Genetic programming approach for the material flow curve determination of copper alloy – CuCrZr

L. GUSEL[*], M. BREZOCNIK, R. RUDOLF, I. ANZEL, Z. LAZAREVIĆ[a], N. ROMČEVIĆ[a]
University of Maribor, Faculty of Mechanical Engineering, Smetanova 17, 2000 Maribor, Slovenia
[a]Institute of Physics, Pregrevica 118, 11080 Belgrade, Serbia

For the control of the forming process it is necessary to know as precisely as possible the flow curve of the formed material. The paper presents the determination of the equation for the flow curve of cooper alloy (CuCrZr) with artificial intelligence approach. The genetic programming method (GP) was used. It is an evolutionary optimization technique based on the Darwinist principles of the evolution of species and the survival of the fittest organisms. The main characteristic of GP is its non- deterministic way of computing. It is probably the most general approach out of evolutionary computation methods. The comparison between the experimental results, analytical solution and the solution obtained genetically clearly shows that the genetic programming method is a very promising approach.

1. Introduction

The manufacturing of products by means of forming is particularly interesting for today's production, since subsequent treatment of the product is either reduced to the minimum or even needless. It seems that these trends will be even more present in the future since not only product cost decreases significantly but also environmental loads are less extensive than in case of many other manufacturing technologies. In order to reach high quality of the metal forming process and full functionality of the product, the properties of the material which the future product will be made of, have to be determined as precisely as possible. Traditional methods often employed to solve complex real problems tend to inhibit elaborate explorations of the search space. They can be expensive and often results in sub-optimal solutions. In most traditional modeling methods, such as multiple regression analysis, a prediction model is determined in advance. Evolutionary computation is generating considerable interest for solving real engineering problems. They are proving robust in delivering global optimal solutions and helping to resolve those limitations encountered in traditional methods. Several researches have been carried out using a neural network or genetic algorithms for modeling, thus forming process parameters [1, 2, 3, 4], but only a few dealing with much more general genetic programming method [5, 6, 7, 8].

In this work, the genetic programming approach for the determination of flow curve of copper alloy is used. Flow stress is one of the main characteristic of the metal materials. Dependence of the flow stress during forming (with constant speed of deformation and temperature) depending on the equivalent strain is called the flow curve. Experimental data obtained during pressure test serve as an environment which, during simulated evolution, the models for the flow curve have to be adapted to. No assumptions about the form and size of flow curve expression are made in advance, but they are left to the self organization and intelligence of evolutionary process. Since GP is a general optimization approach, it has been successfully applied for solving a wide range of different problems [6, 8].

1.1 Method used

Genetic programming was proposed by Koza [9]. Genetic programming is, similarly as genetic algorithm, the evolutionary computation method which imitates biological evolution of living organisms [10, 11]. It is probably the most general approach out of evolutionary computation methods.

In GP the structures subject to adaptation are the hierarchically organized computer programs whose size and form dynamically change during simulated evolution. The aim of GP is to find out the computer program that best solves the problem. Possible solutions in GP are all the possible computer programs that can be composed in a recursive manner from a set of function genes and the set of terminal genes. The set of function genes can include basic mathematical functions, Boolean functions, functions defined with respect to the problem area studied, etc. The set of terminal genes can include numerical constants, logical constants, variables, etc.

The initial population is obtained by the creation of random computer programs consisting of the available

function genes and the available terminal genes. The creation of the initial population is a blind random search for solutions in the huge space of possible solutions. The next step is the calculation of adaptation of individuals to the environment (i.e., calculation of fitness for each computer program). In GP the computer programs change in particular with reproduction and crossover.

After finishing the first cycle which includes: 1. creation of initial population, 2. calculation of fitness for each individual of the population, and 3. genetic modifying of contents of the programs, an iterative repetition of points 2 and 3 follows.

After a certain number of generations, the computer programs are usually ever better adapted to the environment. The evolution is terminated when the termination criterion is fulfilled. This can be a prescribed number of generations or sufficient quality of the solution. Since evolution is a non-deterministic process, it does not end with a successful solution in each run. The number of runs required for the satisfactory solution depends on the difficulty of the problem.

LISP language is especially well suited for GP because there is no syntactic distinction between programs and data. However, any other programming languages that can manipulate computer programs as data and that can then compile, link, and execute the new programs can be successfully used for GP.

2. Experimental procedures

The flow curve represents a link between the flow stress σ_f and the deformation φ. The flow curve is a basis for the calculation of the forming force and work. In general, the flow curves are determined by means of experiments, the most important of which are the tensile test, the pressure test, and the torsion test [12, 13]. Selection of the test method depends on the forming process. The method, in which the stress-strain conditions best coincide with the conditions during actual forming, is selected.

We decided on the pressure test with the assumption of the uniaxial stress state and homogenous change of shape (during pressing the test piece does not buckle but remains cylindrical). This is achieved by good lubrication. For the test we made cylindrical blanks with the ratio $d/h = 1.33$, where h and d is the initial height and diameter of the cylindrical blank, respectively. The material of the blanks, whose flow curve was sought for, was the copper alloy CuCrZr. The blanks were pressed in a special tool at temperature $T = 20\ ^\circ C$ and effective strain $\varphi = s^{-1}$. Teflon was used as lubricant. The pressing force was measured by means of electrical variables, therefore we made a measuring body to which we glued the resistance measuring blades. To determine the flow curve the actual

cross section of the test piece must be known in addition to the force. Therefore, during each measuring it is necessary to read the change of the test piece height.

Table 1 shows the experimental results. The effective strain φ is independent variable and the flow stress σ_f is dependent variable (i.e., measurement result). The flow curve determined by the test is mathematically not yet defined, but is given in the tabulated form or graphically in the form of a diagram. For practical use it is very suitable if the function $\sigma_f - \varphi$ can be given also in the analytical form. We selected the suitable approximation curve which should approach the experimental curve at certain interval as much as possible. The simplest model is the linear function which in most cases strongly deviates from the curve obtained experimentally.

Table 1. Experimental results (T =20°C, φ = s^{-1}).

Measurements	Effective strain φ	Flow stress σ_f [N mm^{-2}]
1	0.035	400.7
2	0.20	493.2
3	0.27	499.8
4	0.35	524.0
5	0.40	528.6
6	0.43	530.9
7	0.55	547.7
8	0.62	558.9
9	0.85	579.5
10	1.15	592.0

Therefore, it is more proper to use the power function in the following form [13]:

$$\sigma_f = C \cdot \varphi^n , \qquad (1)$$

where C is resistance constant defining the position of the curve in the diagram and n is the strain hardening coefficient defining the curve slope. By logarithming the equation (1) the curve in the double logarithm diagram has the form of a straight line:

$$\ln \sigma_f = \ln C + n \ln \varphi . \qquad (2)$$

It can be seen that for the approximate defining of the flow curve parameters it is enough to know the coordinates of two points (values for σ_f and φ) on the straight line. On the basis of the experimental results in table 1 the flow curve was approximated with the equation:

$$\sigma_f = 575 \cdot \varphi^{0,11} \qquad (3)$$

2.1 Genetic programming determination of flow curve

2.2.1 Selection of evolutionary parameters

First, ingredients from which the genetic process attempts to construct a model for flow curve must be chosen. The terminal genes consist of independent variable φ and random floating-point numbers between -10.0 and 10.0. The function genes consist of arithmetic operations of addition, subtraction, multiplication, division, and exponential function. Of course, all above arithmetic operations are protected against the extreme values. By selected genes, the evolution tries to construct the best possible equation (i.e., model) for the flow curve. For all runs the population size is 500 and the maximum number of generations to be run is 51 (i.e., generation 0 with 50 additional generations). The probability of reproduction was 0.1, probability of crossover 0.8 and probability of mutation 0.1. Maximum depth for initial random organisms was 6, maximum depth of mutation 6, and maximum permissible depth of organisms after crossover was 10.

We decided that satisfactory solution for this problem is reached when the sum of percentage errors, taken over ten measurements, between the value returned by the genetically evolved models for the independent variable (φ associated with the particular measurement and the correct value of the dependent variable (σ_f) associated with the particular measurement, is less or equal to 10%. This is the criterion of the success for this problem. Therefore, permissible average error per measurement is 1%. We made several runs of genetic programming system. The termination criterion for a run is triggered either by running the specified maximum number of generation, or by the satisfactory solutions by at least one program in the population. Possible successful models for the flow curve will be valid over intervals determined by the largest and smallest experimental values of the independent variable φ.

3. Results and discussion

The evolutionary searching for solutions starts by creation of 500 random models for the flow curve consisting of terminal genes and function genes. The creation of the initial population is a blind search for solutions in the enormous space of possible solutions. Although the result of a blind random search for solutions is bad, some models already in the initial generation are better adapted to the experimental data (i.e., environment) than the other individuals in the population. In the next generations the genetic combining of the successful solutions leads to better and better models for the flow curve, which match the experimental data relatively good. On the other hand the evolution gradually excludes bad solutions for the flow curve from the population.

Fig. 1 shows the genetic development of the models for flow curve during evolution. It can be seen that in the initial generation (i.e., generation 0) the randomly generated equation for the flow curve is very far from the desired form. Afterwards the solutions are more and more accurate and in the generation 30 the evolution develops the model for flow curve, which meets the set criteria concerning the required accuracy. This model is equal to:

$$\sigma_f = 528 + 57 \cdot \varphi - 153 e^{-4.8\varphi} \qquad (4)$$

The model (4) was generated in generation No. 30 and has the average percentage deviation Δ (i) = 0.615 %. Percentage deviation is in fact the percentage error between a single experimental value and the value predicted by the genetic model. The GP model (4) is relatively simple and describes very well the knowledge hidden in the experimental data. It is not important for the purpose of our paper that the model (4) does not include corresponding constants (e.g., C) which will make the solution physically more consistent.

Of course, the model (4) is not the only successful solution developed by evolution. Some other models for the flow curve were more precise, but they were also more complex. One of these is GP model obtained with genes function set $F = (+,-,*, /, ZEXP)$ is quite complicated and is written in LISP as:

(- ((+ (* 10.2122 φ) (+ -6.59451 7.19885)) (+ (% (ZEXP φ) (- 2.9898 -5.80549))) (+ (% (ZEXP 5.59895) (* (+ 9.20569*

2.16958) (ZEXP φ))) (% ((3.4554 - 4.89915) (* 8.219 φ)) (% (ZEXP (+ φ 8.95216)) (ZEXP φ)))) (- (+ (+ (- (%(ZEXP*

φ)) (% ((+ 9.20569 2.16958) (ZEXP φ)) (ZEXP φ))) 3.4554) (% (* (+ φ 9.219) (%(* φ (% 6.56958)))) (- (+ (* 8.219*

φ) (+ -6.59451 9.19885)) (* φ -5.45288)))) (+ (ZEXP (- (- 2.8998 -5.90548) (% 6.56958 1.10215))) (% (* (+ (* 8.219 φ)

(+ -6.58451 9.19885)) (+ (% (ZEXP φ) (- 2.9898 -5.80548))) (* (+(-6.58451 8.18885))) + (% (ZEXP φ)))))

$$(5)$$

GP model (5) can be written as a mathematical expression:

$$\sigma_f = 544.2887 - 89.6025 e^{-\varphi} + 0.023 \frac{e^{\varphi}}{\varphi} + 0.234 \cdot e^{-\varphi^{\varphi}} +$$

$$e^{\varphi} \left(7.9914 + \frac{1.3421\varphi}{\varphi \left(0.234 + 5.452\varphi + 8.218\varphi^{0.11} \right)} + \frac{1}{e^{\varphi} + 11.983} \right)$$

$$(6)$$

The model (5) was generated in generation No. 48 and has the average percentage deviation $\Delta (i) = 0.571$ %. This is slightly better accuracy than accuracy obtained by GP model (5.1) but because GP model (6) is considerably more complex than GP model (4) we choose GP model (4) to be the most suitable model regarding the accuracy and complexity. To choose the most suitable GP model out of all adequate genetic models is one of the great advantages of the genetic programming method.
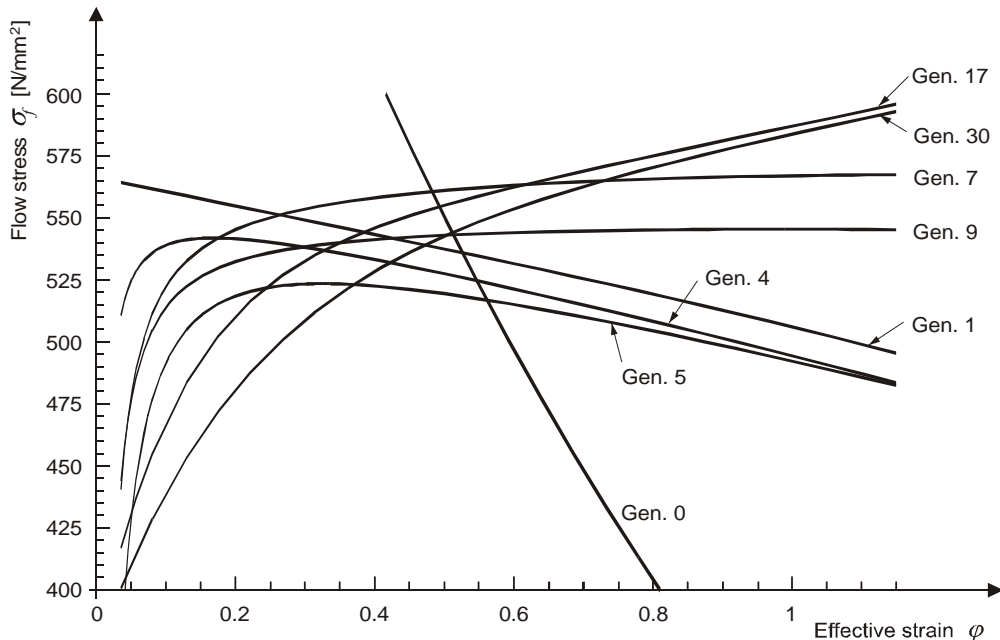

Fig. 1. Development of flow stress curve during evolution.

Fig. 2 shows the depth curve of the best models (generated with function genes F = {+, -, *, /, ZEXP}) in each generation. In generation 0 created randomly, the best models have a depth of 3. Then, from generations 2 to 15 the depth of the best models increases and reaches the value of 9. In the next generations (from 15 to 38) the

depths of models is changing between 8 and 9. Finally, after generation 38, the depth is constant and reaches a maximum depth of 10. The higher number of the model depth usually means higher complexity of the genetic model.
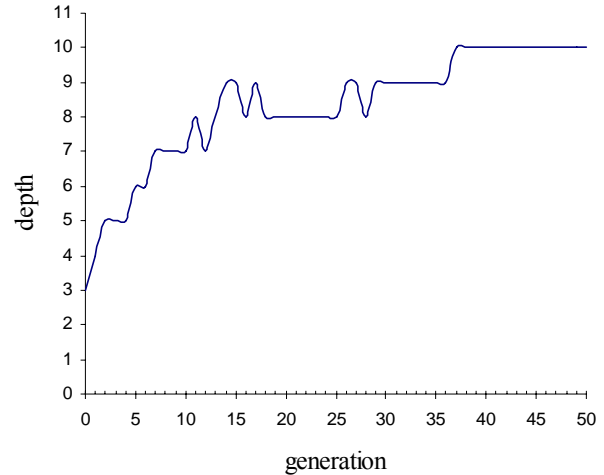

Fig. 2. The depth curve of the best models.

Table 2. Comparison between experiment data, analytical solution and GP model.

Measurements	Experiment: flow stress [N mm^{-2}]	Analytical model: flow stress [N mm^{-2}]	GP model (4) flow stress [N mm^{-2}]	Analytical error [%]	GP model (4) error [%]
1	400.7	397.663	400.656	0.764	0.011
2	493.2	481.705	480.817	2.386	2.575
3	499.8	497.872	501.526	0.387	0.345
4	524.0	512.289	519.435	2.286	0.879
5	528.6	519.870	528.369	1.679	0.044
6	530.9	524.022	533.087	1.313	0.412
7	547.7	538.403	548.432	1.727	0.134
8	558.9	545.545	555.538	2.448	0.605
9	579.5	564.812	573.863	2.601	0.982
10	592.0	583.908	592.937	1.386	0.158
			Average:	1.698	0.615

Table 2 shows a comparison between the experimental results, the results obtained with the analytical model (3), and the results obtained by the genetically developed model (4). Only in case of the measurement 2 the analytical model is more accurate than the model obtained genetically. In case of all other measurements the genetic model is considerably more accurate and deviates from the experimental values for less than 1 %.

The average error is 1.698 % in case of the flow curve obtained analytically and only 0.615 % in case of the equation for the flow curve obtained with simulated

evolution. Therefore, the model obtained without influence of human intelligence is about three times more precise than the analytical model.

In solving the problem we had to do with only one input variable (i.e., effective strain). Of course, genetic programming can be successfully used also in case it is necessary to develop a model with several independent input variables.

4. Conclusions

For successful planning of the forming process and for high quality of products it is important to know accurately the properties of the material during forming.

In the paper we presented the determination of the flow curve for copper alloy – CuCrZr by means of genetic programming. The analysis of the results showed that the equation obtained with artificial intelligence approach describes the flow curve more precisely than the solution derived analytically. It should be mentioned that no assumption about the form and complexity of the flow curve equation were made in advance, but they are left that to the intelligence of the evolution. Only two genetically developed models out of many successful solutions were presented in the paper. The GP model (4) is distinguished by small complexity, simple shape and very good accuracy, while the model (6) by even high precision but also very complex form.

The research showed that very simple and in the same time very accurate models are often hard to reach. This is due to the fact that evolution is a stochastic process, therefore, rationality in the development of the models is rare. However, the genetic programming method gives us a chance to choose the most suitable model (from all adequate models obtained during evolutionary process) for the particular problem. Sometimes we need very simple model with good accuracy and sometimes, in many metal-forming processes, not the model complexity but the accuracy of the model is of vital importance.

References

[1] B. Tanguy, J. Besson, R. Piques, A. Pineau, Engineering Fracture Mechanics **72**(49), 49 (2005).

[2] T. Spalek, P. Pietrzyk, Z. Sojka, Acta Physica Polonica A, **108**, 95 (2005).

[3] N. N. Zadeh, A. Darvizeh, A. Jamali, A. Moeini, Journal of Materials Processing Technology **164-165**, 1561 (2005).

[4] S. Mondal, M. Maiti, Computers & Industrial Engineering **44,** 105 (2002).

[5] M. Brezocnik, M. Kovacic, L. Gusel, Materials and Manufacturing Processes **20,** 497 (2005).

[6] M. Brezocnik, J. Balic, Robotics and Computer Integrated Manufacturing **17** (1-2), 113 (2001).

[7] Y. S.Chang, S. P. Kwang, B. Y. Kim, Future Generation Computer Systems **21,** 1117 (2005).

[8] H. Pierreval, C. Caux, J. L. Paris, F. Viguier, Computers & Industrial Engineering **44,** 339 (2003).

[9] J. R. Koza, Genetic Programming, The MIT Press, Massachusetts, 1992.

[10] J. R. Koza, Genetic Programming II. The MIT Press, Massachusetts, 1994.

[11] Z. Michalewicz, Genetic Algorithms + Data Structures = Evolution Programs, Springer-Verlag,

Berlin, 1996.

[12] K. Lange, Handbook of metal forming, McGraw Hill, New York, 1991.

[13] E. Doege, H. Meyer-Nolkemper, I. Saed, Fliesskurvenatlas metallischer Werkstoffe, Carl Hanser Verlag, München, 1998.

_____

[*]Corresponding author: lzorica@yahoo.com