MGP–CC: a hybrid multigene GP–Cuckoo search method for hot rolling manufacture process modelling

Hossam Faris¹, Alaa F. Shetab¹ and Ertan Öznergiz³

¹King Abdullah II School for Information Technology, The University of Jordan, Amman, Jordan; ²Computers and Systems Department, Electronics Research Institute (ERI), Giza, Egypt; ³Marine Engineering Operations Department, Faculty of Naval Architecture and Maritime, Yildiz Technical University, Istanbul, Turkey

ABSTRACT

Maintaining high level of quality in hot rolling manufacturing processes is very challenging problem to keep competitiveness in the iron and steel industrial market. Monitoring the quality of the output product helps enhancing the product outcomes, increase the company profit and improve the economic growth of the country. In this paper, we propose a new hybrid approach based on multigene genetic programming (MGP) and Cuckoo search (CS) algorithms for developing three rigorous models for roll force, torque and slab temperature in the hot rolling industrial process at the Ereğli Iron and Steel Factory in Turkey. MGP is a robust variation of the standard genetic programming (GP) algorithm while CS is a new nature-inspired metaheuristic search algorithm. The performance of the developed models is evaluated and compared with those obtained for the standard MGP and GP approaches.

1. Introduction

The problem of modelling the hot rolling industrial process is complex and challenging due to the highly dynamic and nonlinear relation between the involved variables in the process such as friction coefficient, yield stress, disturbances, etc. Moreover, the data sets collected to predict the performance of the process is often noisy and multi-variate (Moussaoui, Selaimia, & Abbassi, 2006). This makes getting accurate measurement of the rolling force/torque very difficult task. However, accurate and reliable models for the rolling force, torque and temperature are vital in designing pass online schedules and maintaining high quality of the production.

Steel market industry has witnesses a high competition every year. Thus, maintaining exceptional level of production quality is essential to reserve a high market share (Cowling, 2003; Liberatore, 1989). In hot rolling strip production in particular, the demand for higher quality products motivates the need for quality control techniques such as reliable and accurate prediction models of the rolling force/torque and temperature (Cowling, 2003; Feng, Liu, Luo, Tang, & Liu, 2008; Kirihata, Maccagno, & Jonas, 1998; Kwak, Kim, Park, Lee, & Hwang, 2000; Schroder, 2003). Typical mathematical models have some limitation since some factors are difficult to be integrated such as the reduction rate and the chemical composition. A typical solution to come around this problem in some rolling processes is to adopt fragmented look-up tables. However, this solution has some disadvantages such as the large size of the generated look-up tables and the lack of effective interpolation capability.

In literature, many authors proposed quality monitoring models for the hot rolling process based on different machine learning techniques such as artificial neural networks (ANN) and fuzzy logic in order to overcome the problems mentioned earlier (Barrios, Torres-Alvarado, & Cavazos, 2012; Bouhouche, Yahi, Hocine, & Bast, 2008; Faris, Sheta, & Öznergiz, 2013; Gorni, 1997; Öznergiz, Özsöy, Delice, & Kural, 2009; Rong, Dan, & Yi, 2005; Sheta, Braik, Öznergiz, Ayesh, & Masud, 2013; Tiensuu, Juutinen, & Röning, 2011). However, the disadvantage of most of these techniques is the lack of interpretability and considered as black box models. For that reasons authors in Faris et al. (2013) and Sheta and Faris (2014) investigated the application of Symbolic Regression approach using genetic programming (GP) for predicting features of a hot rolling process including roll force, roll torque, and slab temperature.

One of the difficulties and weaknesses in GP is finding the numeric constants that form the GP model along
with the set of arithmetic functions (Koza, 1992). In standard GP, numeric constants are generated randomly. Therefore, various studies in the literature proposed different ways to solve this problem. Some of the effective approaches proposed by researchers is the hybridization between GP and metaheuristic algorithms. Two main types of these approaches could be distinguished: the co-evolution approaches and the relay approaches. One example of the co-evolution approaches can be found in Mukherjee and Eppstein (2012) where authors used differential evolution to co-evolution of the coefficients in GP. On the other hand, relay approaches perform optimization of the constants in a separated task when GP finishes evolving the structure of the Symbolic Regression model. An example of this approach is proposed in Alonso, Montaa, and Borges (2009) where evolution strategies is applied to optimize GP constants.

In this paper, we propose a new hybrid relay approach based on combining multigene genetic programming (MGP) and Cuckoo search (CS) algorithm for developing more accurate symbolic regression models for a hot rolling industrial process. MGP and CS are executed separately in sequence. MGP is applied first for finding the structure of the symbolic regression models then CS is used for optimizing the numeric constants in these models. MGP is an evolutionary computation algorithm and a robust variation of the standard GP while CS is a new nature-inspired metaheuristic search algorithm. The objective of the proposed approach is to simulate the behaviour of different features of a hot rolling manufacturing process of the Ereğli Iron and Steel Factory in Turkey. Three models for the force, torque and slab temperature in the plate mill are accurately developed and compared to those obtained by MGP and the standard GP approaches.

The paper is organized as follow. Section 2 describes the hot rolling process in Ereğli Iron and Steel Factory. Section 4 introduces the MGP approach and its main processes. Section 5 gives a brief description of the CS algorithm and explains its application for tuning the generated MGP models. Section 6 describes the data collected for the purpose of the experiments in this study. In Section 7, we list the evaluation criteria used to evaluate the final developed models. Finally, Section 8 presents the experiments settings and discusses the obtained results.

2. Problem description

The hot-rolling mill plant in Ereğli Iron and Steel Factory is shown in Figure 1. The plant consists of two slab furnaces, pre-rolling mill, edger, reversible mill, seven strip rolling stands, a cooling system, a hot leveller, and a shearing system. The plant has also a data acquisition and a computer control system modified by General Electrics. The Ereğli Iron and Steel Factory in Turkey has two cold and two hot-rolling rolling mill plants. Cold rolling plants have a total capacity of 2.3 million tons per year and hot-rolling mill plants, 540.000 tons per year, corresponding to a total product capacity of 3 million tons per year.

Steel strips with a thickness of 15–16 mm can be produced in the rolling mill plant. In a normal production cycle, each slab passes 5 times in forward and backward directions in the reversible mill. In this plant, the dimensions of slabs are monitored continuously during every passes with an X-ray system, the temperature of slab with a pyrometer, roll force and torque with four load cells placed along the mill. But averages of these measured values for each pass are used for identification. The chemical composition of the low-carbon steels used in this study is given in Table 1.

3. Genetic programming

GP is an evolutionary algorithm-based methodology for automatically solving problems inspired by biological evolution (Koza, 1991, 1994). GP has many advantages when used to model nonlinear systems (Espejo, Ventura, & Herrera, 2010; Hussein, Sheta, & Abd-Elwahab, 2001; Kotanchek, Smits, & Kordon, 2003; Sheta et al., 2013). They include:

- **Flexibility**: is one of the main advantages of GP over many other heuristic approaches. That is because the general structure of the tree representation of GP individuals can fits a wide range of problems. Moreover, using the terminal set and the function set can do both, represent knowledge and perform computation.
- **Interpretability**: some modelling and prediction techniques are considered as black box methods since they generate models which are hard to understand such techniques traditionally include ANNs, SVMs. Unlike those techniques, tree models generated by GP are

| Carbon | Manganese | Sulphur | Phosphorus | Silicon |
|--------|-----------|---------|------------|---------|
| 0.12%  | 0.25%     | 0.2%    | 0.025%     | 0.05%   |
more simple, easier to evaluate and their complexity is controllable by setting the maximum GP tree depth and length.

Therefore, GP has been applied successfully to a large number of manufacturing and industrial processes. Some of these applications can be found in Hussein et al. (2001), Hussain, Sheta, Kamel, Telbany, and Abdelwahab (2000), Faris and Sheta (2013), Faris et al. (2013), Sheta and Faris (2014) and Faris (2013). In particular for the case of hot rolling process, the results in Faris et al. (2013) and Sheta and Faris (2014) show that the developed GP mathematical models are very competitive taking into consideration their simplicity and interpretability.

4. Multigene genetic programming

Multigene GP is a robust variation of the standard GP (Searson, Leahy, & Willis, 2010). Multigene GP can be applied to evolving a population of trees (possible solutions) while minimizing some error criteria. For a system with \( x \) input of dimension \( n \times m \) to produce a model output \( y \) with dimension \( n \times 1 \), we could produce a tree structure which introduce the mathematical relationship:

\[ y = f(x), \]

where \( n \) is the number of observations taken and \( m \) is the number of input variables. In multigene GP, prediction of the output variable \( \hat{y} \) is resulted by a weighted output for each trees/genes in the multigene individual plus a bias term. Each tree is a function of zero or more of the \( N \) input variables \( x_1, \ldots, x_N \). The main processes of the Multigene GP algorithm flow chart is shown in Figure 3.

4.1. Initial population and representation

The evolutionary cycle of the algorithm starts by generating random individuals (possible solution or model) each of which consists of \( M \) randomly generated trees. A Multigene GP model can be represented in the form shown in Equation (2). A multigene regression model can be written as:

\[ \hat{y} = \alpha_0 + \sum_{i=1}^{M} \alpha_i \times \text{tree}_i, \]

where \( \alpha_0 \) represents the bias or offset term while \( \alpha_1, \ldots, \alpha_M \) are the gene weights and \( M \) is the number of genes (i.e. trees) which constitute the available individual. The weights (i.e. regression coefficients) are automatically determined by a least squares (LS) procedure for each multigene individual. In multigene symbolic regression each symbolic model is represented by number of GP trees weighted by linear combination. Each tree is considered as a ‘gene’ by itself. An example of multigene model is shown in Figure 2. The presented model can be introduced mathematically as given in Equation (3).

\[ \hat{y} = \alpha_0 + \alpha_1 \left[ \sin(X) + \frac{Y}{5} \right] + \alpha_2 [5 \times Y]. \]

Each symbolic model will have set of linear coefficients such as \( \alpha_0, \alpha_1, \ldots \). These parameters will be estimated ordinary LS techniques. This shows the powerful nature of multigene GP since it combines the classical linear regression theory with the ability to capture non-linear behaviour based the evolutionary process. It was reported that there are many advantages of multigene symbolic regression than standard GP approach for symbolic regression (Hinchliffe et al., 1996).

A multigene symbolic models are consisted of a function set and a terminal set. The function set is likely to contain of a number of simple arithmetic operators functions and/or and nonlinear mathematical functions such as the ones given in Equation (4).

\[ F = \{ +, -, \times, /, \sin, \cos, \exp, \log \}. \]

The function \( F \) is combined with the terminal set to help the algorithm to develop suitable tree structure
which represent a model for the problem. This multigene symbolic model has the advantages that it is linear in the parameters with respect to the coefficients $\alpha_0, \alpha_1$ and $\alpha_2$.

4.2. Fitness evaluation

The root mean square error (RMSE) is used as a fitness function in order to evaluate GP individuals. RMSE can be described by Equation (5).

$$\text{RMSE} = \sqrt{\frac{1}{n} \sum_{i=1}^{n} (y_i - \hat{y}_i)^2},$$

where $y$ and $\hat{y}$ are the actual and the estimated values based on proposed model and $n$ is the number of measurements used in the experiments, respectively.

4.3. Tournament selection

A selection mechanism consists of two components: a selection scheme and a selection pressure control strategy. This strategy is usually critical in designing a selection mechanism and has been widely studied by the evolutionary computation community (Legg, Hutter, & Kumar, 2004). Tournament selection is a beneficial and robust selection mechanism (Miller, Sutton, & Werbos, 1995). The selection pressure of tournament selection changed according to the tournament size. The more the number of individuals are selected the higher the resulting selection pressure.

4.4. Crossover

During the evolutionary process, genes are generated using crossover operator. In our case, we used two point crossover. Crossover makes exchange of genes between individuals. A two point high level crossover adopted in Searson et al. (2010) is used. The example in Equations (6)–(8) shows the operation.

- Individuals for crossover

$$G_1 G_2 G_3 G_4 G_5 (G_6 G_7 G_8).$$  \hfill (6)

- Selected genes for crossover

$$G_1 G_2 G_3 G_4 G_5 (G_6 G_7 G_8).$$  \hfill (7)

- Generated individual

$$G_1 G_7 G_8 G_4 G_5 (G_6 G_2 G_3).$$  \hfill (8)

4.5. Mutation

Mutation in multigene GP operates in a similar fashion to standard GP evolutionary process. A random gene is selected then a normal sub-tree mutation is performed. The new generated gene replaces the parent gene in the new generation.

4.6. Coefficients estimation

In this process, the linear coefficients $\alpha_0, \alpha_1, \ldots, \alpha_M$ of each individual in the population are estimated using ordinary LS techniques based on the training data. In order to save the computation time. This operator is applied on a small fraction of each generation. By default it is set to 5%.

5. Optimizing multigene GP models by CS

CS is one of the recent and powerful metaheuristic optimization algorithms (Yang & Deb, 2009, 2010). CS is inspired by the natural brood of Cuckoo birds and the Lévy flight behaviour of some birds species and fruit flies. The CS algorithm is designed mainly on the following three assumptions (Yang & Deb, 2009):

1. A Cuckoo bird selects randomly a nest and put one egg (solution) in it. Therefore, every nest forms one candidate solution.
2. Some nests that has the best quality solutions will be passed to the next generation.
3. The number of available host nests is fixed, and the ratio that a host discovers a cuckoo egg can be approximated by a fraction $p_a$ of the $n$ nests being neglected and replaced by new random nests (solutions) at their positions.

New solutions $x_{i}^{t+1}$ are generated using a Lévy flights as shown in Equation (9). Lévy flights are performed by random walks with step-lengths are drawn from a Lévy distribution, a heavy-tailed probability distribution, which has an infinite variance with an infinite mean, Equation (10).

$$x_{i}^{t+1} = x_{i}^{t} + \alpha \odot \text{Lévy}(\lambda),$$

$$\text{Lévy} \sim \lambda = t^{-\lambda}.$$  \hfill (9)

In Equation (9), $\alpha > 0$ is the step size which should be related to the scales of the problem of interest. In most cases, $\alpha$ is supposed equals 1. The product $\odot$ denotes entry-wise multiplications.

In this paper, the fitness function used in CS algorithm is the same as MGP which is the RMSE criteria. While nests are encoded as sequence of the coefficients $\alpha$s that
connects the trees in the final MGP developed models shown in Figure 4. The goal is to find the best coefficients that minimize the RMSE of the MGP model. CS algorithm enjoys some advantages that make it very competitive to other metaheuristic algorithm. Such advantages include:

- CS has only one parameters to be tuned beside the population size which is the discovery rate $P_d$ (Yang & Deb, 2009).
- In contrast to other metaheuristic approaches that use uniform distributions or Gaussian to generate new explorative moves in the search space, CS performs random walk via Lévy flight which is more efficient in exploring the search space if the search space is large enough (Yang & Deb, 2009).

The CS algorithm works as described by Algorithm 1 where $s$ is the step size and $P_d$ is the probability of discovering the egg. $E_i$ is derived form Lévy Flight distribution which is a random walk with step length depending on the current state plus a transition probability (Yang & Deb, 2009).

6. Hot rolling data set

The hot rolling collected data set contains 640 collected from 128 different slabs by a General Electric’s data acquisition system. The collected measurements were practically measured for three sub-processes: roll force $f$, roll torque $G$ and slab temperature $T$ of the rolling process. The thickness and width distribution of the data ranged from 31.68 to 168.6 mm and 948.76 to 1457.26 mm, respectively. In order to compare our obtained results with other reported results from the literature (Öznergiz et al., 2009; Sheta, Öznergiz, Abdelrahman, & Babuska, 2009), same data set partitioning for the training and testing was adopted, 500 for training and 140 for testing.

Each of force $f$, the torque $G$ and the slab temperature $T$ has six input variables. The inputs and output of the proposed model structures can be summarized as follows:

- Inputs to force $f$ and torque $G$ models: $u_1, u_2, u_3, u_4, u_5$ and $u_6$ inputs stand for: entry temperature ($T_i$), width ($W_s$), carbon equivalent ($C_e$), gauge ($h_i$), draft ($i$) and roll diameter ($R$), respectively. The output of each sub-process is stated as $M_{Force}, M_{Torque}$ and given in Equation (11).

$$M_{Force} = f(T_i, W_s, C_e, h_i, i, R),$$

$$M_{Torque} = f(T_i, W_s, C_e, h_i, i, R).$$

(11)

- Input to slab temperature $T$ model: $u_1, u_2, u_3, u_4, u_5$ and $u_6$ inputs stand for: Entry temperature ($T_i$), width ($W_s$), carbon equivalent ($C_e$), gauge ($h_i$), torque ($G_i$), power ($E_i$), respectively. The output of each sub-process is stated as $M_{Temp}$ and given in Equation (12).

$$M_{Temp} = f(T_i, W_s, C_e, h_i, G_i, E_i).$$

(12)

7. Performance evaluation

Other performance criteria are used to evaluate the quality of the developed Multigene GP models. The set of criteria used are given as follows:

1. Variance-accounted-for (VAF)

$$VAF = \left[ 1 - \frac{\text{var}(y - \hat{y})}{\text{var}(y)} \right] \times 100\%.$$  

(13)

2. Mean squares error (MSE)

$$\text{MSE} = \frac{1}{n} \sum_{i=1}^{n} (y - \hat{y})^2.$$  

(14)

3. Mean absolute error (MAE):

$$\text{MAE} = \frac{1}{n} \left( \sum_{i=1}^{n} |y_i - \hat{y}_i| \right),$$  

(15)

where $y$ and $\hat{y}$ are the actual and the estimated values based on proposed model and $n$ is the number of measurements used in the experiments, respectively.

8. Experimental results

In our experiments we use the GPTIPS Tool\textsuperscript{1} for developing Multigene GP models for rolling force, torque and slab temperature. The parameters of the application are tuned as listed in Table 2. The best results are obtained using seven genes in each individual. The function set used to combine the terminals in GP individuals is $F = \{+, -, /, *, \sqrt{\text{, sin, cos}}\}$.
Algorithm 1 Basic steps describing the CS

1: **procedure** CS
2: **Objective function** \( f(x), x = (x_1, x_2, \ldots, x_d)^T \)
3: **Generate initial population of** \( n \) **host nests** \( x_i (i = 1, 2, \ldots, n) \)
4: **while** ( \( t < \) Max Generation) or (stop criterion not met)
5: \( \) **Get a cuckoo randomly by Lévy flights** evaluate its fitness \( F_i \)
6: **Choose a nest among** \( n \) **(say,)** \( j \)
7: **If** \( (F_i > F_j) \)
8: **replace** \( j \) **by the new solution**
9: **end if**
10: \( A \) **fraction** \( (P_a) \) **of the worse nests are abandoned and new ones are built**
11: **Keep the best solutions/nests**
12: **Rank the solutions/nests and find the current best**
13: **Pass the current best solutions to the next generation**
14: **end while**
15: **end**

**Table 2.** Parameters settings for MGP experiments.

| Parameter                   | Setting     |
|-----------------------------|-------------|
| Population size             | 50          |
| Number of generation        | 100         |
| Selection mechanism         | Tournament (size = 3) |
| Max. tree depth             | 10          |
| Probability of crossover    | 85%         |
| Probability of mutation     | 10%         |
| Elite                       | 2%          |
| Max. no. of genes allowed   | 7           |

**Figure 5.** MGP convergence curves for force model.

**Figure 6.** MGP convergence curves for torque model.

**Figure 7.** MGP convergence curves for temperature model.

In Figures 5–7 we show the convergence of MGP toward the best obtained force torque and temperature models, respectively. Figures 8–10 show the actual and estimated force, torque and temperature values based on the developed MGP models.

In order to enhance the accuracy of the developed MGP models, CS algorithm is used to tune the parameters of these models. This approach will be referred to as MGP–CS. CS is applied for 500 cycles. The actual and estimated force, torque and temperature values based on the developed MGP–CS models are shown in Figures 11–13, respectively.

The developed MGP and MGP–CS models are evaluated using VAF, MSE and MAE; and compared with the previously obtained results for the standard GP approach in Faris et al. (2013) and Sheta and Faris (2014). Based on the comparison results shown in Tables 3–5 we can notice that MGP achieved better results than the standard GP approach in the force and torque cases while there
Figure 8. Actual and MGP estimated force in training and testing cases.

Figure 9. Actual and MGP estimated torque in training and testing cases.
Figure 10. Actual and MGP estimated temperature in training and testing cases.

Figure 11. Actual and MGP–CS estimated force in training and testing cases.
Figure 12. Actual and MGP–CS estimated torque in training and testing cases.

Figure 13. Actual and MGP–CS estimated temperature in training and testing cases.
was no improvement in the temperature case. Moreover, combining CS algorithm with MGP in the MGP–CS approach generated the best quality models for the roll force, torque and temperature since it increased the VAF measurement and decreased MAE and MSE for all testing cases.

In Table 4, there is an enhancement also on the roll torque by decreasing the MAE and MSE values for the testing case. However, in the case of slab temperature which is shown in Table 5 we can notice that the performance is almost the same compared to the standard GP approach. Finally, it can be concluded that the MGP–CS modelling approach has a promising potential in improving the quality of the roll force and torque prediction models in the hot rolling process. This can be explained by the advantage of Multigene GP that combines the classical linear regression theory with the ability to capture non-linear behaviour based the evolutionary process.

### 9. Conclusion

In this paper, MGP is combined with a CS algorithm to develop estimation models for a Hot-Rolling industrial process. Three sub models are proposed for the process; the rolling force, torque and slab temperature. The developed MGP–CS models are evaluated and compared with the standard GP models based on different evaluation criteria. Experimental data collected from the Ereğli Iron and Steel Factory in Turkey was used for the verification of the developed models. Compared to MGP and standard GP, MGP–CS models showed better performance results for rolling force, torque and temperature compared to those obtained for the other approaches. This advantage can support implementing higher quality models efficiently in iron and steel manufacturing plants.

### Note

1. GPTIPS Tool is a free, open source MATLAB toolbox for performing symbolic regression by GP.

### Acknowledgments

Authors would like to thank to Prof. Dr Can ÖZSOY and the Ereğli Iron and Steel Factory for providing the required technical assistant and the data for this developed research work.

### Disclosure statement

No potential conflict of interest was reported by the authors.

### References

Alonso, C. L., Montaa, J., & Borges, C. E. (2009). Evolution strategies for constants optimization in genetic programming. 21st International Conference on Tools with Artificial Intelligence, 2009, ICTAI’09, IEEE, pp. 703–707.

Barrios, J. A., Torres-Alvarado, M., & Cavazos, A. (2012, February). Neural, fuzzy and grey-box modelling for entry temperature prediction in a hot strip mill. Expert Systems With Applications, 39(3), 3374–3384.

Bouhouche, S., Yahi, M., Hocine, B., & Bast, J. (2008). Soft sensor-based artificial neural networks and fuzzy logic: Application to quality monitoring in hot rolling. In Proceedings of the 10th WSEAS international conference on automatic control, modelling & simulation (ACMOS’08) (pp. 149–154). Stevens Point, WI: World Scientific and Engineering Academy and Society (WSEAS).

Cowling, P. (2003). A flexible decision support system for steel hot rolling mill scheduling. Computers & Industrial Engineering, 45(2), 307–321.

Espejo, P. G., Ventura, S., & Herrera, F. (2010). A survey on the application of genetic programming to classification. IEEE Transactions on Systems, Man, and Cybernetics, Part C: Applications and Reviews, 40(2), 121–144.

Faris, H. (2013). A symbolic regression approach for modeling the temperature of metal cutting tool. International Journal of Control and Automation, 6(4), 431–440.

Faris, H., & Sheta, A. (2013). Identification of the Tennessee Eastman chemical process reactor using genetic programming. International Journal of Advanced Science and Technology, 50, 121–140.

Feng, X., Liu, Y., Luo, H., Tang, H., & Liu, C. (2008). Numerical simulation during expansion for hot rolling sheet strip. In Proceedings of the first international workshop on knowledge discovery...
Gorni, A. (1997). The application of neural networks in the modeling of plate rolling processes. *Journal of Materials-Engineering*, **49**(4), 1–9.

Hinchliffe, M., Hiden, H., McKay, B., Willis, M., Tham, M., & Barton, G. (1996). Modelling chemical process systems using a multigene genetic programming algorithm. In J. R. Koza (Ed.), *Late breaking papers at the genetic programming 1996 conference*, Stanford University, July 28–31 (pp. 56–65). Stanford University, CA: Stanford Bookstore.

Hussein, A., Sheta, A., & Abd-Elwahab, A. (2001). Modeling of a winding machine using neural networks. Proceeding of the WSEAS international conference on scientific computation and soft computing, pp. 528–533.

Koza, J. (1991). *Evolving a computer program to generate random numbers using the genetic programming paradigm*. Proceedings of the fourth international conference on genetic algorithms, La Jolla, CA: Morgan Kaufmann.

Koza, J. R. (1992). Genetic programming. Cambridge, MA: MIT Press.

Koza, J. R. (1994). *Genetic programming II*. Cambridge, MA: MIT Press.

W. J. Kwak, Y. H. Kim, H. D. Park, J. H. Lee, & S. M. Hwang (2000). FE-based on-line model for the prediction of roll force and roll power in hot strip rolling. ISIJ International-The Iron and Steel Institute of Japan, **38**(2), 187–195.

Legg, S., Hutter, M., & Kordon, A. (2004). Tournament versus fitness uniform selection. In *Proc. 2004 Congress on evolutionary computation* (CEC-2004) (pp. 2144–2151). Portland, OR: IEEE.

Liberatore, R. L. (1989). Hot rolling mill optimization model. *Computers & Industrial Engineering*, **17**(1–4), 130–135.

Miller, W., Sutton, R., & Werbos, P. (1995). *Neural networks for control*. Cambridge, MA: MIT Press.

Moussaoui, A., Selaimia, Y., & Abbassi, H. A. (2006). Hybrid hot strip rolling force prediction using a Bayesian trained artificial neural network and analytical models. *American Journal of Applied Sciences, 3*(6), 1885–1889.

Mukherjee, S., & Eppstein, M. J. (2012). *Differential evolution of constants in genetic programming improves efficacy and bloat*. Proceedings of the 14th annual conference companion on genetic and evolutionary computation, ACM, pp. 625–626.

Öznergiz, E., Özsoy, C., Delice, I. I., & A. Kural (2009). Comparison of empirical and neural network hot-rolling process models. *Proceedings of the Institution of Mechanical Engineers, Part B: Journal of Engineering Manufacture*, **223**, 305–312.

Rong, Z., Dan, B., & Yi, J. (2005). A bp neural network predictor model for desulfurizing molten iron. Proceedings of the First international conference on advanced data mining and applications, ser. ADMA’05, Berlin, Heidelberg: Springer.

Schröder, D. K. H. (2003). *A basic understanding of the mechanics of rolling mill rolls*. Austria: ESW & BRC.

Searson, D. P., Leahy, D. E., & Willis, M. J. (2010). GP TiPS: An open source genetic programming toolbox for multigene symbolic regression. In *Proceedings of the international multi-conference of engineers and computer scientists 2010 (IMECS 2010)* (pp. 77–80, Vol. 1). Hong Kong.

Sheta, A. F., Braik, M., Öznergiz, E., Ayesh, A., & Masud, M., (2013). Design and automation for manufacturing process: An intelligent business modeling using adaptive Neuro-Fuzzy inference system. Springer, United Kingdom, Book Chapter. Business Intelligence and Performance Management-Theory, Systems and Industrial Applications Book.

Sheta, A., & Faris, H. (2014). Improving production quality of a hot rolling industrial process via genetic programming model. *International Journal of Computer Applications in Technology, special issue on Computational Optimisation and Engineering Applications, 49*(3/4), special Issue on: ‘Computational Optimisation and Engineering Applications’.

Sheta, A., Öznergiz, E., Abdelrahman, M. A., & Babuska, R. (2009). Modeling of hot rolling industrial process using fuzzy logic. In D. Che (Ed.), *Proceedings of the ISCA 22nd international conference on computer applications in industry and engineering (CAINE)* (pp. 81–86). ISCA.

Tienasu, H., Juutilainen, I., & Rönöing, J. (2011). *Modeling the temperature of hot rolled steel plate with semi-supervised learning methods*. Proceedings of the 14th international conference on discovery science, ser. DS’11, Berlin, Heidelberg, Springer, pp. 351–364. Retrieved from http://dl.acm.org/citation.cfm?id = 2050236.2050264

Yang, X.-S., & Deb, S. (2009). *Cuckoo search via Levy flights*. World congress on nature biologically inspired computing, pp. 210–215.

X. S. Yang, & Deb, S. (2010). Engineering optimisation by cuckoo search. *International Journal of Mathematical Modelling and Numerical Optimisation, 1* (4), 330–343.