M. HAWRYLUK*#, D. WILK-KOŁODZIEJCZYK**, K. REGULSKI**, M. GŁOWACKI**

DEVELOPMENT OF AN APPROXIMATION MODEL OF SELECTED PROPERTIES OF MODEL MATERIALS USED FOR SIMULATIONS OF BULK METAL PLASTIC FORMING PROCESSES USING INDUCTION OF DECISION TREES

The article discusses the development of an approximation model of selected plastic and mechanical properties obtained from compression tests of model materials used in physical modeling. The use of physical modeling with the use of soft model materials such as a synthetic wax branch with various modifiers is a popular tool used as an alternative or verification of numerical modeling of bulk metal forming processes. In order to develop an algorithm to facilitate the choice of material model to simulate the behavior of real-metallic materials used in industrial production processes the induction of decision trees was used. First of all, the Statistica program was used for data mining, which made it possible to determine / find the relationship between the percentage of particular constituents of the model material (base material and modifiers) and yield strength, critical and maximum strain, and provide the opportunity to indicate the most important variables determining the shape of the stress – strain curve. Next, using the induction of decision trees, an approximation model was developed, which allowed to create an algorithm facilitating the selection of individual modifying components. The last stage of the research was verification of the correctness of the developed algorithm. The obtained research results indicate the possibility of using decision tree induction to approximate selected properties of modeling materials simulating the behavior of real materials, thus eliminating the need for costly and time-consuming experiments carried out on metallic material.

Keywords: regression trees induction, properties approximation, physical modelling, soft model materials

1. Introduction

A proper elaboration of an industrial metal forming process requires that numerous tests and trials on the actual material should be performed, which is connected with enormous costs as well as a big investment of time. The most important designing stage is the final verification of the elaborated metal forming process performed on the actual material.

At present, there is a search for methods which, on the one hand, would facilitate the designing of metal forming processes and, on the other hand, would eliminate the experiment on the actual material as a verification tool [20,30]. The search is being conducted in two main directions. One, based on mathematical methods as well as new calculation techniques, makes it possible to construct mathematical models of various metal forming processes and phenomena occurring in the deformed material. Here, one should mention numerical modelling based on FEM, etc. [2,19,20,25,32], as well as a whole spectrum of IT tools [2,21,18,32]. The popularity of numerical modelling in the analysis and design of industrial processes results, on the one hand, from the increasing availability of computers with high computing powers, and, on the other hand, from the increasingly simple use of programs applying such numerical methods. Despite the

unquestionable usability and the inevitable implementation of numerical modelling in the analysis and design of processes, one should remember both the potential and the limitations of this method. In turn, the modern IT technologies constantly provide new methods and tools making it possible to partially replace the costly and time-consuming material experiments with a virtual one. Also, more and more new formalisms of knowledge representation in computer systems are being created and developed, e.g. graph theory, fuzzy logic, artificial neural networks, regression trees and genetic algorithms, thus making it possible to construct expert systems supporting various areas of human activity [28,33]. The basic limitation of a direct use of techniques based on mathematical methods in the designing process is the lack of certainty that the obtained results are correct. This uncertainty can be caused by improper assumptions, an inappropriate model or calculation errors, which make the obtained results more or less correct. Despite the fact that numerical modelling and IT tools significantly reduce the role and scope of the experiment on the actual material, it is the stage which remains the most expensive and time-consuming in the whole designing process [3,26].

As it turns out, an alternative for numerical modelling and IT tool verification can be physical modelling methods, with the

* WROCLAW UNIVERSITY OF SCIENCE AND TECHNOLOGY, LUKASIEWICZA 5, 50-371, WROCŁAW, POLAND
** AGH UNIVERSITY OF SCIENCE AND TECHNOLOGY, DEPARTMENT OF APPLIED COMPUTER SCIENCE AND MODELLING, AL. MICKIEWICZA 30, 30-059 KRAKOW, POLAND

\# Corresponding author: marek.hawryluk@pwr.edu.pl

use of soft modelling materials (based on plasticine and synthetic waxes with modifiers), which are much cheaper and faster, constituting another direction in the development of methods supporting the analysis and design of metal forming processes. This method can be an independent tool in the design and analysis of metal forming, which considers both the shape and properties of the ready product, or it can work in combination with numerical modelling, providing it with the necessary information on the behaviour of the deformed material, the boundary conditions and the structural changes; it can also play the role of a verification tool [17,30]. The physical modelling methods make it possible to shorten the designing time and reduce the costs of the actual experiment, owing to the use of non-metallic soft modelling materials, which, through various modifier additions, provide the possibility to obtain the characteristics of most metals and their alloys. For example, synthetic wax filia very well simulates the behaviour of lead formed at ambient temperature, or typical steel, e.g. C45, undergoing a hot deformation process [20,34]. The literature provides a large number of physical modelling applications in the analysis of specific industrial (mainly volumetric) processes, in which compression is the dominant stress state. Non-metallic soft modelling materials, owing to their unique construction, have found their application in the simulation of forging, extrusion, pressing and upsetting processes [2,4,21]. In turn, it is very rarely that one can come across a study which discusses the tests results of a simulation of processes involving other states of stress. Those infrequent cases include: physical modelling of tube blank rolling [8], longitudinal rolling of metal sheets [24], longitudinal rolling of sections [5] and bending thick metal sheets [21], or helical rolling of tubes [23]. The selection of the modelling material also strongly depends on the simulated phenomenon, while being less dependent on the measurement method. For example, in a simulation of cracking, paraffin wax is used as the modelling material, which very well reveals the surface micro- and macro-cracks [4,16]. In turn, when the experiment requires significant plastic deformations, a ductile modelling material should be applied [1], e.g. plasticine. By modifying the composition of modelling materials based on plasticine and synthetic wax filia (through additions of kaolin, lanolin, paraffin, chalk, etc.) and changing the deformation rate and temperature, one can obtain models of flow stress-strain curves for different actual materials [15]. It is assumed that, if the shape of the work-hardening curve for the modelling material is close to the shape of the curve for the given metallic material, it means that the plastic similarity condition has been fulfilled, guaranteeing a proper representation of the physical modelling result in the industrial process. Usually, such a selection of particular modifying components added to the base material (plasticine, synthetic wax filia) has been made based on the many years of experience and knowledge of the researcher performing the studies. Despite the fact that the general rules and effect of the operation of particular modifiers are well-known, the available literature hardly provides studies on the use of even selected IT tools which assist in such a choice in a more standardized

way, not based on the researcher's experience. It turns out that, for the approximation of mechanical properties and thus the prediction of the stress-deformation curves of modelling materials with modifiers, IT tools can be suitable, especially decision tree induction [3,10,11,14,29]. And so, conducting research and studies concerning the use of such type of IT tools to support the selection and prediction of the shape of stress-deformation curves seems to be fully justified.

The aim of the study is to develop an appropriate approximation model of the properties of modelling materials enabling the support of the decisions made in the selection of modifiers for the base material with the use of decision tree induction.

This should contribute to shortening the time of modifier selection through time-consuming physical modelling experiments. In turn, supporting the analysis of the modelling material properties will enable a faster and better matching of the actual metallic materials (steel, and its alloys, aluminum, titanium, etc.) with the work-hardening curves. And so, the application of physical modelling results will become more efficient and more frequently used, as a tool which is more reliable and which better reflects the reality than the virtual computer simulations.

1.1. Characteristics of modelling materials

The commonly used plasticines and filia-based waxes characterize in low Young modulus, high elastic deformation, high sensitivity to deformation and temperature and, usually, deformation weakening. Such materials exhibit not only the desired elastic properties, but they are also suitable for the modelling of hot physical deformation of actual metals. The modelling of cold metal forming process is much more difficult. While it is already possible to produce materials with work hardening, they are still very sensitive to the deformation rate and temperature, which makes physical modelling difficult [21,12,13,36]. Modelling materials based on plasticines exhibit higher structural heterogeneity, and so they are used mainly in the qualitative evaluation of the examined processes, especially the material flow images. In turn, modelling material based on waxes characterize in a lower degree of structural heterogeneity, thus exhibiting more stable properties. Because of this, they are applied in tests of the force parameters of the analyzed processes [30,36].

The possibility to transform the physical modelling results into the industrial processes is determined by the preservation of the similarity conditions, mainly in the plastic, elastic and geometrical scope, as well as the thermal and dynamic friction conditions [30]. In practice, the selection of the modelling material is determined by the modelled process – through the selection of those similarity conditions which are the most crucial from the perspective of the process, as ideal preservation of all the conditions is impossible. In plastic processing, such a condition is the material's similarity condition in the plastic scope. It is fulfilled when the modelling material during deformation behaves in the same way as the actual material. The work-hardening curve is a graphic representation of the material's behaviour during de-

formation. And so, the selection of the modelling material can be based on the criterion of the shape of the work-hardening curve, which should be as close to that of the actual material as possible [21,30]. A well-known plastic similarity criterion is the work-hardening curve model according to Hollomon [9] and the yield stress-deformation curve model by Alder and Phillips [7]. In their studies [21,35], the authors also discuss a new approach to the evaluation of the material's similarity in the plastic scope, in which a quantitative evaluation of the matching degree of the shape of the flow stress-strain curves of the modelling material and the actual material is assumed. These papers also demonstrate that the degree of similarity of the modelling material to the actual one determines the succeeding experiment results and the value of error of the elaborated model. In the case of modelling materials, in order to determine the work-hardening curves, the upsetting test is usually performed, while for actual materials, both the upsetting and tensile tests are conducted. The preparation of the appropriate modelling material is not difficult, and the literature provides information on several technologies of modelling material preparation. In fact, the latter are very similar, while requiring access to appropriate devices, usually found in special physical modelling laboratories [21].

Presented below are exemplary results of the use of a modelling material data base for the physical modelling of operation II of extrusion in a multi-operational process of producing a constant velocity joint boot (CVJB). Figure 1 shows the work-hardening curves for the base material – filia with an addition of paraffin and lanolin, as well as their reference to the actual material – steel UC1, used for CVJB forgings. Table 1 presents the values of coefficients t and C and the chemical composition

of the modelling materials applied for the modelling of the extrusion process of a CVJB made of steel UC1, described in detail in the studies [34]. This new description of the condition of plastic similarity has been repeatedly verified and a very good agreement. Two parameters are used in the proposed description:
– a scale coefficient (factor)

$$C = \frac{1}{k} \sum_{i=0}^{k} \frac{\sigma_i^{rz}}{\sigma_i^{m}} \qquad (1)$$

– and a similarity coefficient (factor)

$$t = \frac{\sum_{i=1}^{k} \dfrac{\left| \sigma_i^{rz} - C\sigma_i^{m} \right|}{\sigma_i^{rz} + C\sigma_i^{m}}}{2} \qquad (2)$$

where:
σ_i^{rz} – the flow stress of a "real" material at the point i
σ_i^{m} – the flow stress of a model material at the point i
k – the number of points on the flow curves of the model and "real" materials, for which the similarity coefficient is determined.

According to such a quantitative evaluation of the matching degree of the shape of the yield stress-deformation curve of the modelling material and the actual one, the closer to 0 the value of coefficient t, the better matching of the modelling material to the actual one. In a theoretical case of an ideal matching of both curves, this coefficient should equal zero [21].

Based on the presented results, one can see that the lowest similarity coefficient value t was obtained for filia, which means that the work-hardening curve for this material is the best matched to the flow curve of UC1 of all the preliminarily selected materials.

Fig. 2a shows a comparison of a forging after the second operation obtained from physical modelling (filia wax) and a forging from the multi-operational process of forging a CVJB (steel UC1). Fig. 2b presents the tools used in the physical modelling and the samples (preforms) made of filia (Fig. 2c). For which, on the surface of symmetry, straight horizontal lines of a different

TABLE 1

Values of parameters *t* and *C* and chemical composition of modelling materials used for lead deformation modelling

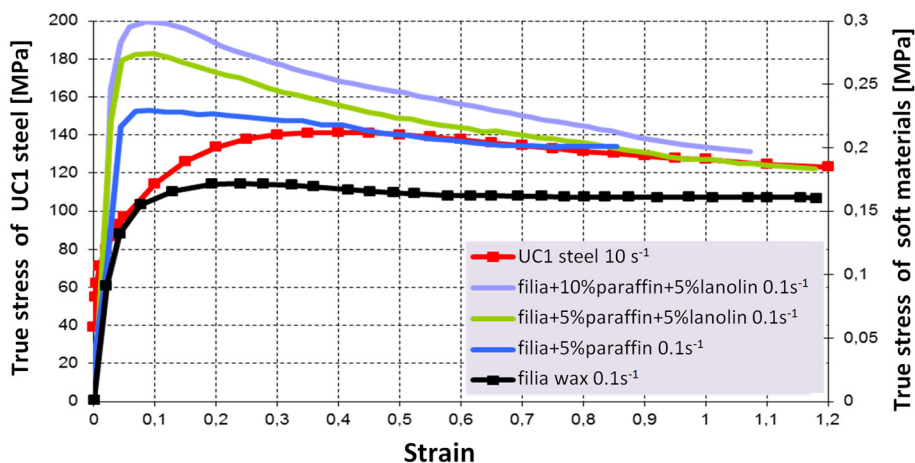| Chemical composition	t	C
filia + 5% paraffin + 5 % lanolin	0.06	269.3
filia + 10 % paraffin + 5 % lanolin	0.07	250
filia + 5% paraffin	0.26	285
Filia	0.023	366.7

Fig. 1. Courses of yield stress in the function of deformation of selected wax mixtures and steel UC1
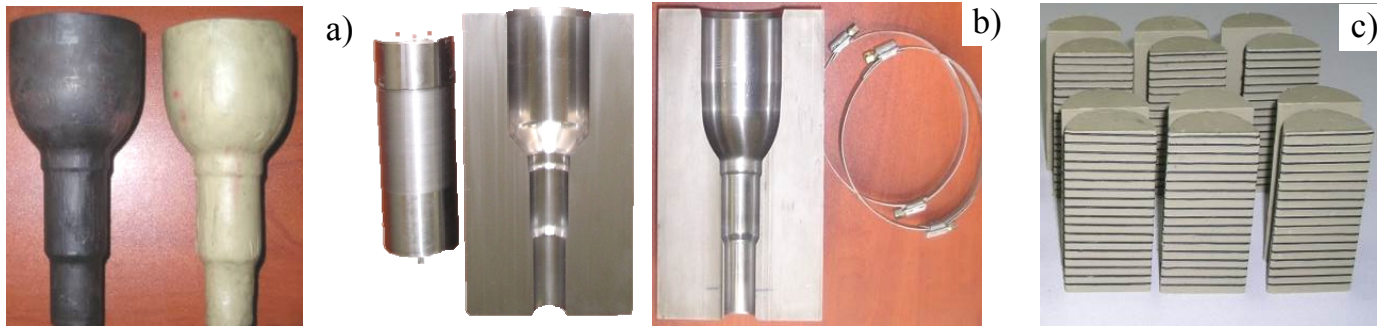
Fig. 2. a) Comparison of a forging made of steel after II extrusion operation with a forging made of filia wax, b) a tool set: a punch, two matrix halves, tightening belts, c) preforms made of modelling materials with horizontal flow lines

colour were plotted at the distance of 5 mm from each other, in order to provide the possibility to perform an evaluation of the manner of material flow based on their deformation.

1.2. Decision trees

Data mining techniques is a very broad concept. It includes statistical tools [6], but also algorithms of rule induction with the use of decision trees (CART, CHAID and others) [10,14,29] and rough sets or fuzzy logic, as well as artificial neural networks, a support vector machine SVM and classifiers such as k nearest neighbours (kNN) or the Bayesian classifier [27]. All these methods have been repeatedly verified and applied in various industrial areas. In the discussed problem, only some of the mentioned tools were used, of which the most important ones are decision trees and regression trees (CART).

The CART algorithm is one of many algorithms of decision tree induction. It is based on hierarchic, binary divisions of a data set for a better segregation of the cases being representatives of the dependent variable value. The algorithm aims at an ideal situation, when the created partition (leaf) includes cases of the same dependent variable value. CART is a universal algorithm in respect of the type of dependent variable – for quantitative variables, regression trees are constructed, in which the division criterion is based on the quantitative feature variance, whereas for discrete dependent (qualitative) variables, a classification tree is built on the basis of the selected node purity index (Gini index, G2 – maximum likelihood statistical significance or Chi^2) [3,11]. These methods are, however, not as effective in prediction as neural networks or the support vector machine. They do not achieve as good results, mainly because of the fact of discretization of the quantitative variables and thus the enforced generalization. Decision trees are a graphic representation of the rules obtained based on the data structure analysis. However, these algorithms enable not only the creation of rules, but also the determination of the significance of the particular variables in the model, which is sometimes as important as the model itself. A variable is described as significant in the classification process, i.e. one which provides information on the class, depending on its readiness for the participation in the dependent variable divisions, which is measured during the construction of

the tree. The established significance makes it possible to create a ranking of independent variables in respect of the effect on the dependent variable. The significance is the degree of covariance with the dependent variable.

The undoubted advantages of classifiers based on trees are: their graphic representation, legible and easy to interpret and verify based on the domain knowledge; a possibility to determine the significance of predictors; insensitiveness to noise and outliers; the result in the form of a set of rules possible to use in other applications.

A decision tree is a graphic method of supporting a decision process; it is a tree-like structure, in which the internal nodes contain tests on the attribute values, and the leaves describe the decisions about the classification of objects. A decision tree is a graphic encasement of a series of conditional statements. Decision trees constitute an advanced form of knowledge representation, which provides a wide range of interpretation possibilities, both at the stage of knowledge acquirement itself (data mining) and in the phase of its application in the decision process. The aim of the studies is finding a relation between the particular components of the modelling material and the selected mechanical properties [9,30].

Each internal vertex of the tree contains the so-called separation point, i.e. a test on the predictor variable, which divides the data set into partitions. The division of nodes in decision trees, as in the discussed case, takes place based on the least square criterion (LSD – Least Significant Difference).

$$R(t) = \frac{1}{N_w(t)} \sum_{i=1}^{} w_i f_i \left(y_i - \overline{y}(t) \right)^2 \qquad (3)$$

where:
$N_w(t)$ – weighted number of cases in the node t,
w_i – value of weighting variable for case i,
f_i – value of frequency variable,
y_i – value of response variable,
$y(t)$ – is the weighted average in the node t.

The constructed and selected tree makes it possible to create the rules. The tree interpretation is direct: for each leaf (conclusion), we track all the consecutive branches (graph arc). Each encountered vertex represents a test, thus being a basis for the creation of a rule premise.

2. Test methodology

In order to achieve the objective of the study, the research was divided into 4 following stages.

2.1. Elaboration of modelling material data base

In order to construct a modelling material data base, the authors, based on the research procedures, prepared samples of base materials with modifiers, in the form of cast cylinders, which underwent upsetting tests performed on a specially constructed press for physical modelling experiments. The work-hardening curves of the modelling materials were determined by means of upsetting tests on made on cylinder samples, 60 mm high and 60 mm in diameter. The basic tests were performed at the temperature of 22°C and at the deformation rate of 0.01 s^{-1}. Technical vaseline was used as the lubricant, placed in specially prepared openings. The tests were conducted on a specially constructed press, whose power device is the motoreducer produced by Lenze, power 1 kW and output torque 15 Nm. The measurement system is equipped with a computer and an inverter with a 16-bit measurement chart and Hottinger amplifiers together with an application written in the Labview environment, which enables control, measurement and archivisation of data. The force was measured by means of the sensor ZEWPN with the range of 0-5 kN, and the displacement – with the use of an induction sensor with the range of 0-200 mm. Based on the force measurement and the displacement, the yield stress-deformation flow curve diagrams were elaborated for the selected material compositions. For a given material composition, a minimum of 3 repetitions were made, from which the averaged run was selected [21].

2.2. Statistical analysis of work-hardening curve results for modelling materials

The creation of a precise model is always preceded by collecting data. Experimental data is always the basis for the process of machine learning, regardless of the assumed data mining technique. Decision trees, neural networks, the support vector machine, Bayesian classifiers, kNN and the discriminant analysis – all these tools are based on training data, which, in the analyzed case, are the results collected from laboratory experiments. The preliminary tests confirmed that one of the most available IT tools frequently applied in related areas is decision tree induction, owing to its not very extensive scope of data and a small size of data sets. Artificial neural networks, also popular in the field of approximation, require much larger volumes of training vectors, especially when a larger set of input parameters is examined, that is, a larger number of independent variables is taken into consideration. It should also be emphasized that, from the perspective of acquiring knowledge of the relations present within the scope of a given phenomenon, neural networks are useless, as they operate according to the black box rule – they are capable of predicting, yet, to a human, this knowledge is unavailable, as opposed to the rules obtained by means of trees, which can be used in a universal manner, independent of the model itself.

2.3. Construction of a model of approximating the values of modelling material mechanical properties based on decision trees followed by an algorithm of base material modifier selection

Another stage of the studies of the decision support tool based on decision tree induction in the scope of selecting the proper material will be a construction of the appropriate model capable of approximating the mechanical property values depending on the additions. This stage is scheduled to consider the three main properties of modelling materials: $\sigma_{\max}$ – maximal yield stress, ε_{kr} – critical deformation, ε_{gr} – limit deformation, which are the ones which mostly determine the shape of the yield stress-deformation curve, depending on the three basic base material modifiers: kaolin, paraffin and lanolin.

The property approximation model aims at answering the questions: In which way do the mechanical properties of the base modelling material – synthetic filia wax – change after the introduction of a given amount of a specific modifier? Is it possible to use the given model to quantitatively determine the mechanical properties of the modelling material by pointing to the percentage composition of the additions? The role of the approximation model is to replace the mathematical function when the determination of the function form seems impossible.

Based on a properly constructed model, it will be possible to create an algorithm capable of selecting the additions (base material modifiers) in such a way so that the most similar modelling material to the given actual material can be created, when the modelling material representation is based on the approximation model.

2.4. Verification of the elaborated algorithm with the use of decision tree induction – physical modelling experiment in an upsetting test

An additionally performed laboratory experiment, consisting in physical modelling, will be a way to validate the results obtained by means of the elaborated model. The rules obtained from the decision tree induction algorithm will enable the determination of the desired material composition – establishing the addition content. Under laboratory conditions, the planned modelling material with modifiers was created and its properties were examined, which made it possible to determine the quality of the approximation model.

3. Discussion of results

3.1. Modelling material data base

Fig. 3 shows exemplary results obtained in the upsetting test on the effect of modifiers: kaolin and kaolin with lanolin, on the level of stress and deformation for the base material – filia wax. The yield stress for filia, with the deformation rate of 0,01 s^{-1}, after the maximal values is reached with the deformation of 0,35, slightly lowers. Adding kaolin to filia causes an increase of its reinforcement, which grows together with the increase of the kaolin content.

For mixtures containing over 8% kaolin, an increase of stress was obtained in the whole deformation scope. Such curves can be used to model the deformation process of actual materials under the conditions of cold metal forming. In turn, in the case of adding only lanolin to pure filia, the obtained curves exhibit a significant decrease of yield stress after it reaches its maximal value. This decrease intensifies with the decrease of strain rate, as well as for filia with paraffin. But addition of paraffin caused simultaneously increase of σ_{max}. The lanolin content increase also causes a reduction of critical deformations (Fig. 4a). Such mixtures can be used for the modelling of hot metal forming processes. A similar effect on the shape of the flow stress-strain curves for filia is demonstrated by paraffin, the difference being that the latter causes a significant increase of the stress level (Fig. 4b).

In the performed research, the effect of the deformation rate on the shape of the flow stress-strain curves of the examined mixtures was also determined. It can be inferred from the courses presented in Fig. 5 that a change in the deformation rate from 0,01 to 0,1 s–1 does not cause a change in the character of the work-hardening curves of the examined mixtures, while causing quite a significant increase of the yield stress. In turn, Table 2 presents compilatory results of the elaborated data base, which includes the maximal stress, the critical deformation and the limit deformation.

From the point of view of the behaviour of modelling materials, σ_{max} describes the maximal recorded force value in the upsetting test, ε_{kr} describes the deformation value after which a weakening of a given material composition was observed (no given value in this column means that, for the given modelling material, the critical deformation was not reached), ε_{gr} – the limit deformations were determined at the moment of the occurrence of cracks on the surface of the side surface of the cylinder).

The investigations performed on the modelling materials showed that by modifying their composition (with additions of kaolin, lanolin, paraffin) and changing the deformation rate and temperature, one can obtain models of flow stress-strain curves for different actual materials. For example, an addition of kaolin causes a clear increase of the yield stress level and a simultaneous increase of yield point. Increasing the content of these additions, e.g. kaolin in the amount over 20%, lowers the limit deforma-
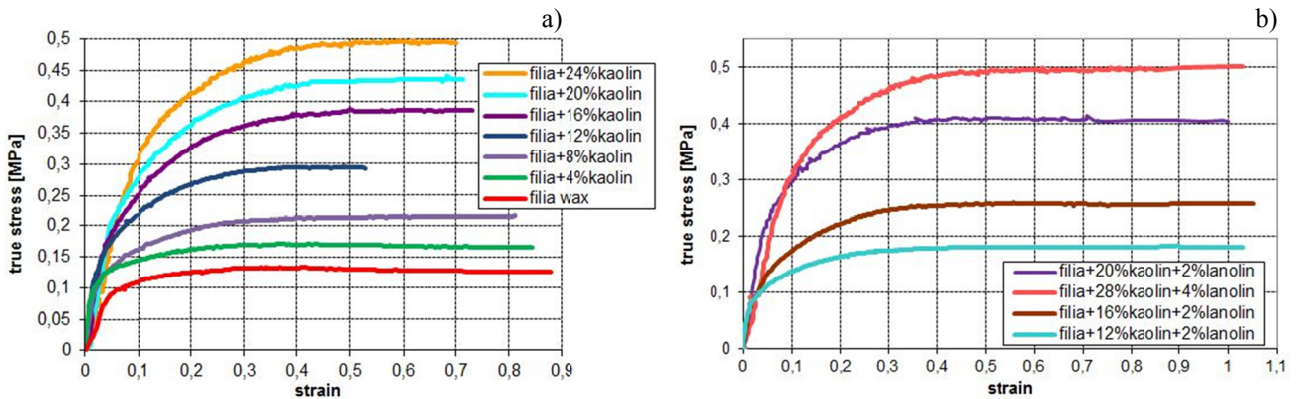

Fig. 3. The flow stress-strain curves for filia with: a) different kaolin contents, b) kaolin and lanolin
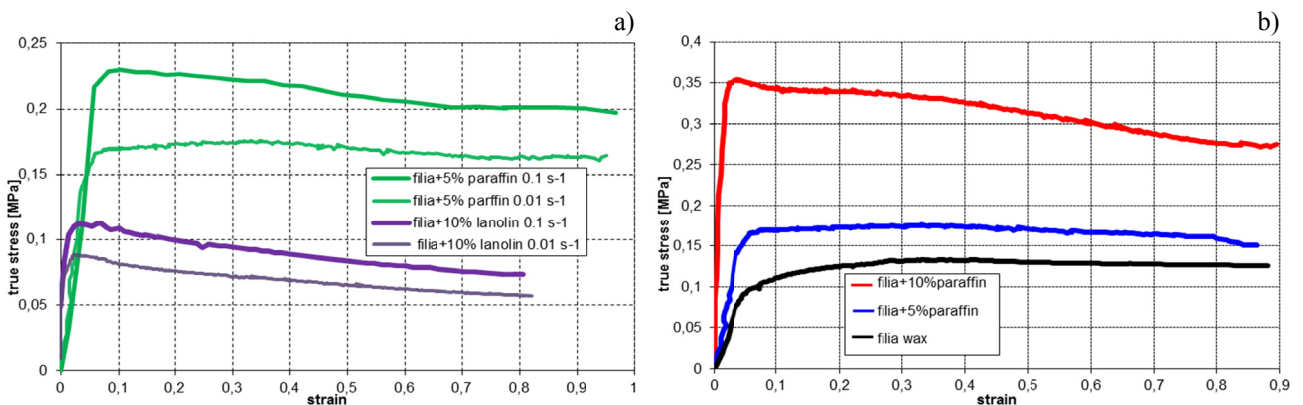

Fig. 4. The flow stress-strain curves for filia: a) with different lanolin contents, b) with different paraffin contents
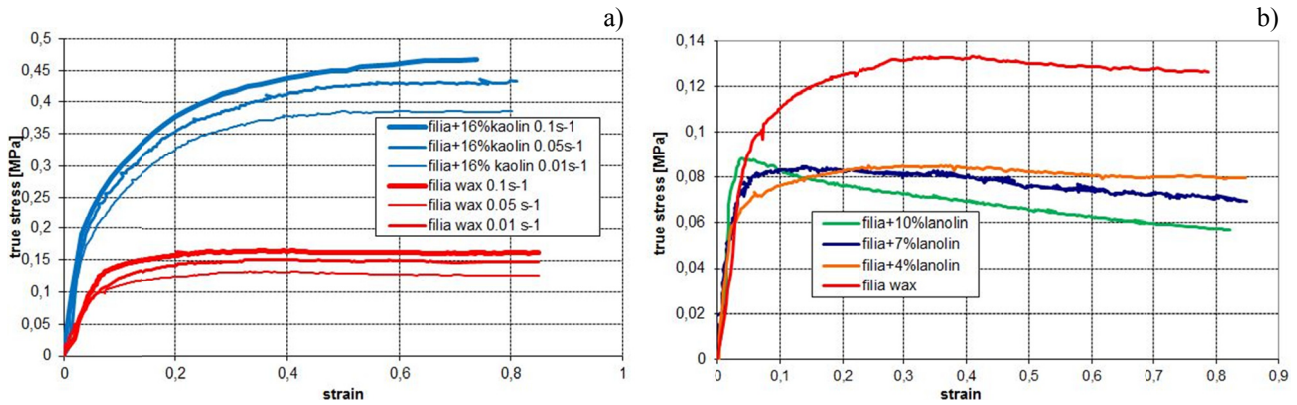
Fig. 5. Effect of a strain rate change on the work-hardening curves: a) for filia and filia with 16 % kaolin, b) for filia with 4-10% lanolin

TABLE 2

Properties of filia-based modelling materials – summary results [36]

Base material	Modifier	Test conditions		Properties		
		T [°C]	$\dot{\varepsilon}$ [s⁻¹]	σ_{max}[MPa]	ε_{kr}	ε_{gr}
filia	kaolin 2-4 %	22	0.01	0.128-0.136	0.38-0.39	1-1.05
filia	kaolin 8-24%	22	0.01	0.209-0.507	—	0.8-0.9
filia	kaolin 12-28% + lanolin 2-4%	22	0.01	0.182-0.505	0.91	0.95-1.05
filia	lanolin 4-14%	22	0.01-0.1	0.083-0.095	0.03-0.29	1.2
filia	—	22	0.01-0.1	0.13-0.17	0.3-0.41	0.9-1
filia	kaolin 16%	22	0.01-0.1	0.385-0.47	—	0.8-0.95
filia	paraffin 5-10%	22	0.01-0.1	0.176-0.356	0.03-0.09	0.9-1.05
filia	—	17-22	0.01	0.13-0.18	0.3-0.36	0.9-1.1
filia	kaolin 20%	18-24	0.01	0.41-0.46	—	0.8-0.9

tions, and such a material cracks faster that the actual material. Also, increasing the kaolin content is limited; with the content of this powder filler above 20% of the total weight of the modelling material, we observe percolation, that is tacking of modifier particles. Percolation makes proper mixing impossible, and thus, also, the appropriate distribution of the filler particles in the matrix of the modelling material. It can also affect the intensity of the sedimentation phenomenon, which causes anisotropy of the modelling material. Reducing the percolation phenomenon or obtaining higher limit deformations in the mixtures with a high kaolin content is possible through the introduction of a small amount of lanolin. On the other hand, it is difficult to find materials for which, in the whole deformation scope, an intense increase of yield stress is observed. In turn, an addition of lanolin or paraffin causes the obtained curves to exhibit a strong decrease of the yield stress after it reaches its maximal value. A change of deformation rate by one order of magnitude, e.g. from 0.01 to 0.1 s⁻¹, and of temperature by a few degrees Celsius, does not cause a significant change in the character of the work-hardening curves of the examined mixtures, while only causing a change in the level of the yield stress. The collected results of experimental tests performed on the modelling material (Tab. 2) makes it possible to quantitatively describe the properties of the examined modelling materials: σ_{max} – maximal yield stress value, ε_{kr} – critical deformations, ε_{gr} – limit deformations. These properties can constitute the basis for the selection of the modelling material for any actual material [21].

3.2. Statistical analysis of work-hardening curve test results for modelling materials

In the first place, a statistical analysis will be performed on the results obtained in the upsetting tests for one of the base modelling materials. For the examinations, filia was selected, owing to a larger number of experimental data and the observed higher repeatability and stability of the properties. Next, the material test results were collected and analyzed for each content of this base material with the introduced modifiers. For each set of data, a correlation matrix was constructed, based on which a preliminary evaluation was performed. Of those two base materials, for the further examinations, the one was selected for which the results were properly verified based on the experience of a scientist working with physical modelling with the use of soft modelling materials.

The performed preliminary statistical analysis of the obtained results shows that, in order to teach the approximation model (in the succeeding stage of investigations) how the properties of the modelling material are formed, one should methodically approach the study of the effect and amount of the particular modifying additions. And so, the particular modifiers were introduced and considered in the elaborated model with the selected step, where their percentage composition was successively changed. In the case of unsatisfactory results, caused e.g. by the lack of data for the given scope, additional investigations are planned in the area of physical modelling and supplementa-

tion of the data base for the missing compositions of the modelling materials. The considerations of the effect of the particular additions on the properties of the modelling material were based on the data referring to the laboratory plastometric tests for filia, of which a fragmentary compilation is presented in Table 3. As it was mentioned before, this material was selected for the tests owing to the stability of its properties under the conditions of a changing composition.

104 samples for different filia contents were laboratory tested. Such a number is high enough to point to certain existing relations. A good way of starting the tests is generating the matrix of the scatter diagram, presented in Fig. 6.

acteristic – histogram ε_{kr} – of the critical deformation pointing to a strong disproportion in the existing values, which results from the material characteristics. It can also be seen that the set of samples has a weak representation in the scope of elevated lanolin contents, which makes it difficult to draw conclusions on the effect of this component.

The correlation matrix explicitly demonstrates those components which significantly affect the properties (Tab. 4).

TABLE 3

Exemplary test results of modelling material properties: filia with constant test conditions: $T = 22°C$; $\varepsilon = 0.01$ s^{-1}

Modifier content [%]:			Properties		
Kaolin	Lanolin	Paraffin	$\sigma_{\max}$	ε_{kr}	ε_{gr}
0	0	0	0.13	0.35	0.7
2	0	0	0.17	0.71	0.71
4	0	0	0.22	0.72	0.72
6	0	0	0.25	0.73	0.73
20	4	2.5	0.55	0.042	0.88
20	4	5	0.58	0.043	0.93
20	4	7.5	0.6	0.046	0.96
20	4	10	0.65	0.5	0.97

TABLE 4

Correlation matrix of modelling material components and properties

		Kaolin	Lanolin	Paraffin
$\sigma_{\max}$		0.69	−0.26	0.56
		$p = 0.000$	$p = 0.008$	$p = 0.000$
ε_{kr}		0.34	−0.08	−0.38
		$p = 0.000$	$p = 0.407$	$p = 0.000$
ε_{gr}		0.30	0.18	0.58
		$p = 0.002$	$p = 0.063$	$p = 0.000$

Fig. 7. Compilation of 3-D scatter diagrams showing the total effect of the selected components on the particular properties
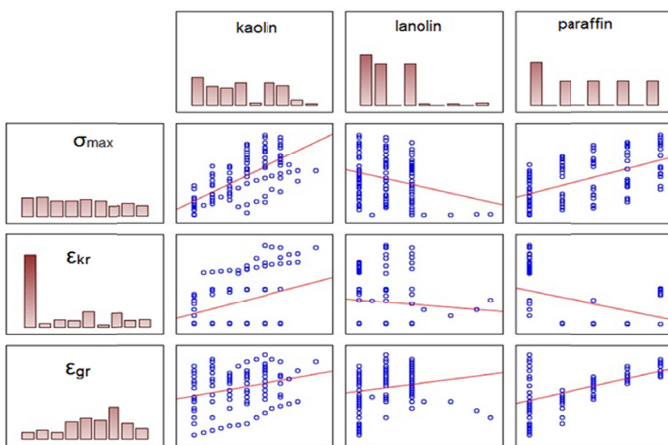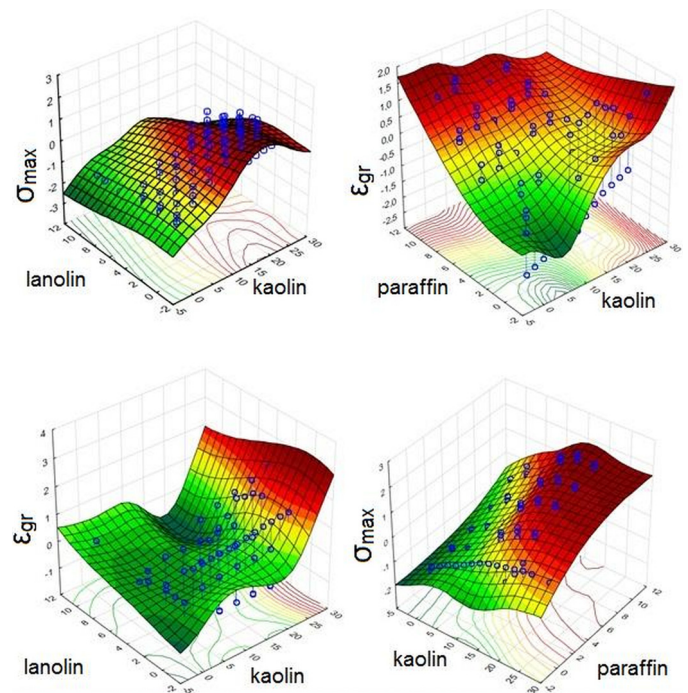
Fig. 6. Scatter diagram matrix for filia components and mechanical properties: $\sigma_{\max}$ – maximal yield stress value, ε_{kr} – critical deformation, ε_{gr} – limit deformation

These diagrams (Fig. 6) make it possible to establish that there is a strong relation between the kaolin content and $\sigma_{\max}$ – the maximal yield stress value as well as ε_{gr} – limit deformation. Other two pairs of relations between paraffin and $\sigma_{\max}$ and ε_{gr} are also clearly represented. Increasing the content of kaolin and paraffin causes a simultaneous increase of $\sigma_{\max}$ and ε_{gr}. On the diagram, one can also notice another experimental data char-

We can point out (Tab. 4) a statistically significant effect of kaolin and paraffin on the maximal value of yield stress ($\sigma_{\max}$). Also, kaolin raises the value of critical deformation. An elevated paraffin content lowers the critical deformation value (ε_{kr}). The effect of lanolin should not be considered due to the lack of statistical significance.

With the aim of a simultaneous examination of different effects of the particular components on the properties, one can create a compilation of 3-D diagrams, shown in Fig. 7.

3.3. Construction of a model of approximating the values of modelling material mechanical properties based on decision trees followed by an algorithm of base material modifier selection

In the case of the discussed data, the effect of the percentage compositions of the components on the particular properties was examined. The CART algorithm was used to generate three regression trees, one for each dependent variable. Fig. 8 shows the arrangements of these trees for each dependent variable. The tree structure graphically presents the division of the training set into value classes, depending on the test functions, which are the value scopes of explanatory variables.

The tree arrangements confirm the significant feature of the CART algorithm — it generates binary trees, owing to which the division into partitions according to one explanatory variable (e.g. paraffin content) can be repeated with the use of another test function value. The more expanded the tree, the bigger the number of rules, and the higher the number of division levels in the tree, the bigger the number of premises in each rule. Each leaf (final node) denotes one decision rule. The tree structure visualizes the knowledge base for the model of parameter value approximation. The tree for σ_{max} has 16 final nodes, which means that, on its basis, it is possible to generate 16 decision rules pointing to the expected value σ_{max} depending on the chemical composition of the modelling material. The tree for ε_{kr} generates 10, and the tree for ε_{gr} – 13 rules. In total, the generated knowledge base contains 39 decision rules. Each final node characterizes in a mean value and a variance value of the dependent variable. The quality of the rules can be determined with the use of the variance value – the lower the variance, the more reliable the rule. Unfortunately, the variance value is not an absolute value, which means that the rule quality can be measured in this way only within one tree.

Table 5 presents the results in the particular nodes for each tree. The diagrams of these characteristics, shown in Fig. 9, enable the selection of the nodes depending on the expected values of dependent variable, that is a specific plastic property. Knowing the expected value, we select the node number which fulfills the specific conditions, and next, based on the tree, we read out the rules which enable the adaptation of the factor composition.

TABLE 5

Characteristics of final nodes in regression trees for each variable

	σ_{max}			ε_{kr}			ε_{gr}	
	Mean	Variance		Mean	Variance		Mean	Variance
8	0.09	0.0008	4	0.27	0.0047	8	0.75	0,0016
9	0.15	0.0002	6	0.73	0.0007	9	0.83	0,0004
12	0.26	0.0017	8	0.82	0.0005	7	0.89	0,0005
13	0.20	0.0010	9	0.96	0.0051	12	0.84	0,0010
11	0.29	0.0013	12	0.07	0.0090	13	0.90	0,0005
14	0.28	0.0039	18	0.03	0.000005	16	1.02	0,0010
16	0.38	0.0006	19	0.03	0.000002	17	0.95	0,0024
17	0.50	0.0013	17	0.04	0.000001	18	0.87	0,0002
20	0.27	0.0128	15	0.04	0.000006	19	0.94	0,0001
22	0.39	0.0049	11	0.45	0.0109	22	0.95	0,0004
23	0.49	0.0026				23	0.97	0.0003
26	0.55	0.0040				24	1.01	0.0004
27	0.64	0.0050				25	0.98	0.0002
30	0.72	0.0018						
31	0.79	0.0012						
29	0.65	0.0063						

For example: Let us assume that we want to obtain the highest maximal value of yield stress – and so, we select node no. 31, for which the mean σ_{max} is the highest. We read out the composition from the tree: kaolin >11%; paraffin >8.75%; lanolin < = 3% (the tree fragment which makes it possible to read out such a rule was shown in Fig. 9). If we are interested in the high-
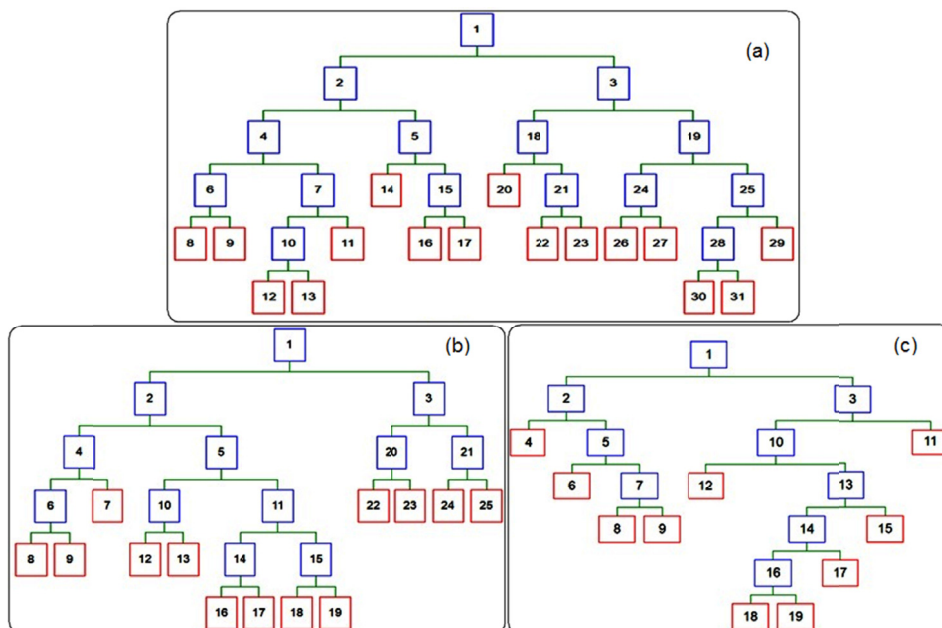

Fig. 8. Arrangement of decision trees for particular dependent variables: a) σ_{max}, b) ε_{gr}, c) ε_{kr}

est ε_{gr} – limit deformation, we choose node 16 in the third tree. From the tree, we read out the composition: kaolin >9%; 1.25% > paraffin > 6.25%; 0.5% < lanolin < = 3%. Unfortunately, the conditions for paraffin are contradictory to the rules from the first tree. And so, we can try to select another node – the closest one in respect of the mean value is node 24 (the difference between them being only 0.01). In the case of node 24, the conditions are as follows: kaolin >6%; paraffin >8.75%. This means that there is no contradiction to the rules for the tree σ_{max}. Also, the user can use the tree model in order to determine the composition of the additions. Each tree leaf represents a class of cases with similar values of the parameter being the dependent variable. The cases within one leaf (class) have the same composition – that is, the content of additions for each case within this class is the same [32].
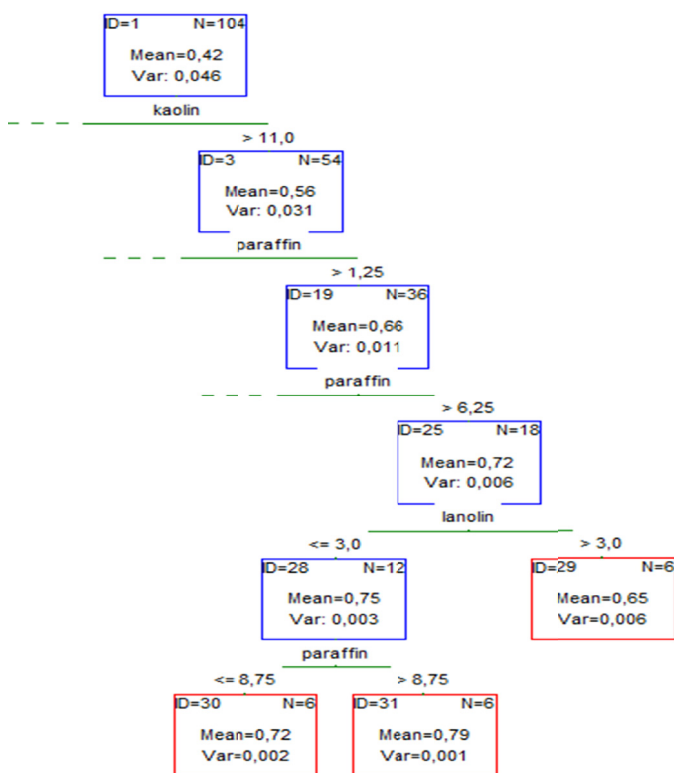

Fig. 9. Regression tree fragment enclosing three out of sixteen leaves for dependent variable σ_{max}

The user can apply the tree to determine the addition composition in order to obtain the desired level of stress or deformation. With the purpose of a simpler selection of the desired level of each parameter, it is possible to apply value visualization for the particular tree nodes. The user can take advantage of the diagram shown in Fig. 10 by selecting the level, and next, verify the rule leading to the desired value. The CART algorithm also makes it possible to qualitatively evaluate the effect of the particular components on the properties. This effect is expressed by the significance, i.e. the readiness for the participation in the divisions of the dependent variable, which is measured during the construction of the tree.

The significance of the particular components for each property has been presented in Fig. 11.
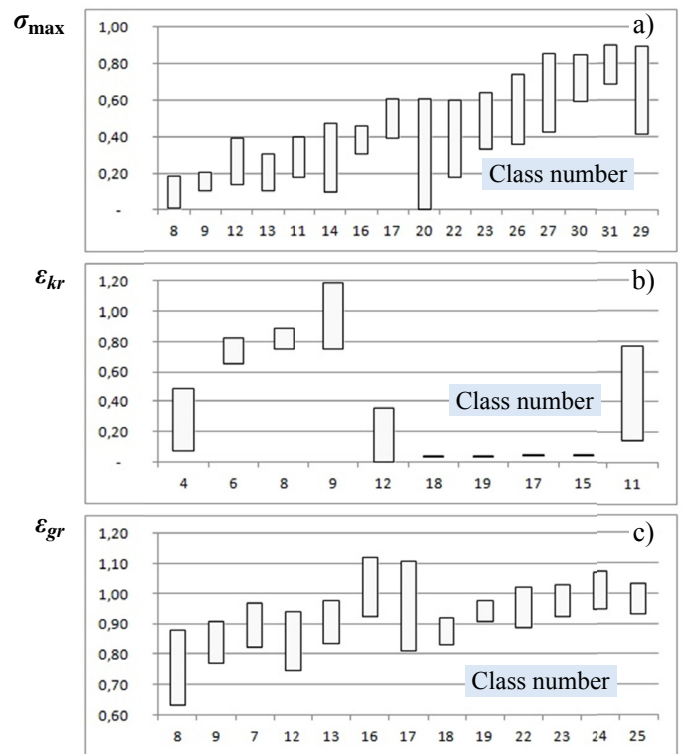

Fig. 10. Characteristics diagrams for final nodes in regression trees for variables: a) σ_{max}, b) ε_{kr}, c) ε_{gr}
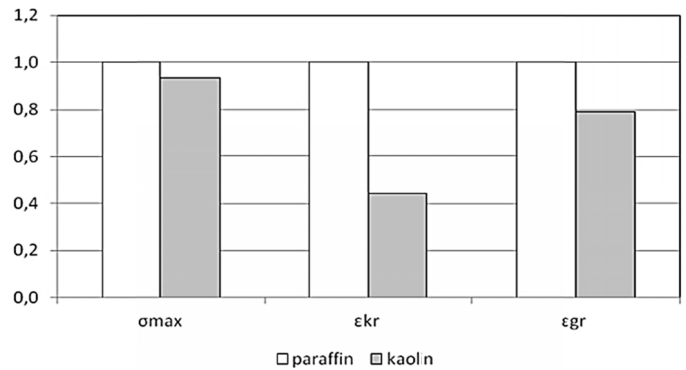

Fig. 11. Significance of particular components for each property

The significance results for lanolin have not been presented because, as it was established before, an insufficient number of measurements for different contents of the same component causes the lack of statistical significance for the correlation with other variables. Expanding the training data set can certainly lead to a change in the form of the trees, thus changing the rank of the particular variables.

3.4. Verification of the elaborated algorithm with the use of decision tree induction – physical modelling experiment in an upsetting test

With the aim to verify the decision tree model, the work-hardening curves were determined for materials with different compositions than those used for the modelling. The comparison

TABLE 6

Comparison of the experimental values of the maximal yield stress value with the values determined by means of a regression tree

Components [%]			σ_{max} [MPa]		ε_{kr}		ε_{gr}	
Kaolin	Lanolin	Paraffin	Empirical	Estimated	Empirical	Estimated	Empirical	Estimated
4	2	7.5	0.41	0.39	0.042	0.044	0.98	0.95
0	0	5	0.18	0.16	0.033	0.071	0.86	0.90
8	4	5	0.28	0.29	0.044	0.042	0.9	0.90
0	0	0	0.13	0.10	0.35	0.279	0.7	0.76
12	4	0	0.1	0.27	0.76	0.739	0.89	0.96
4	4	5	0.24	0.29	0.041	0.042	0.94	0.90
0	1	0	0.082	0.10	0.34	0.279	0.87	0.84
0	0	10	0.36	0.29	0.04	0.456	0.98	1.01
16	0	0	0.43	0.27	0.8	0.820	0.8	0.84
0	10	0	0.088	0.10	0.35	0.279	0.78	0.84

was to concern: σ_{max} – maximal yield stress, ε_{kr} – critical deformation, ε_{gr} – limit deformation examined during the measurement, as well as developed the approximating model.

The results demonstrate that the tree is capable of detecting relations and disregarding the training data. Precision coefficients of models of a regression trees illustrating the fitting of models were presented in table 7.

TABLE 7

Precision coefficients of models of a regression trees

	σ_{max} [MPa]	ε_{kr}	ε_{gr}
Mean absolute error (MAE)	0.057	0.070	0.038
Correlation (r)	0.776	0.882	0.901

We can risk stating that, even in a situation when the tree extrapolated the results beyond the scope, the error remains within the acceptable range at the stage of material design.

Another verification method is the determination, based on the tree, of the rules concerning the addition content in the application for the limit values of the particular properties. These rules were determined on the basis of the diagrams shown in Fig. 8 and by means of the trees for σ_{max}, ε_{kr} and ε_{gr}. The rules in their general form have been presented in Table 8.

TABLE 8

Rules for the selection of additions depending on the desired property level

Kaolin	Lanolin	Paraffin	σ_{max}	ε_{kr}	ε_{gr}
>11	<3	5-9	max	—	—
>13	>1	>1	—	max	—
>9	0.5-3	<1	—	—	max
<3	—	<1	min	—	—
>2	<1	>1	—	min	—
<15	<0.5	<3	—	—	min

Based on the experimental data, it can be stated that these rules are true in their general form, and so, it is possible to verify the correctness of the drawn conclusions. If the user wants to create a material, he/she must determine an exemplary set of properties to be obtained. Next, by means of the rules obtained from the trees, he/she will determine the amount of additions and the estimated mean value for each property.

4. Conclusions

The paper presents the possibilities of applying IT tools, especially decision tree induction used for the creation of an approximation model of the properties of a modelling material, such as filia, applied in physical modelling of industrial metal forming processes. Statistical tools, as well as the CART algorithm implemented in the STATISTICA packet, were applied. The physical modelling methods make it possible to shorten the designing time and reduce the cost of the actual experiment, owing to the use of non-metallic soft modelling materials, which, through the addition of various modifiers, make it possible to obtain characteristics for most metals and their alloys. It is assumed that, if the shape of the work-hardening curve for the modelling material is close to the shape of the curve for the given metallic material, it means that the condition of plastic similarity has been fulfilled, which guarantees a proper representation of the physical modelling results in the selected industrial process. Despite the fact that the general rules and effect of the particular modifiers on the shape of the work-hardening curve are well-known, usually, such a selection of the particular modifying component added to the base material (plasticine, synthetic wax filia) takes place on the basis of the knowledge and experience of the scientist. The presented results of preliminary tests demonstrated that, for this purpose, decision tree induction is an appropriate tool.

The obtained results make it possible to estimate the values of the selected modelling material parameters (σ_{max} – maximal yield stress value, ε_{kr} – critical deformation, ε_{gr} – limit deformation) based on the determined percentage composition of three components: kaolin, paraffin and lanolin. As it has been demonstrated, such properties can constitute the basis for the selection of the modelling material for any actual material. The approximation model of the base material properties was elaborated with

the use of the regression tree induction algorithm. In this way, a set of rules was generated, which enable the determination of the addition content, depending on the expected work-hardening properties. The examinations were performed by means of a scant number of training data, and yet, at the verification stage, an acceptable level of error was achieved for the presented application, which confirms the unique abilities of the CART algorithm in the standard abstraction.

The elaborated knowledge base makes it possible to perform a qualitative and quantitative evaluation of the effect of the particular components on each property, thus supporting the development of IT tools applied in physical modelling, which is either an independent tool or an alternative for numerical modelling verification in the design and analysis of metal forming processes.

REFERENCES

[1] A.E.M. Pertence, P.R. Cetlin. Similarity of ductility between model and real materials, Journal of Material Processing technology **103**, 434-438 (2000).

[2] B.P.P.A. Gouveia, J.M.C. Rodrigues, P.A.F. Martins, T. Bay Physical and numerical simulation of the round-to-square forward extrusion, Journal of Mechanical Working Technology **112**, 244-251 (2001).

[3] D. Wilk-Kołodziejczyk, K. Regulski, G. Gumienny, B. Kacprzyk, S. Kluska-Nawarecka, K. Jaśkowiec, Data mining tools in identifying the components of the microstructure of compacted graphite iron based on the content of alloying elements, International Journal of Advanced Manufacturing Technology **95**, 3127-3139 (2018).

[4] G.C. Gingher, G. Padjen. Hot strip mill edging practices and plasticine modelling, 34th MWSP Conf. Proc., ISS-AIME **30**, 3-12 (1993).

[5] H.W. Shin, et. al. A Study on the rolling of I-section Beams, International Journal of Machine Tools and Manufacture **34** (2), 147-160 (1994).

[6] I. Olejarczyk-Wożeńska, A. Adrian, H. Adrian, B. Mrzygłód. Parametric representation of TTT diagrams of ADI cast iron. Archives of Metallurgy and Materials **57**, 613-617 (2012). DOI: 10.2478/v10172-012-0065-9.

[7] J. Alder, K.A. Phillips. The effect of strain-rate and temperature on the resistance of aluminum, copper and steel to compression, Journal of the Institute Metals **83**, 80-88 (1954).

[8] J. Boucly, J. Oudin, Y. Ravalard. Simulation of ring rolling with new wax-based model material on a flexible experimental machine, Journal of Mechanical Working Technology **16** (2), 119-143 (1988).

[9] J.H. Hollomon. Tensile deformation. Trans. AIME **162**, 268-290 (1945).

[10] J.R. Quinlan. Induction on decision trees. Machine Learning, Kluwer Academic Publishers, Boston (1986).

[11] K. Regulski, D. Szeliga, J. Kusiak. Data exploration approach versus sensitivity analysis for optimization of metal forming processes. Key Engineering Materials **611-612**, 1390-1395 (2014).

[12] K. Swiatkowski, R. Cacko. Investigations of new wax-based model materials simulating metal working process **72** (2), 267-271 (1997).

[13] K. Swiatkowski. Physical modelling of metal working processes using wax-based model materials **72** (2), 272-276 (1997).

[14] L. Breinman, et al. Classification and regression trees. Chapman and Hall, London (1993).

[15] Ł. Wojcik, K. Lis, Z. Pater. Plastometric tests for plasticine as physical modelling material, Open Engineering **6** (1), 653-659 (2016).

[16] M. Arentoft. Prevention of defects in forging by numerical and physical simulation, Technical University of Denmark Institute of Manufacturing Engineering 1996.

[17] M. Arentoft, P. Henningsen, N. Bay, T. Wanheim. Simulations of defects in metal forming, Journal of Mechanical Working Technology **45**, 527-532 (1994).

[18] M. Hawryluk, B. Mrzygłód. A durability analysis of forging tools for different operating conditions with application of a decision support system based on artificial neural networks (ANN). Eksploatacja i Niezawodność - Maintenance and Reliability **19** (3), 338-348 (2017), http://www.ein.org.pl/sites/default/files/2017-03-04p.pdf.

[19] M. Hawryluk, J. Jakubik. Analysis of forging defects for selected industrial die forging processes. Engineering Failure Analysis **59**, 396-409 (2016).

[20] M. Hawryluk, S. Polak, Z. Gronostajski, K. Jaśkiewicz, Application of physical similarity utilizing soft modeling materials and numerical simulations to analyse the plastic flow of UC1 steel and the evolution of forces in a specific multi-operational industrial precision forging process with a constant-velocity joint housing. Experimental Techniques, https://doi.org/10.1007/s40799-018-0288-4.

[21] M. Hawryluk. The influence of the condition of plastic similarity on the accuracy of physical modeling of extrusion processes, PhD thesis, Wroclaw University of Science and Technology, Wrocław 2006.

[22] M. Itoh. Determination of forming limits of thick sheet in compression bending by using model material, The Technical University of Denmark Institute of Manufacturing Engineering, Journal of Mechanical Working Technology **12**, 269-277 (1985).

[23] M. Young-Hoon, et al. Physical modelling of edge rolling in plate mill with plasticine. Steel Research **64** (11), 557-563 (1993).

[24] M.J. Adams, et al. A Two roll mill as a rheometer for pastes, material resume social symposium process, Material Research Society **289**, 237-257 (1989).

[25] O.C. Zienkiewicz. Finite Element Method, McGraw Hill (1977).

[26] P. G. Maropoulos. Review of research in tooling technology, process modelling and process planing. Part II: Process plan-

ing. Computer Integrated Manufacturing Systems **8** (1), 13-20 (1995).

[27] D. Wilk-Kołodziejczyk, K. Regulski, G. Gumienny, Comparative analysis of the properties of the nodular cast iron with carbides and the austempered ductile iron with use of the machine learning and the support vector machine, The International Journal of Advanced Manufacturing Technology **87**, 1077-1093 (2016).

[28] T. Gangopadhyay, D. Kumar D, I. Pratihar. Expert system to predict forging load and axial stress. Applied Soft Computing **11** (1), 744-753 (2014).

[29] T. Hill, P. Lewicki. STATISTICS methods and applications. Stat-Soft Tulsa (2007).

[30] T. Wanheim. Physical modelling of metalprocessing. Proces-technics Institut, Laboratoriet for Mekaniske Materialeprocesser; DanmarksTekniskHojkole, 1988; Danmark.

[31] V. Vazquez, T. Altan. New concepts in die design – physical and computer modelling applications, Journal of Material Processing Technology **98**, 212-223 (2000).

[32] Z. Górny, S. Kluska-Nawarecka, D. Wilk-Kołodziejczyk, K. Regulski. Methodology for the construction of a rule-based knowledge base enabling the selection of appropriate bronze heat treatment parameters using rough sets. Archives of Metallurgy and Materials **60**, 309-315 (2015). DOI: 10.1515/amm-2015-0050.

[33] Z. Gronostajski, et al. The expert system supporting the assessment of the durability of forging tools. The International Journal of Advanced Manufacturing Technology **82** (9), 1973-1991 (2015).

[34] Z. Gronostajski, et. al. Analysis of forging process of constant velocity joint body. Steel Research International 1:547-554 (2008).

[35] Z. Gronostajski, M. Hawryluk. Analysis of metal forming processes by using physical modelling and new plastic similarity condition. 10th ESAFORM Conference on Material Forming AIP Conference Proceedings, ISSN 0094-243X **907**, Zaragoza, Spain 608-613 (2007).

[36] Z. Gronostajski, M. Hawryluk. Materials used in physical modeling, Mechanical Review **7-8**, 42-46 (2005).