JOURNAL OF MATERIALS PROCESSING TECHNOLOGY 209 (2009) 1538–1549

journal homepage: www.elsevier.com/locate/jmatprotec

Optimum shape design of die and preform for improved hardness distribution in cold forged parts

*Haluk Tumer, Fazıl O. Sonmez**

Department of Mechanical Engineering, Bogazici University, Istanbul, Bebek 34342, Turkiye

ARTICLE INFO

Article history:
Received 12 September 2007
Received in revised form
24 February 2008
Accepted 4 April 2008

Keywords:
Process optimization
Hardness
Cold forging
FEM
ANSYS

ABSTRACT

Hardness of a cold formed part increases due to plastic deformation induced during the forming process. Hardness affects forgability of the part as well as its mechanical performance in use. In this study, a method was proposed to optimize the shape of preform and die to minimize variation in hardness distribution. The method was applied to backward extrusion. A finite element model of the mechanical behavior of the material during processing was developed to predict the effective strain distribution in the part. The hardness distribution was determined using an analytical relation between effective strain and hardness. An optimization code based on the Nelder-Mead search algorithm was developed and integrated with the finite element model. The results showed the effectiveness of the proposed method.

1. Introduction

Knowing that hardness is a measure of strength and resistance to wear, it should be considered as an important quality parameter. Accordingly, it highly influences the performance of products in their service life. High levels of hardness in a finished product may increase its performance for instance in metal cutting applications but its accompanying low toughness makes this part unsuitable for applications requiring high impact resistance. Uneven hardness distribution is another concern. The regions of high hardness may make the part brittle while the regions of low hardness may detract from its wear resistance. In cold forming operations, hardness is not only an output of the process and a quality parameter of the end product, but it is also a parameter that must also be monitored during the process. During cold forging of a part, the material might become so hard that the available press power may become unable to deform the workpiece further or the dies and the part itself might be damaged. In order to overcome this problem, forging operations are usually completed in substeps followed by annealing. A cold forming process should therefore be carefully designed so that hardness of the material during as well as after the processing remains within the acceptable limits.

The change in hardness depends on the permanent deformation induced during cold forming. By controlling deformations in a cold forging process, one may control hardness and thus achieve a desired level of hardness. For this purpose, one should know the relationship between deformation and hardness for a cold forged part. In that case, if the controllable process parameters are related to deformation, one may control deformation, and thus hardness. This means, one may find the values of process parameters resulting in the desired hardness distribution through an optimization procedure.

In the literature, there were a number of studies on optimization of cold forging processes. Most of them focused on

* *Corresponding author*. Tel.: +90 212 359 7196; fax: +90 212 287 2456.
E-mail address: sonmezfa@boun.edu.tr (F.O. Sonmez).
0924-0136/$ – see front matter © 2008 Elsevier B.V. All rights reserved.
doi:10.1016/j.jmatprotec.2008.04.017

two major objectives: The first one was to obtain an accurate final geometry and the second one was to decrease the applied pressure in order to lower the required press capacity, and thus to prevent dies from failures and improve tooling life. The objective of the study conducted by Thiyagarajan and Grandhi (2005) was to reduce the difference between the desired and achieved geometries of the finished product and to minimize the strain variance by optimizing preform shapes in multi-level forging processes. They defined the boundary curves of preform shapes by means of coordinates of control points lying on these boundary curves. They applied the reduced basis technique to decrease the number of optimization variables. Response surface methodology (RSM) was chosen to build an approximation model and perform optimization. The proposed model was applied on multi-level forging of a steering link. They reported that a preform shape that enabled complete die fill was obtained. Zhao et al. (2002) proposed an optimization method for preform shape design. The objective was to minimize the effective strain variation in an H-shaped forging. They defined the preform die shape through B-spline curves. The coordinates of control points were design variables. Broyden–Fletcher–Goldfarb–Shanno (BFGS) algorithm was employed to minimize the objective function. The final maximum and minimum strains were obtained to be 1.3608 and 0.350, respectively; they were 1.605 and 0.296 for the non-optimized case. Although there are similarities between Zhao and Wang's study and the present study, e.g. FEM is employed to obtain effective strains; our objective is to obtain a uniform hardness distribution in a cold forged part, which is directly related to the part quality, rather than to minimize the variation in effective strain. Shi et al. (2004) proposed a shape optimization technique to improve stamping quality by minimizing the risk of rupture, wrinkles and unstretched areas. They utilized the modified sweeping simplex as search algorithm. They integrated this algorithm with a FE model and applied it on a front fender to validate its results. Celano et al. (2001) developed a technique for optimum process design of multi-pass cold drawing process using the simulated annealing algorithm. The objective of that study was to determine the optimum number of passes and pass-schedule to keep the drawing stresses on the wire below a safe limit. They also showed the effectiveness of their approach by comparing it with a set of industrial sequences for wires of different materials. Gao and Grandhi (2000) proposed a sensitivity analysis based optimization method for improving the grain size variation in a forged part. The optimization variables were billet shape and die velocity. They also set optimization constraints for die underfill and excessive material waste. They applied their method on a turbine disk of waspaloy (a nickel based, age-hardenable super alloy) and obtained optimum billet geometry and die velocity. Roy et al. (1997) carried out a comprehensive study on optimization of multi-stage cold forming processes. Multi-pass cold wire drawing, multi-pass cold tubular profile drawing and multi-pass cold forging of an automotive outer race were the selected processes. They used a micro genetic algorithm to optimize the process variables.

The literature survey shows that there has been no study on controlling hardness of cold forged parts by optimizing geometries of preform and die. In this study, a methodology for optimum process design of cold formed parts was proposed to obtain uniform hardness distribution. The process to be optimized was chosen to be backward cold extrusion. First, a finite element model of the process was developed in ANSYS environment. This FE model enabled calculation of the effective strains induced in the part during processing. Then, the strain–hardness relationship proposed by Demir and Sonmez (2004) and Sonmez and Demir (2007) was used to convert effective strain values to hardness and thus to determine the hardness distribution. The next step was to develop an optimization code based on the Nelder-Mead algorithm (Mathews, 1999). The objective function was defined as the sum of variations in hardness from the average hardness in the part, and some of the geometric parameters defining die and preform shapes were chosen as optimization variables. At last, the FE model and the optimization code were integrated. Having carried out optimization, the optimum preform and die shapes were determined.

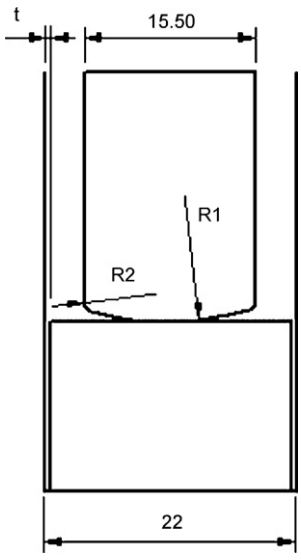

Fig. 1 – Die and workpiece geometries (dimensions are in mm).

2. Problem formulation

The backward (cup) extrusion process was considered in this study because of its prevalence in the industry, higher variation in the strain distribution as a result of highly nonuniform deformation state induced in the workpiece and compatibility to computationally efficient axisymmetric modeling.

The geometries of the die and workpiece are depicted in Fig. 1, while the geometry of the extruded part is shown in Fig. 2. The geometries described in these figures were adopted from the study of Hur et al. (2003), in which a design method for FE analysis of backward extrusion process was proposed.

The thickness of the walls, the height and depth of the cup are assumed to be specified according to the requirements of the product. Since a flat punch tip increases the press capacity required for inducing the desired deformations, the punch tip is to be curved. The radii of curvatures, *R*1 and

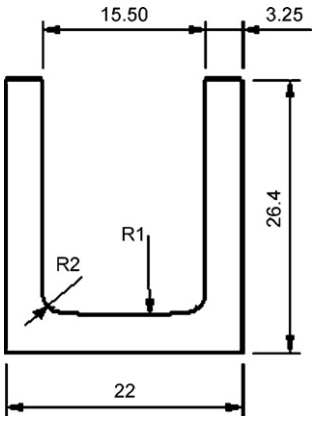

Fig. 2 – Geometry of the finished product (dimensions are in mm).

R2, defining the tip geometry of the punch are assumed to be variables for which appropriate values can be chosen during the process design phase. Only, upper and lower limits may be set because of the product requirements during the product design phase. The initial shape of the workpiece is cylindrical. In order to place the workpiece inside the die walls, the workpiece diameter should be smaller compared to that of the die walls. Accordingly, a gap exists between the preform and the inner walls of the die. The clearance, t, is to be determined during the process design phase. R1, R2 and t are therefore considered as independent process optimization variables.

Since there may not be a feasible geometry corresponding to an arbitrary set of values for the optimization variables, they are not free variables. For this reason, constraints are imposed on the optimization variables considering geometric constraints, design and process requirements, and possible instability sources in FE analysis. For fillet radius R1, the minimum allowable value is 7.75 mm. which is equal to the radius of the punch. If both R1 and R2 are chosen to be 7.75 mm, the shape of the punch tip becomes spherical and smaller radii of curvatures may not be defined. On the other hand, there is no upper limit set for R1. If it tends to go to infinity, the tip becomes flat.

The upper limit for R2 is 7.75 mm, which is the half of the punch radius. If R2 takes the value of its upper limit, the curve defined by R2 eliminates R1, and the punch tip becomes spherical. The lower bound of R2 is 3 mm. A lower value may lead to difficulties in FE analysis. This limit was set due to a number of failed FE analyses below this value.

A very small clearance makes the placement of the preform inside the die very difficult. For this reason, a clearance of 0.1 mm is taken as the lower limit of t. That means the workpiece should be machined to a diameter being at least 0.2 mm smaller than the diameter of the die walls. The upper limit of t is 0.5 mm. In practice, this limit is too large, since it may cause difficulties in workpiece positioning and uncertainties in deformation behavior. The upper limit is taken to be large, because the feasible domain should be chosen as wide as possible, and should not be restricted based on the common practice. Otherwise, one may exclude the globally optimal point from the feasible domain. As will be mentioned later, the upper limit of 0.5 mm is not active since the optimum value of the die gap converges to its lower bound.

3. Hardness–strain relations

Hardness increases with increased amount of deformation. Altan et al. (2005) stated that hardness distribution in cold forged parts could be predicted if the hardness–effective strain relation of the material and the effective strain distribution in the finished part were known.

Sonmez and Demir (2007) proposed analytical relations between the effective strain induced in a cold formed workpiece and the resulting Brinell or Vickers hardness. This relation is based on the flow curve constants, which are easily obtainable.

Vickers hardness (HV) is defined as the load applied on a pyramidal indenter divided by the surface area of the permanent impression. Since the top angle of a pyramid is 136°, and its base area is equal to 0.9272 times of its lateral area, HV can be expressed as

$$\mathrm{HV} = 0.9272 P_\mathrm{m} \tag{1}$$

where P_m is the mean pressure (the applied load, F, over the projected area, A). The base area of the pyramidal impression, A, is calculated by

$$A = \frac{s^2}{2} \tag{2}$$

where s is the length of the base diagonal. HV is then given by

$$\mathrm{HV} = 0.9272 \left(2\frac{F}{s^2}\right) \tag{3}$$

Consequently, HV can be calculated if the diagonal of the indentation s is measured.

Tabor (1948) assumed that the stress developed at a certain point is representative of the whole deformation and it is linearly related to the average pressure.

$$P_\mathrm{m} = \alpha\sigma_\mathrm{e} \tag{4}$$

Here, α is a constant and σ_e is the representative stress.

Considering that the indenter used in Vickers hardness measurements is pyramidal, if the magnitude of the acting force is increased, the depth of the impression changes whereas the impression shape remains unchanged. So the ratio h/a (the indentation depth/contact length) remains constant regardless of indentation depth (see Fig. 3). That makes $\varepsilon(x/a, y/a)$ independent of the indentation size. Representative yield strain ε_e corresponding to representative yield stress is independent of the load and hardness, and according to Tabor (1948) it is equal to 0.08.

The true stress–true strain relation under uniaxial loading is described as

$$\sigma = K\varepsilon^n \tag{5}$$

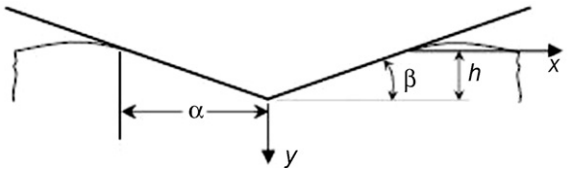

Fig. 3 – Geometry of a Vickers indentation (Sonmez and Demir, 2007).

where K and n are material constants. The relation given by Eq. (5) is also valid for the representative stress and representative strain (Tabor, 1948). Thus, one may state

$$\sigma_e = K\varepsilon_e^n \tag{6}$$

From Eqs. (1), (4), and (6)

$$HV = cK\varepsilon_e^n \tag{7}$$

The constant c in Eq. (7) was obtained by Tabor (1948) as 2.9. This equation is valid only for materials without any prior deformation. Since plastic strains are induced during a cold forging process, the effect of work hardening should be considered in the HV equation. Tabor (1948) found that the prior uniaxial plastic strain, ε_o, could be added to ε_e. Eq. (7) then becomes

$$HV = cK(\varepsilon_o + \varepsilon_e)^n \tag{8}$$

Substituting the values of the constants, Eq. (8) may also be written as follows

$$HV = 2.9K(\varepsilon_o + 0.08)^n \tag{9}$$

Sonmez and Demir (2007) adapted this equation to 3D deformation problems by substituting effective strain for ε_o in Eq. (9), which is given by

$$\varepsilon_o = \sqrt{\frac{2}{3}e_{ij}e_{ij}} \tag{10}$$

where e_{ij} are deviatoric strain components expressed as

$$e_{ij} = \begin{bmatrix} \frac{1}{3}(2\varepsilon_{11} - \varepsilon_{22} - \varepsilon_{33}) & \varepsilon_{12} & \varepsilon_{13} \\ \varepsilon_{12} & \frac{1}{3}(2\varepsilon_{22} - \varepsilon_{11} - \varepsilon_{33}) & \varepsilon_{23} \\ \varepsilon_{13} & \varepsilon_{23} & \frac{1}{3}(2\varepsilon_{33} - \varepsilon_{11} - \varepsilon_{22}) \end{bmatrix} \tag{11}$$

Substituting Eq. (11) into Eq. (10), one obtains

$$\varepsilon_o = \sqrt{\frac{2}{3}(e_{11}^2 + e_{22}^2 + e_{33}^2 + 2e_{12}^2 + 2e_{13}^2 + 2e_{23}^2)} \tag{12}$$

Effective strain as a function of principal strains, e_1, e_2 and e_3 is given by

$$\varepsilon_o = \sqrt{\frac{2}{3}(e_1^2 + e_2^2 + e_3^2)} \tag{13}$$

In this study, after determining the effective strain through FE analysis Eq. (9) is used to calculate Vickers hardness. Use of this analytical relationship proposed by Sonmez and Demir (2007) improves the applicability of the approach proposed in this study. Since this relationship (given in Eq. (8)) is based on two universal (ε_o and c) and two flow curve constants (K and n), one may apply this relationship to another material after obtaining its flow curve constants. Besides, these constants are available in the literature for a wide range of materials.

4. Methodology

4.1. *Optimization procedure*

In the process optimization procedure, first of all, the objective function is analytically expressed. Then, the optimization variables and constraints are defined. The next step is to choose a proper search algorithm. In this study, the Nelder-Mead algorithm was selected, because it is a robust zero-order search algorithm not requiring numerical derivatives of the objective function. To integrate the finite element model and the optimization algorithm, a code was developed using the built-in programming language in ANSYS. This code carries out FE analyses, writes the results into output files, and also evaluates the results to modify the values of the optimization variables according to the decision criteria of the Nelder-Mead algorithm.

Fig. 4 describes the optimization procedure followed in this study. Initially, the optimization code selects random values for the optimization variables within the feasible domain and creates an initial geometry according to these selected values. Then, predefined loads, boundary conditions, material properties are defined in the finite element model. Following that, a finite element analysis is conducted, and first, second and third principal plastic strains are obtained at each node. Using the plastic strain data, the code calculates the resulting effective strain using Eq. (13). To obtain Vickers hardness, Eq. (9) and the calculated effective strain data are used. As mentioned before, the objective function is defined as the sum of variations from the average Vickers hardness value. The details of the calculation method for the objective function are given in the following section. Because, there are constraints on the optimization variables, this is a constrained optimization problem. In cases of constraint violations, a penalty is calculated and added to the value of the objective function. In this way, the problem is transformed to an unconstrained optimization problem, for which the search algorithm is suitable. Penalties are activated under two conditions: firstly, if an optimization variable takes a value outside its feasible range, and secondly, if the FE analysis fails. Further details of the penalty method are given in the next section. Because the Nelder-Mead algorithm requires $k+1$ different current configurations, k being the number of design variables, this procedure is repeated for four randomly created initial configurations. Having completed $k+1$ FE analyses and obtained objective function values, the program calculates a new set of values for the optimization variables and creates a new geometry to be analyzed according to the decision criteria of the Nelder-Mead algorithm. This procedure is repeated until the

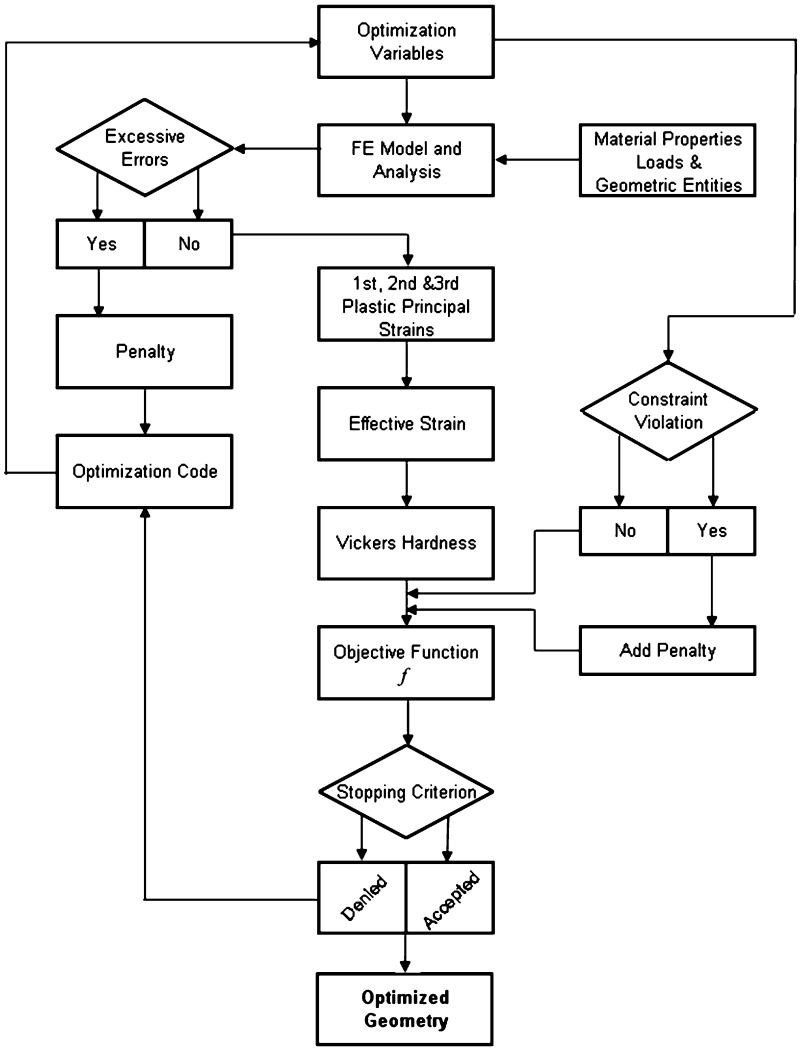

Fig. 4 – The optimization procedure.

stopping criterion is satisfied, which requires the difference between the objective function values of the best and worst current configurations to be small.

Since Nelder-Mead is a local search algorithm, it does not necessarily converge to the global optimum. For this reason, the optimization process should be repeated many times starting from different points within the feasible region of process variables. Then, the lowest value is chosen as the global minimum of the objective function.

4.2. *Definition of the objective function and penalty functions*

In order to minimize the variation in Vickers hardness through an optimization method, this variation should be defined analytically. Therefore, an approach similar to that of Zhao et al. (2002) was used except that weighting factors were not utilized. Basically, the objective function is expressed as the sum of variations from an average Vickers hardness value.

The Vickers hardness is related to the effective strain as expressed in Eq. (8). Since FE results are obtained at the nodes, Eq. (8) is reformulated for the nodes as

$$HV_i = cK(\varepsilon_{oi} + \varepsilon_e)^n \tag{14}$$

Here, HV_i is the Vickers hardness and ε_{oi} is the effective strain at the ith node. Note that the unit of HV is kgf/mm^2.

The average Vickers hardness within the extruded part, HV_{avg}, is given by

$$HV_{avg} = \frac{1}{n}\sum_{i=1}^{n} HV_i \tag{15}$$

and n is the total number of nodes. One may define the deviation of the hardness at node i from HV_{avg} as

$$HVE_i = |HV_{avg} - HV_i| \tag{16}$$

The degree of inhomogeneity of hardness within the whole part is formulated as

$$HVE_T = \frac{1}{HV_{avg}^2}\sum_{i=1}^{n} HVE_i^2 \tag{17}$$

Note that HVE_T is defined as a power function to intensify the effect of deviation in hardness, and it is divided by HV_{avg}^2 for normalization.

The objective function, f, is defined as

$$f(R1, R2, t) = w_1 HVE_T + w_2 \sum_{k=1}^{l} P_k + c\eta \quad (18)$$

Here, l is the number of applied penalty functions.

During the optimization process, the value of the objective function is recalculated whenever the values of optimization variables are changed by the search algorithm. For this purpose, non-linear FE analyses involving large deformation are performed. The search algorithm may generate a set of values for the variables such that for some reason FE analysis fails, e.g. the geometry may not be constructed due to a negative value for the radius of curvature. In that case, effective strains at the nodes, (ε_{oi}), cannot be calculated. ANSYS automatically assigns zero value for them. According to Eq. (14) and Eq. (15), HV_i become equal to the hardness of the starting material. Consequently, the algorithm obtains zero variation in Vickers hardness and it sticks to this fake minimum point. In order to prevent this, an analysis error term, η, is defined that becomes active and takes a large value compared to other terms in the objective function, if FE analysis fails. η is used with a control coefficient c. In such a case, c is equal to 1.0 and the third term becomes nonzero. Otherwise, c is equal to zero.

In Eq. (18), w_1 is the weighting constant for HVE_T and it is equal to 1000, w_2 is the weighting coefficient for the penalties of constraint violations and it is equal to 10. For all penalty functions the same weighting constant is used since all optimization variables have about the same degree of importance in the objective function.

Suppose that for the selected optimization variable, x, there exist a lower and an upper bound denoted by x_ℓ and x_u, respectively. The inequality constraint is expressed as

$$x_\ell < x < x_u$$

This relation requires that two penalty functions be defined. For the lower bound, the penalty function is defined as

$$P_k = \left\langle \frac{-x + x_\ell}{x_\ell} \right\rangle \quad (19)$$

and the penalty function for the upper bound is

$$P_{k+1} = \left\langle \frac{x - x_u}{x_u} \right\rangle \quad (20)$$

Because the type of the penalty functions is external, they become active only if their related constraint is violated. Otherwise, they are equal to zero. This condition is controlled by the operator "$\langle\,\rangle$". If the value of the term inside this operator is positive, it yields the same value, otherwise it yields zero. Note that all penalty functions are defined in a manner such that they become equal to zero, if their related variable takes a value within its feasible range.

For $R1$, there exists only one penalty function since it is constrained in one direction only whereas two penalty functions are required for both $R2$ and t, which are constrained in two directions. Consequently, five penalty functions are employed in the objective function.

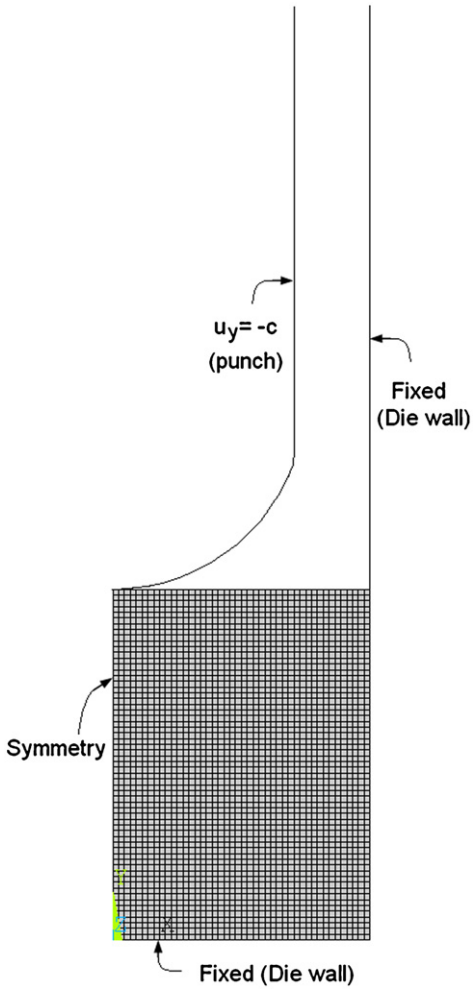

Fig. 5 – FE model with the applied boundary conditions.

5. Finite element modeling

Fig. 5 shows a geometric representation of the punch, workpiece with its mesh structure and dies in ANSYS. Note that an axisymmetric model is used instead of a 3D model, since use of an axisymmetric model greatly reduces the modeling and analysis time compared to that of an equivalent 3D model without compromising the accuracy.

5.1. *Meshing*

Selection of an appropriate finite element type is essential for obtaining reliable analysis results. Selected element should satisfy a set of requirements: Firstly, since axisymmetric approach is adopted in this study, the element type has to be compatible with axisymmetric modeling. Secondly, the element should have large deflection, large strain capabilities. Lastly, a high order element is more suitable for highly nonlinear deformation. Accordingly, the element type chosen for

meshing the workpiece is Plane183, being a high order, 8-node 2D rectangular element. In order to determine the mesh density that may enable accurate calculation of the strain state, a convergence analysis was carried out for the element size and 0.3 mm edge length was found to be appropriate. Using a smaller element size did not yield appreciably different results.

Assuming that the die and the punch do not undergo plastic deformation and their elastic deformations have a negligible effect on the deformation of the workpiece, their surfaces are defined as non-deformable by using rigid lines surrounding their representative areas, and thus they are not meshed. Only the workpiece is meshed. Consequently, the number of finite elements to be evaluated during the analysis is reduced, and the required solution time is decreased.

5.2. *Creation of contact elements*

The next step in the FE analysis is the creation of contact elements for the die-workpiece and punch-workpiece interfaces. The contact type used in this study is rigid-to-flexible and surface-to-surface contact. The dies and punch are defined as rigid target bodies whereas the workpiece is a deformable contact body. In order to establish contact pairs, the boundary lines of the bodies have to be meshed. For this purpose, CONTA172 was selected for the deformable lines, and TARGE169 was selected for the non-deformable lines. In order to create a contact, the groups of nodes which probably will come in contact should be specified.

5.3. *Boundary conditions*

In the FE model, there are only displacement boundary conditions and the axisymmetry condition applied on lines and nodes. The lines representing the die walls are fixed in all degrees of freedom. The punch lines are restrained from rotating and moving along the x-axis. They are only allowed to move through the prescribed vertical displacement. The punch movement along the y-axis is controlled with a pilot node. Friction forces exist between the contacting surfaces impeding movement.

In ANSYS, a load step is applied in increments with a certain number of substeps. A convergence analysis was also performed to determine the number of substeps necessary for accuracy. About 4000 substeps were used to obtain accurate results.

5.4. *Friction coefficient*

In the literature, one of the proposed methods to determine the friction constant for die–workpiece interface was to conduct ring compression tests. Gouveia et al. (1999) conducted FE simulations of forward extrusion process. They utilized Coulomb friction model with a constant friction coefficient of 0.18, which was obtained by ring upsetting experiments. Having conducted finite element analysis of the process, they carried extrusion experiments to validate their model. Experimental measurements showed that the friction coefficient was not in agreement with the expectations from the ring compression tests. Consequently, they stated that this disagreement might be due to the differences in the deformation behavior and the pressure, and they also stated that applying a lower friction coefficient would give results closer to that of experiments.

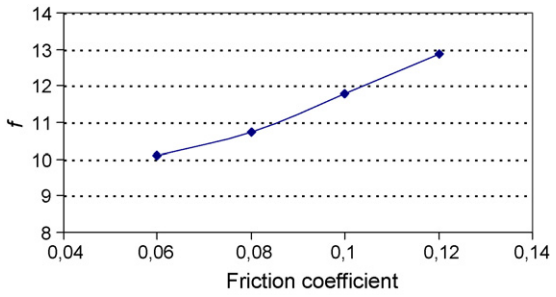

Fig. 6 – The effect of friction on the objective function.

Hur et al. (2003) simulated backward extrusion process. The friction coefficient was assumed to be constant and equal to 0.1. Petruska and Janicek (2003) studied strain inhomogeneity through hardness measurements on cold formed products. The type of the forming process they considered was forward extrusion. They simulated friction by Coulomb's model, and the friction coefficient was equal to 0.15. Roy et al. (1997) simulated extrusion process of automotive outer race preforms. They applied a constant friction coefficient of 0.1 in their FE model.

In this study, Coulomb's model was assumed to correctly reflect the friction at the die–workpiece and punch–workpiece interfaces as in the previous studies. However, there is no clear agreement on the value of the friction coefficient. In practice, friction coefficient may depend on the type of the materials in contact, and also on lubrication and surface roughness. In this study, the friction constant was assumed to be 0.1, which was within the range of values adopted in the previous studies.

In order to observe the effects of variations in the friction coefficient, a number of FE simulations were conducted. The results are given in Fig. 6. The change in the friction coefficient from 0.1 to 0.08 results in a reduction of 8.9% in the objective function value. On the other hand, if it is increased to 1.2, an 8.6% increase is observed. The difference between the objective function values for the cases with friction coefficients 0.06 and 0.12 is about 27.5%. Consequently, it can be stated that the friction coefficient should be minimized to obtain a more uniform hardness distribution, which indicates the importance of lubrication in cold forging processes.

5.5. *Material model*

In this study, the selected material model is multi-linear isotropic hardening (MISO) model. MISO is a rate independent model suitable for large strain applications. In MISO, the stress–strain curve is described by a set of linear sub-elements instead of a power equation. To define the stress–strain curve to be used in the non-linear analysis, datum points from the stress–strain curve are required. The more points are entered, the better the non-linear behavior is approximated. The MISO flow curve of ST37, shown in Fig. 7, is defined by 38 datum points from the flow curve of the material.

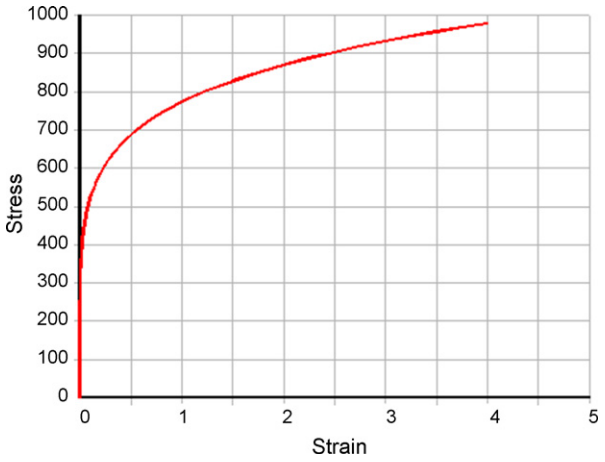

Fig. 7 – MISO flow curve of St37.

6. Results and discussions

6.1. *Initial appraisal of the constraints*

After having developed the optimization code to be run in ANSYS environment, a group of optimization iterations were conducted. Optimization results showed that some variables tend to converge to their either upper or lower limits. If a variable is monotonically increasing to its upper limit, or monotonically decreasing to its lower limit during the optimization process, there are two alternatives: one is to extend the initially chosen feasible domain to search for better configurations beyond these limits unless this violates design requirements, process restrictions, and other criteria related to computational difficulties. The other alternative is to assign this limit value to the parameter assuming that this is the optimum value for this parameter; it thus becomes a constant in the following optimization runs. In this way, the global minimum can be located more easily and computational cost can be decreased.

Table 1 gives the initial configurations in the first run, and the resulting optimum values. A second run was conducted under the same conditions with a different set of starting values for the variables as given in Table 2.

Note that t and $R2$ converged to their lower bound in the first and second runs. For this reason, the corresponding constraints and penalty functions are modified as will be discussed in the following section.

Table 1 – Starting points and objective function values, f, of the first run and the resulting optimum values

	R1 (mm)	R2 (mm)	t (mm)	f
1	30.891	5.585	0.291	15.599
2	57.974	4.140	0.174	13.136
3	15.874	5.059	0.495	15.558
4	53.388	4.302	0.386	14.171
Optimum	78.00	3.00	0.11	11.43

Table 2 – Starting points and objective function values, f, of the 2nd run and the resulting optimum values

	R1 (mm)	R2 (mm)	t (mm)	f
1	85.468	5.075	0.309	14.874
2	102.921	4.007	0.235	13.320
3	43.702	5.585	0.291	15.583
4	85.780	4.140	0.174	13.156
Optimum	109.65	3.01	0.12	11.57

In contrast to $R2$ and t, $R1$ does not show a tendency to converge to its lower bound. Therefore, optimization is continued without any modification on the constraint and corresponding penalty function of $R1$.

6.2. *Optimum values of the design variables*

6.2.1. *Die gap t*

At the initial phases of the optimization, the smallest allowable value of t was taken as 0.1 mm. Since t converged to its minimum limit, the feasible range was to extended to $0.05 < t < 0.5$ mm. However, in this case the preform should be machined to a higher degree of precision, and this will increase the manufacturing cost. Using the new lower bound for t, the optimization process was repeated. As shown in Table 3, t again converged to its minimum. Note that although $R1$ and $R2$ converged to the same values (109 mm and 3 mm, respectively) as shown in Tables 2 and 3. The optimum objective function value is smaller for the smaller lower limit. This implies that choosing a smaller clearance is conducive to a more uniform hardness distribution.

The lower limit of t was further shifted to 0.01 mm. The same initial configurations were used as in the previous case. Only penalty functions were updated according to the new domain of t. The resulting configuration and the corresponding objective function value are given in Table 4. t again converged to its lower bound, 0.01 mm. One may not choose a smaller clearance, because this requires too tight machining tolerances for billets unless highly increased machining costs can be tolerated.

One may conclude that the theoretical optimum value of t tends to approach zero. In practice, the optimum value of t is the tightest tolerance allowed by the manufacturing process used for preparing the workpiece. In this study, t is taken to be equal to 0.01 mm. From this point on, it becomes a constant, and it is not evaluated in further optimization runs. As a result, the optimization problem is simplified from 3D to 2D.

Table 3 – Starting points and the results for the new feasible range for t; $0.05 < t < 0.5$ mm

	R1 (mm)	R2 (mm)	t (mm)	f
1	111.163	5.075	0.309	14.872
2	136.835	4.007	0.235	13.341
3	56.514	5.585	0.291	15.574
4	113.586	4.140	0.174	13.171
Optimum	108.59	3.01	0.05	11.19

Table 4 – The results for the new feasible range for t; 0.01 < t < 0.5 mm

R1 (mm)	R2 (mm)	t (mm)	f
82.392	3.008	0.011	10.868

6.2.2. *Corner radius R2*

Tables 1–4 show that R2 converges to 3.0, as t converges to its minimum allowed value. Because the value for the lower limit of R2 was chosen in order to avoid difficulties that sharp corners may cause in finite element calculations and also wearing away of the punch after repeated use, one may assume that there is a leeway to relax this constraint also. But first, a new optimization run was conducted with the die gap t having fixed to 0.01 mm. Starting points and the resulting optimal values are given in Table 5. It is observed that R2 converges to 3.0 as in the previous cases.

After having ensured that the optimum value for R2 does not lie close to its maximum boundary, the upper limit was shifted to 5.0 mm whereas the new lower limit was chosen to be 2.0 mm. Another optimization run was conducted using these new limits. The corresponding results given in Table 6 show that R2 again converges to its lower limit. In this case, the limits of R2 were revised as 1.5 and 4.0 mm, and the optimization process was repeated. Table 7 shows that R2 converges to 1.51 mm. This result implies that R2 tends to converge to the minimum possible value allowed by its constraints. On the other hand, too small R2 values may lead to stress concentration, shorter tool life and decreased reliability of FE results because of poor convergence. Accordingly, 2.0 mm was considered as a suitable value for R2.

In the following analysis, R2 becomes a constant parameter instead of being an optimization variable. Consequently, the only remaining optimization variable is R1. The optimization problem then becomes a one dimensional problem, which decreases the computational effort.

6.2.3. *Bottom radius R1*

Having obtained the optimum values of R2 and t, the remaining optimization variable is the punch bottom radius R1. Because the number of design variables is now one ($k = 1$), the algorithm uses two ($k + 1$) current configurations. The results of the optimization runs that were carried out to find the optimum value of R1 are given in Table 8, which shows that the optimal value of R1 is between 31.525 and 31.248. In this interval, the change in the objective function value is less than 0.001. Thus, the optimum value of R1 may be assumed to be 31.5 mm.

The values given in Table 9 were selected from the results of different runs, in which R1 exceeded 100 to show that beyond 100 mm, the more R1 is increased, the higher objective function results were obtained. In addition, for the values of R1 higher than 200 mm, the curvature on the punch tip becomes nearly flat. Thus, a dramatic change in the objective function values is not expected beyond this limit. One should also note that the objective function is not very sensitive to changes in R2 when it is large. Thus, if the bottom should be almost flat as a design requirement, the increase in the objective function will not be large.

6.3. *Evaluation of the results*

The optimum values of R1, R2, and t were found to be 31.5, 2.0 and 0.01 mm, respectively. These values of the optimization variables result in an objective function value of 9.573. The worst four configurations generated by the algorithm during the iterations and their resulting objective function values are given in Table 10. The worst one is 18.007. There is about 46.8% reduction in the variation of the Vickers hardness distribution, which is a dramatic improvement between the worst and the best case.

Table 5 – Starting points and the results for the case of 3.0 < R2 < 7.75

	R1 (mm)	R2 (mm)	f
1	66.822	5.385	13.820
2	89.593	5.235	13.587
3	136.835	4.007	12.141
Optimum	167.80	3.00	10.99

Table 6 – Starting points and the results for the new feasible range for R2; 2.0 < R2 < 7.75 mm

	R1 (mm)	R2 (mm)	f
1	66.822	4.077	12.174
2	89.593	3.854	11.850
3	136.835	2.010	9.949
Optimum	137.05	2.00	9.95

Table 7 – Starting points and the results for the new feasible range for R2; 1.5 < R2 < 4 mm

	R1 (mm)	R2 (mm)	f
1	79.170	2.301	10.229
2	128.443	2.955	10.928
3	113.586	1.675	9.750
Optimum	36.17	1.51	9.34

Table 8 – Optimization results for R1

Run	Initial configurations		Final configurations	
	R1	f	R1	f
1	113.163	9.868	83.701	9.805
	89.593	9.806	86.647	9.805
2	101.828	9.836	85.966	9.805
	136.835	9.899	86.513	9.804
3	8.257	17.549	31.525	9.573
	79.17	9.809	31.248	9.573
4	79.17	9.809	45.335	9.695
	56.514	9.795	45.324	9.695
5	8.0	18.007	31.438	9.573
	200.0	9.917	31.25	9.573

Table 9 – Optimization results for R1 > 100

Run number	R1 (mm)	R2 (mm)	t (mm)	f
1	104.000	2.000	0.010	9.843
2	136.835	2.000	0.010	9.899
3	152.000	2.000	0.010	9.904
4	200.000	2.000	0.010	9.917

In the best case, the highest and lowest Vickers hardness values are 284.73 and 148.75, whereas the extremum values in the worst case are 286.81 and 148.95. Because our objective is to reduce variation in hardness, not its maximum value, it can be stated that the applied procedure has a little effect with regard to decreasing the maximum hardness value or the difference between that of the highest and the lowest.

Note in Table 10 that the worst two objective function values were obtained using the optimum values of R2 and t, which shows the sensitivity of the objective function to R1 near its lower bound. In contrast, this sensitivity to R1 decreases when it takes a large value as indicated before. For instance, for two

Table 10 – The worst four configurations

Run number	R1 (mm)	R2 (mm)	t (mm)	f
1	8.000	2.000	0.010	18.007
2	8.257	2.000	0.010	17.549
3	57.774	5.553	0.309	17.061
4	30.891	5.585	0.291	15.599

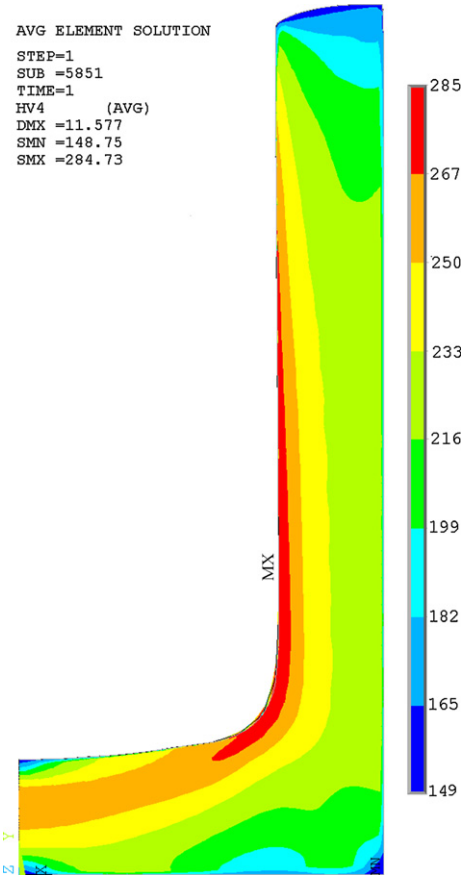

Fig. 8 – The optimum Vickers hardness distribution.

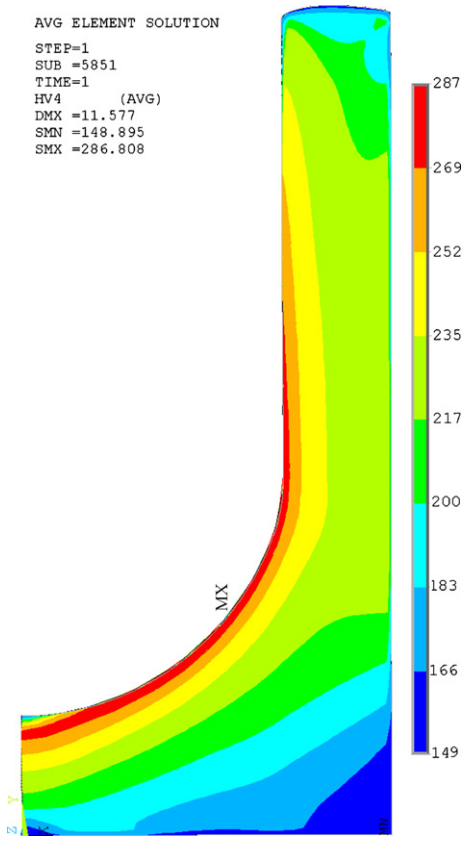

Fig. 9 – The worst Vickers hardness distribution.

different values of R1 (83.701 and 86.647), the same objective function value is obtained (9.805), although there is a 3.5% difference between them. On the other hand, a change in R1 from 8 to 8.257 (3.2% difference) resulted in an increase of 2.6% in the objective function value.

Although the objective function is sensitive to R1 near its minimum bound, the effects of R2 and t cannot be neglected. For the case that R1 is equal to 30.891 mm, which is close to its optimum value, the corresponding function value is 15.599, which is the fourth worst result in the whole optimization procedure. This result was due to the values of R2 and t being far from their optimum values. As mentioned before, the weights of R1, R2 and t were taken to be same. The results show that this was an appropriate choice.

In Figs. 8 and 9, Vickers hardness distribution plots are given for the best and worst geometries, respectively. Hardness of the inner region increases during extrusion because of large deformation. The FE mesh in the deformed shape is shown in Fig. 10.

6.4. *Assumptions and discussions*

In this study, the following assumptions were made which might detract from the accuracy of results:

First of all, material flow curve constants *K* and *n* for St37 are 773 and 0.171, respectively. These constants define the deformation behavior of the material. Besides, they are also used in the analytical relationship between Vickers hardness and effective strain. The effects of the experimental errors or the

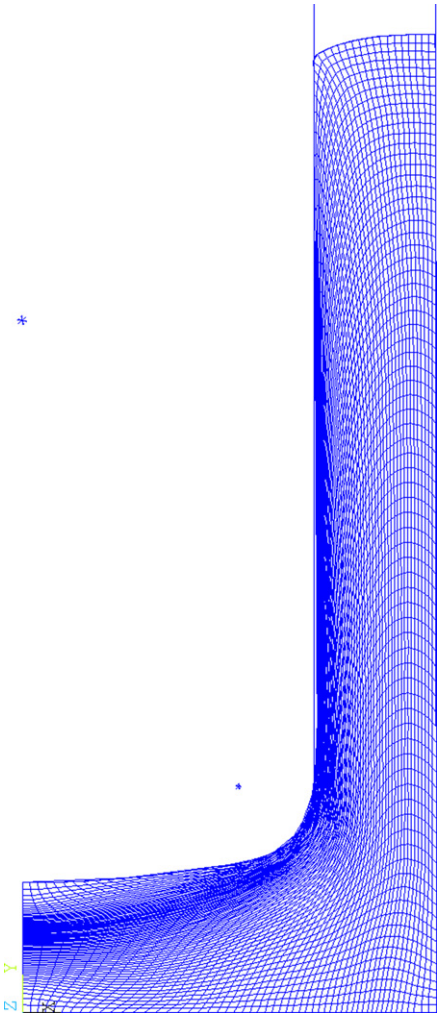

Fig. 10 – The FE mesh in the deformed shape.

errors due to curve fitting were assumed to be negligibly small on the numerically obtained hardness results.

Secondly, up to 8% error may occur in the predictions of the analytical Vickers hardness–effective strain relationship given in Eq. (8) (Sonmez and Demir, 2007). Assuming that the errors in the lowest and highest predicted hardness values are about the same, the error in the objective function value due to the error in hardness prediction is expected to be small.

The effect of friction coefficient was high on the results as discussed before. In the optimization runs, the FE model used a constant friction coefficient equal to 0.1. One should correctly determine the friction coefficient in any real application.

The initial billet material was assumed to be homogeneous and fully annealed, thus the initial hardness distribution in the billet was uniform. In addition, the change in hardness of the deformed workpiece was assumed to be a function of induced effective strain, and independent of strain path.

Die and punch surfaces were modeled using rigid lines. In practice, these surfaces deform elastically, but these deformations are negligible compared to large plastic deformations induced in the workpiece. Otherwise, areas representing the die and the punch would need to be meshed in order to obtain their deformation history, which would highly increase the required computational capacity.

In this study, optimization variables were mainly constrained by geometric requirements. One may integrate new constraints and improve effectiveness of the proposed methodology. Forgability of the product may be monitored by observing induced stresses in the material. Then, a constraint may be set for maximum allowable stress above which internal cracks develop in the workpiece. In addition, a set of constraints may be defined to prevent the die and the punch being exposed to stresses higher than their yield strength. The required press power may also be obtained, and it may be used as a constraint related to economic considerations.

The methodology proposed in this thesis is applicable to other types of materials as well as to other cold forming processes. If a researcher has the required information about the mechanical properties of the material in concern and developed a reliable FE model of the process in concern, he/she can adopt the proposed methodology in his/her own study. In addition to different geometric optimization variables, different variables and constraints may be defined to enlarge the scope of the proposed method. For instance, stresses developed in the tooling may be constrained to prevent tooling failures or improve tooling life. The proposed methodology may also be adapted to multi-stage cold forming processes.

7. Conclusions

In this study, a methodology was proposed to improve the hardness distribution in a backward extruded cup by optimizing preform and die shapes. Two variables defining the punch tip geometry ($R1$ and $R2$) and one variable defining the clearance between the die and the workpiece (t) were selected as optimization variables. The objective function was expressed as the sum of variations from the average hardness. The ranges of values that optimization variables could take were constrained because of the limitations related to the process, product requirements, and finite element analysis. Penalty functions were utilized in order to account for constraint violations. Nelder-Mead was selected as the search algorithm, since it is a robust zero-order algorithm which makes decisions based on the values of the objective function and does not require calculation of any derivatives. An optimization code was developed using ANSYS parametric design language incorporating the finite element model and the optimization procedure.

The optimum values for the optimization variables $R1$, $R2$ and t, were obtained as 31.5, 2.0 and 0.01 mm, respectively. The minimum (best) objective function value obtained during iterations was 9.573, and the maximum (worst) objective function value was 18.007, which refers to a possible improvement of 46.8% in the objective function in comparison to that of arbitrarily chosen values of design parameters.

In addition to the achieved improvements in the hardness distribution, two important design guidelines for an improved hardness distribution were suggested based on the outputs of the optimization procedure. Firstly, the smaller the clearance

between the die walls and the preform, the smaller the variation in the hardness distribution. Secondly, a more effective lubrication results in a more homogeneous hardness distribution. In other words, there is a negative correlation between the friction coefficient for the die-workpiece interface and the degree of variation in the hardness distribution.

Acknowledgments

This paper is based on the work supported by the Scientific Research Projects of Bogazici University with the code number 08A602.

REFERENCES

Altan, T., Ngaile, G., Shen, G., 2005. Cold and Hot Forging—Fundamentals and Applications. ASM International, Ohio.

Celano, G., Fichera, S., Fratini, L., Micari, F., 2001. Application of AI techniques for the optimal design of multipass cold drawing processes. Journal of Material Processing Technology 113, 680–685.

Demir, A., Sonmez, F.O., 2004. Prediction of Brinell hardness distribution in cold formed parts. Transactions of ASME. Journal of Engineering Materials Technology 126, 398–405.

Gao, Z., Grandhi, R.V., 2000. Microstructure optimization in design of forging processes. International Journal of Machine Tools and Manufacture 40, 691–711.

Gouveia, B.P.P.A., Rodrigues, J.M.C., Bay, N., Martins, N.P.A.F., 1999. Finite element modeling of cold forward extrusion. Journal of Materials Processing Technology 94, 85–93.

Hur, K.D., Choi, Y., Yeo, H.T., 2003. A design method for cold forward extrusion using FE analysis. Finite Elements in Analysis and Design 40, 173–185.

Mathews, J.H., 1999. Numerical Methods Using Matlab, third ed. Prentice Hall.

Petruska, J., Janicek, L., 2003. On the evaluation of strain inhomogeneity by hardness measurement of formed products. Journal of Materials Processing Technology 143–144, 300–305.

Roy, S., Ghosh, S., Shivpuri, R., 1997. A new approach to optimal design of multi-stage metal forming processes with micro genetic algorithms. International Journal of Machine Tools and Manufacture 37 (1), 29–44.

Shi, X., Chen, J., Peng, Y., Ruan, X., 2004. A new approach of die shape optimization for sheet metal forming processes. Journal of Material Processing Technology 152, 35–42.

Sonmez, F.O., Demir, A., 2007. Analytical relations between hardness and strain for cold formed parts. Journal of Materials Processing Technology 186, 163–173.

Tabor, D., 1948. A simple theory of static and dynamic hardness. Proceedings of the Royal Society of London, Series A, Mathematical and Physical Sciences 192/1029, 247–274.

Thiyagarajan, N., Grandhi, R.V., 2005. Multilevel design process for 3D preform shape optimization in metal forming. Journal of Materials Processing Technology 170, 421–429.

Zhao, X., Zhao, G., Wang, G., Wang, T., 2002. Preform die shape design for uniformity of deformation in forging based on preform sensitivity analysis. Journal of Material Processing Technology. 128, 25–32.