (including age, pre-existing loading and environmental exposure), support conditions, among others. While the derived expressions seem to naturally overcome the bulk of these challenges, the reader is still reminded with the aforementioned observations.

One way to scientifically overcome such challenges is to compile a much larger database than that presented herein. This can be undertaken via: (1) specifically designing future experiments, (2) collaborative efforts between researchers, (3) launching a repository at which researchers can freely access and update on regular basis, and (4) developing analytical and numerical models that can accurately capture the structural response of CFST columns. The results of such numerical studies, if designed properly, could present an attractive and economical solution over those of experimental nature, and hence may facilitate development of improved NIML-based design expressions.

In fact, the intention of this work is not to develop a set of AI-based expressions that are continually updated by including future tests to be carried out after the publication of this work. We suggest updating the proposed expressions in a 5 years cycle which is similar to that adopted in design codes.

Finally, it goes without saying that an exclusive feature of machine learning algorithms is that they are dynamic in nature and can evolve/improve once new data points are collected. Hence, the proposed expressions/codes are expected to undergo a series of improvements and reliability/calibrations before being officially inducted into practical applications. The proposed, as well as future expressions are to account for size effects, instability, uncertainty and composite action between steel tube and concrete filling. The use of such intelligent techniques will be carried out in parallel to traditional methods, as opposed to completely overthrowing testing and simulations.

## CONCLUSIONS

The outcome of this analysis shows the merit of utilizing intelligent agents (GA and GEP) in analyzing complex structural engineering phenomena that involves a lengthy and tedious procedure. This study paves the way for future works and encourages our community to leverage modern technologies towards realizing efficient and up-to-date design solutions. Based on the information presented in this paper, the following conclusions can be drawn.

- AI can overcome many of the limitations associated with those adopted in international design codes which do not account for specific material strengths or slenderness ratio. Integrating GA and GEP can provide structural engineers with novel, modern and optimal solutions.
- While the extension of design provisions examined in American, European, and Australian codes beyond their material and slenderness limits may not lead to consistent and conservative predictions, the newly proposed expressions seem overcome this trend.

Please cite this paper as:

Naser M.Z., Thai S., Thai, T. (2021). "Evaluating Structural Response of Concrete-Filled Steel Tubular Columns through Machine Learning." *Journal of Building Engineering*. <https://doi.org/10.1016/j.jobbe.2020.101888>.

- The performance of AI-derived expressions can be further improved with the availability of new data points as well as algorithms. It is expected that such expressions will undergo a continuous upgrade process every 5-10 years in a process similar to updating codal provisions.

## REFERENCES

- AISC, 2016. Specification for Structural Steel Buildings (ANSI/AISC 360-16) | American Institute of Steel Construction, ANSI/AISC 360-16.
- Alavi, A.H., Gandomi, A.H., Sahab, M.G., Gandomi, M., 2010. Multi expression programming: A new approach to formulation of soil classification. *Eng. Comput.* 26, 111–118. <https://doi.org/10.1007/s00366-009-0140-7>
- AS/NZS2327, 2017. AS/NZS2327 - Composite Steel-Concrete Construction for Buildings. Standard Australia/Standard New Zealand.
- Bailey, C., 2002. Holistic behaviour of concrete buildings in fire. *Proc. Inst. Civ. Eng. - Struct. Build.* 152, 199–212. <https://doi.org/10.1680/stbu.2002.152.3.199>
- Botchkarev, A., 2019. A new typology design of performance metrics to measure errors in machine learning regression algorithms. *Interdiscip. J. Information, Knowledge, Manag.* 14, 045–076. <https://doi.org/10.28945/4184>
- Cheu, R.L., Wang, Y., Fwa, T.F., 2004. Genetic algorithm-simulation methodology for pavement maintenance scheduling. *Comput. Civ. Infrastruct. Eng.* 19, 446–455. <https://doi.org/10.1111/j.1467-8667.2004.00369.x>
- Denavit, M., 2019. Steel-Concrete Composite Column Database [WWW Document]. URL <http://mark.denavit.me/Composite-Column-Database/#references> (accessed 6.18.19).
- El Ansary, A.M., El Damatty, A.A., Nassef, A.O., 2010. A coupled finite element genetic algorithm technique for optimum design of steel conical tanks. *Thin-Walled Struct.* <https://doi.org/10.1016/j.tws.2009.10.004>
- EN 1994-1-1, Eurocode 4, EN 1994-1-1, 2004. Eurocode 4: design of composite steel and concrete structures - Part 1.1: General rules and rules for buildings, European Committee for Standardization. <https://doi.org/10.1680/dgte4.31517>
- Ferreira, C., 2001. Gene Expression Programming: a New Adaptive Algorithm for Solving Problems. Ferreira, C. (2001). *Gene Expr. Program. a New Adapt. Algorithm Solving Probl. Complex Syst.* 13.
- Fike, R., Kodur, V., 2011. Enhancing the fire resistance of composite floor assemblies through the use of steel fiber reinforced concrete. *Eng. Struct.* 33, 2870–2878. <https://doi.org/10.1016/j.engstruct.2011.06.011>
- Frank, I., Todeschini, R., 1994. The data analysis handbook.
- GB50936, 2014. Technical code for concrete filled steel tubular structures. China Building Industry Press.
- Gharsellaoui, S., Mansouri, M., Trabelsi, M., Refaat, S.S., Messaoud, H., 2020. Fault diagnosis of heating systems using multivariate feature extraction based machine learning classifiers. *J. Build. Eng.* <https://doi.org/10.1016/j.jobbe.2020.101221>
- Golbraikh, A., Shen, M., Xiao, Z., Xiao, Y. De, Lee, K.H., Tropsha, A., 2003. Rational selection of training and test sets for the development of validated QSAR models. *J. Comput. Aided. Mol. Des.* 17, 241–253. <https://doi.org/10.1023/A:1025386326946>

Please cite this paper as:

Naser M.Z., Thai S., Thai, T. (2021). "Evaluating Structural Response of Concrete-Filled Steel Tubular Columns through Machine Learning." *Journal of Building Engineering*. <https://doi.org/10.1016/j.jobe.2020.101888>.

- Goldberg, D.E., Holland, J.H., 1988. Genetic Algorithms and Machine Learning. Mach. Learn. <https://doi.org/10.1023/A:1022602019183>
- Goode, G., 2008. Composite columns - 1819 tests on concrete-filled steel tube columns compared with Eurocode 4 - The Institution of Structural Engineers. Inst. Struct. Eng.
- Hajjar, J.F., 2002. Composite steel and concrete structural systems for seismic engineering. J. Constr. Steel Res. 58, 703–723. [https://doi.org/10.1016/S0143-974X\(01\)00093-1](https://doi.org/10.1016/S0143-974X(01)00093-1)
- Hejazi, F., Toloue, I., Jaafar, M.S., Noorzaei, J., 2013. Optimization of earthquake energy dissipation system by genetic algorithm. Comput. Civ. Infrastruct. Eng. 28, 796–810. <https://doi.org/10.1111/mice.12047>
- Huang, H., Burton, H. V., 2019. Classification of in-plane failure modes for reinforced concrete frames with infills using machine learning. J. Build. Eng. <https://doi.org/10.1016/j.jobe.2019.100767>
- Kaveh, A., Kabir, M.Z., Bohlool, M., 2019. Optimum design of three-dimensional steel frames with prismatic and non-prismatic elements. Eng. Comput. <https://doi.org/10.1007/s00366-019-00746-9>
- Kilpatrick, A.E., Rangan, B.V., 1999. Tests on High-Strength Concrete-Filled Steel Tubular Columns. ACI Struct. J. 96, 268–274. <https://doi.org/10.14359/618>
- Knowles, R., Park, R., 1969. Strength of Concrete Filled Steel Tubular Columns. J. Struct. Div. 95, 2565–2588.
- Kodur, V.K.R., 1999. Performance-based fire resistance design of concrete-filled steel columns. J. Constr. Steel Res. 51, 21–36. [https://doi.org/10.1016/S0143-974X\(99\)00003-6](https://doi.org/10.1016/S0143-974X(99)00003-6)
- Koza, J.R., 1992. A genetic approach to finding a controller to back up a tractor-trailer truck, in: Proceedings of the 1992 American Control Conference.
- Leon, R.T., Perea, T., Hajjar, J., Denavit, M., 2011. Concrete-filled tubes columns and beam-columns, in: A Database for the AISC 2005 and 2010 Specifications, Festschrift Gerhard Hanswille, Germany (2011), Pp. 203-212. pp. 203–212.
- Liew, J.Y.R., Xiong, M., Xiong, D., 2016. Design of Concrete Filled Tubular Beam-columns with High Strength Steel and Concrete. Structures 8, 213–226. <https://doi.org/10.1016/j.istruc.2016.05.005>
- Liu, Z., Goel, S.C., 2008. Cyclic Load Behavior of Concrete-Filled Tubular Braces. J. Struct. Eng. 114, 1488–1506. [https://doi.org/10.1061/\(asce\)0733-9445\(1988\)114:7\(1488\)](https://doi.org/10.1061/(asce)0733-9445(1988)114:7(1488))
- Naser, M.Z., 2019a. Properties and material models for modern construction materials at elevated temperatures. Comput. Mater. Sci. 160, 16–29. <https://doi.org/10.1016/J.COMMATSCI.2018.12.055>
- Naser, M. Z., 2019a. Can past failures help identify vulnerable bridges to extreme events? A biomimetical machine learning approach. Eng. Comput. <https://doi.org/10.1007/s00366-019-00874-2>
- Naser, M. Z., 2019b. Properties and material models for common construction materials at elevated temperatures. Constr. Build. Mater. 10, 192–206. <https://doi.org/10.1016/j.conbuildmat.2019.04.182>
- Naser, M.Z., 2019b. AI-based cognitive framework for evaluating response of concrete structures in extreme conditions. Eng. Appl. Artif. Intell. 81, 437–449. <https://doi.org/10.1016/J.ENGAPPAI.2019.03.004>



Please cite this paper as:

Naser M.Z., Thai S., Thai, T. (2021). "Evaluating Structural Response of Concrete-Filled Steel Tubular Columns through Machine Learning." *Journal of Building Engineering*. <https://doi.org/10.1016/j.jobe.2020.101888>.





- Naser, M.Z., 2019c. Heuristic machine cognition to predict fire-induced spalling and fire resistance of concrete structures. *Autom. Constr.* 106, 102916. <https://doi.org/10.1016/J.AUTCON.2019.102916>
- Rafiei, M.H., Khushefati, W.H., Demirboga, R., Adeli, H., 2017a. Novel approach for concrete mixture design using neural dynamics model and virtual lab concept. *ACI Mater. J.* 114. <https://doi.org/10.14359/51689485>
- Rafiei, M.H., Khushefati, W.H., Demirboga, R., Adeli, H., 2017b. Supervised Deep Restricted Boltzmann Machine for Estimation of Concrete. *ACI Mater. J.* 114. <https://doi.org/10.14359/51689560>
- Roy, P.P., Roy, K., 2008. On some aspects of variable selection for partial least squares regression models. *QSAR Comb. Sci.* 27, 302–313. <https://doi.org/10.1002/qsar.200710043>
- Sakino, K., Nakahara, H., Morino, S., Nishiyama, I., 2004. Behavior of Centrally Loaded Concrete-Filled Steel-Tube Short Columns. *J. Struct. Eng.* 130, 180–188. [https://doi.org/10.1061/\(ASCE\)0733-9445\(2004\)130:2\(180\)](https://doi.org/10.1061/(ASCE)0733-9445(2004)130:2(180))
- Sarir, P., Chen, J., Asteris, P.G., Armaghani, D.J., Tahir, M.M., 2019. Developing GEP tree-based, neuro-swarm, and whale optimization models for evaluation of bearing capacity of concrete-filled steel tube columns. *Eng. Comput.* <https://doi.org/10.1007/s00366-019-00808-y>
- Smith, G., 1986. Probability and statistics in civil engineering. Collins, London.
- Tao, Z., Uy, B., Han, L.H., He, S.H., 2008. Design of concrete-filled steel tubular members according to the Australian Standard AS 5100 model and calibration. *Aust. J. Struct. Eng.* 8, 197–214. <https://doi.org/10.1080/13287982.2008.11464998>
- Thai, S., Thai, H., Uy, B., Ngo, T., Naser, M., 2019. Test database on concrete-filled steel tubular columns [WWW Document]. <https://doi.org/10.17632/3XKNB3SDB5.1>
- Thai, Son, Thai, H.T., Uy, B., Ngo, T., 2019. Concrete-filled steel tubular columns: Test database, design and calibration. *J. Constr. Steel Res.* <https://doi.org/10.1016/j.jcsr.2019.02.024>
- Yuan, F., Chen, M., Huang, H., Xie, L., Wang, C., 2018. Circular concrete filled steel tubular (CFST) columns under cyclic load and acid rain attack: Test simulation. *Thin-Walled Struct.* 122, 90–101. <https://doi.org/10.1016/j.tws.2017.10.005>
- Yucel, M., Bekdaş, G., Nigdeli, S.M., Sevgen, S., 2019. Estimation of optimum tuned mass damper parameters via machine learning. *J. Build. Eng.* <https://doi.org/10.1016/j.jobe.2019.100847>
- Zhong, S. tong, Zhang, S., 1999. Application and development of concrete-filled steel tubes (CFST) in high rise buildings. *Adv. Struct. Eng.* 2, 149–159. <https://doi.org/10.1177/136943329900200207>

Please cite this paper as:

Naser M.Z., Thai S., Thai, T. (2021). “Evaluating Structural Response of Concrete-Filled Steel Tubular Columns through Machine Learning.” *Journal of Building Engineering*. <https://doi.org/10.1016/j.jobbe.2020.101888>.

## APPENDIX

*Table A.1 Matlab Codes Obtained from GEP Analysis \**

Circular (concentric loading)
 CC.txt
Square/Rectangular (concentric loading)
 RC.txt
Circular (eccentric loading)
 CE.txt
Square/Rectangular (eccentric loading)
 RE.txt

\*Only positive outcomes should be considered. We recommend the use GA as oppose in GEP in these scenarios.