Strojniški vestnik - Journal of Mechanical Engineering 57(2011)No, StartPage-EndPage
UDC xxx.yyy.z

Paper received: 8.12.2009
Paper accepted: 18.1.2011

Retrieved Features of 2D and 3D Models from CAD as Decision Support for Production Time/Cost Estimation

Predrag Cosic[1,*] - Drazen Antolic[2]
[1] University of Zagreb, Faculty of Mechanical Engineering and Naval Architecture
[2] AD, Ltd, Zagreb

Very frequently a company has to respond quickly to some important requests for offers, generated for individual or batch production, for example: a great number of requested offers (requested prices) either for production of products at once or for production of small batches (the offers for which are rarely repeated), frequent changes of priorities during production, short deadlines for delivery, market demands for bringing the prices of individual or batch production close to the prices of mass production etc.

Regression analysis as a possible approach to time/cost estimation is used for the estimation of requested results based on the previous stochastic results and experiments. On the other hand, the requests for classification consideration of the product shape and process sequencing are important conditions for designing a general model for the estimation of machining times. In fact, it means development of a technological knowledge base. The purpose of this paper is to establish possible connections between features (2D and 3D) and necessary machining time for manufacturing of products.

Retrieving the required variables (features) from a 2D model involves a certain level of subjectivity concerning the number and size of the variables (manner of presenting dimension lines, number of dimension lines, number of views, tolerance fields). On the other hand, for the retrieving of the required variables from a 3D model a strictly defined procedure in the process of 3D model creation is needed.

Estimation of production time is a necessary basis for cost estimation, cost reduction or TCE (Total Cost Estimation). As result of our analysis, we have created eight regression equations with the obtained index of determination, with the most important independent variables different for 2D and 3D model.

Keywords: stepwise multiple linear regression, group technology, knowledge base, production times

1 INTRODUCTION

An experienced process planner usually makes decisions based on comprehensive data without breaking it down into individual parameters.

So, as the first phase it was necessary to establish technological knowledge base, define features of the 2D drawing (independent variables), possible dependent variables, size and criteria for sample homogenization (principles of group technology) for carrying out analysis of variance and regression analysis.

The second phase in the research was to investigate the possibility for easy automatic, direct finding and retrieving of 3D features of an axial symmetric product to the regression model. The defined request resulted in the development of the procedure for retrieving parameters from the 3D model with a low level of subjectivity, a very fast and reliable process via CAD report to the regression model.

As it can be seen from references there are different approaches for data receiving from CAD (STEP) [1], integration of CAPP, CAD/CAM and business activities [2], development of database system of mechanical components [3] and [4], and integrated product engineering [5] for costs estimation and rapid cost estimation [6], application of neural networks in estimation of machining time [7], connection from CAPP, CAD, CAM; DFx to DFA through product development [8] etc.

The most important characteristic of our approach presented in this paper is estimation of production time using group technology and regression analysis.

[*] Corr. Author's Address: Predrag Cosic, Ivana Lucica 5, 10 000 Zagreb, Croatia, predrag.cosic@fsb.hr

2 DRAWING FEATURES AND TECHNOLOGICAL DATABASE FOR PRODUCTION TIME ESTIMATION

Very frequently (especially in the case of SMEs) it is necessary to respond quickly to some important requests for offers, generated for individual or batch production, for example in the case of:

1) a great number of requested offers for manufacturing of products at once,
2) small batches that are very rarely repeated,
3) frequent changes of priorities during production,
4) short deadlines for delivery,
5) market demands for bringing prices of individual or batch production close to the prices of mass production, etc.

It must be stressed that technological knowledge and speed of process planning are often more important than the technological level of equipment, skills and knowledge of people who realize the technology. So, very often in practice we can be faced with the following:

a) A great amount of time spent in planning of the technological process for a product without any specific agreement being made concerning order for manufacturing of the product,
b) Signing of an agreement without estimated precise machining times/costs necessary for product manufacturing and realization in accordance with contracted production,
c) As one of the important criteria in decision support for selection of the best suppliers during process of product development very useful is fast estimation of production times as the main basis for cost estimation,
d) Looking for most important influences for determination of production times/costs,
e) During the process of estimation of company competitiveness, the analysis of cost structure of a product can be one of very important criteria.

3 DEVELOPMENT OF TECHNOLOGICAL DATABASE

Different approaches can be applied in process planning. These may, for example, include very precise definition of IF THEN procedure for the creation of technological knowledge database. Or, we can decide for the use of fuzzy logic and certainties of possible solutions. Or, we can try to solve a restricted area of the problem by heuristic approach. What can it mean? Technological processes are basically based upon product drawings with adequately defined dimensions, tolerances (dimensional and geometrical), surface roughness, batch size, shape and kind of material, heat treatment, requested delivery, disposable equipment, tools, etc. At the same time, process plans are primary result of the planner’s experience, intuition and decision support. The process planner can establish possible connections between drawing features and necessary production times for products manufacturing.

The fundamental idea in the approach [9] and [10] to production time estimation is investigation of the existence of some kind of relationship between the shape and data from the drawing and the process type, process sequencing, primary process, way of tightening, selection of tools, machine tools, production times, etc (Figure 1). The greatest challenge is to establish (or investigate) the most important factors from the drawing for useful, easy, fast and very exact estimation of production times. This is necessary in the process of offers definition for better estimation of the terms of product delivery, production times and costs, manufacturing management, and last but certainly not the least important, the price of the product.

As one of the first steps in our project research, we defined possible shapes of raw material and 30 potential basic technological processes. Parameters of the basic technological processes can be:

- **shape and kind of raw material** (features of drawing, knowledge base),
- **type of workpiece** (features of drawing, shape and dimensions of raw material),
- **necessary operations** for treatment (features of drawing, expected production time, knowledge base),
- **operations sequencing** (features of drawing, necessary operations of treatment, knowledge base),
- **necessary production times** (features of drawing, based on equations).

So, it is necessary to establish the features of the drawing (independent variables), possible dependent variables, size and criteria for sample homogenization (principles of group technology) for carrying out the analysis of variance and, regression analysis. Relations between the tables (Fig. 2), features of the drawing and defined technological data are important elements of the realized technological processes and procedures.

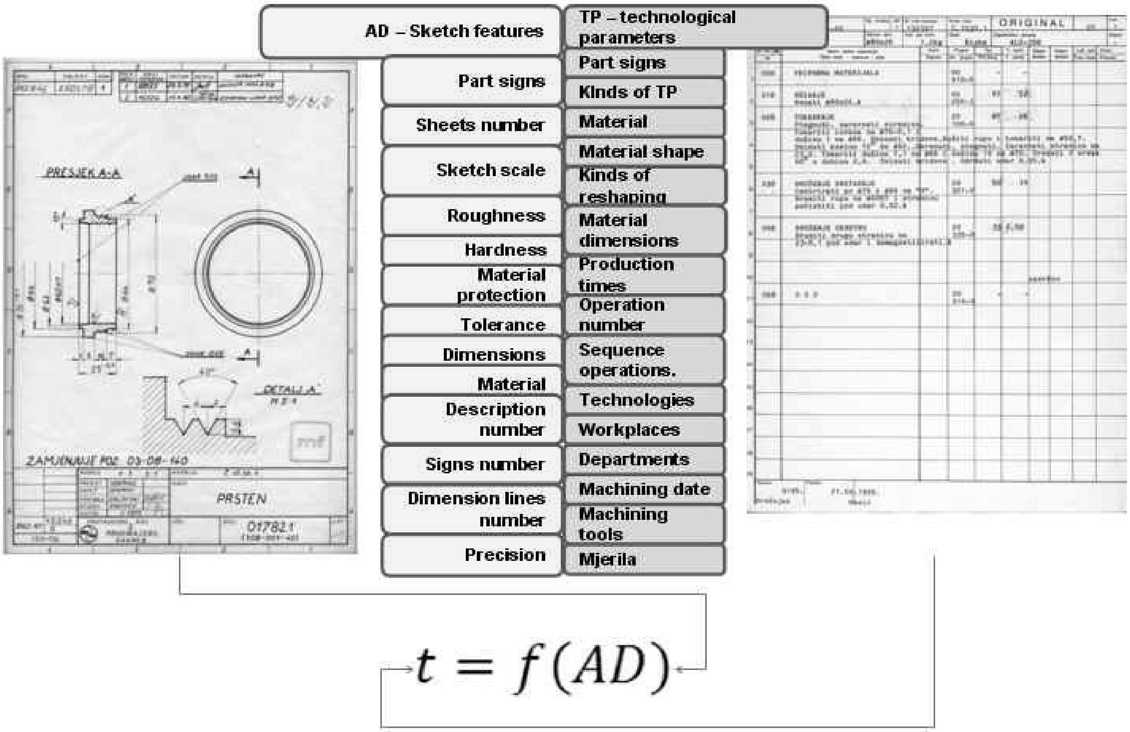

Fig. 1. *Relationship between drawing features and normed technological data [8]*

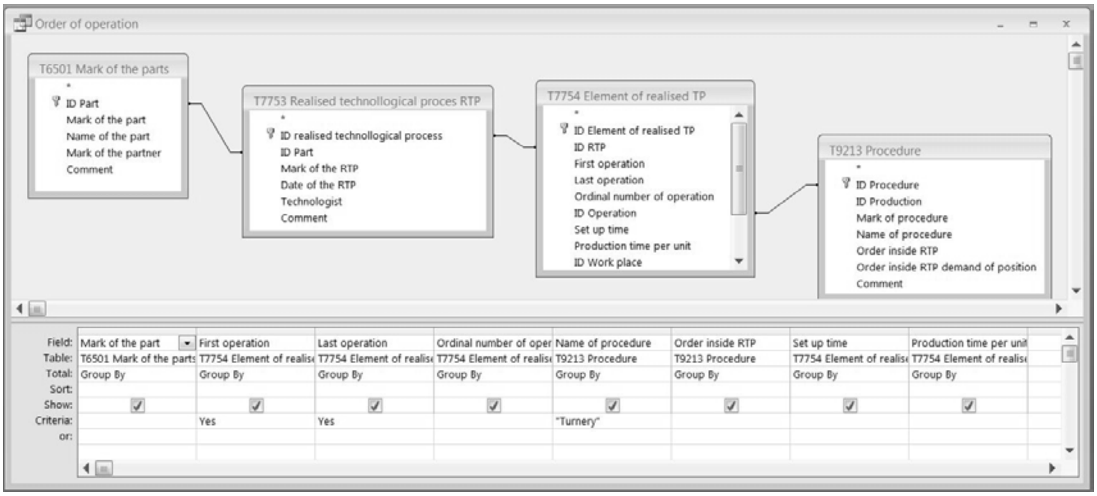

Fig. 2. *Relations between the tables [8]*

4 PREVIOUS RESEARCH OF ESTABLISHED HYPOTHESES

The main questions were the following: How to establish the relationship between the technical drawing of the part and production time estimation? How to estimate production time/costs without detailed process plans? Can the estimation be fast, easy and relatively reliable? What are the expected influence factors for time estimation? What would be exact criteria for efficacy of estimation? In the previous research we had defined domain borders of independent variables, reduced the number of independent variables using correlation/factor analysis and defined the type of smoothing curves with a high index of determination. Of course, a desirable level of generalization in regression analysis will be an important indicator for the quality of regression equation. For different positions within the same group with a significant level of homogeneity, 8 regression equations were established using stepwise multiple linear regression.

The next steps of the research were conducted in the way of automatic recognition and joining a part to the adequate group of parts (logical operators in the database). The research as the second request included more precise measurement and calculation of the production time of the parts. The third request concerned the procedure for the estimation of multiple linear regression with the least variables, the greatest index of determination and good coincidence of calculated and predicted values of dependent variable.

5 HOMOGENIZATION OF THE SAMPLE, CLASSIFIERS SELECTION AND SIGNIFICANCE OF VARIABLES

Blank material may be expressed by three basic groups: quality, shape and dimensions. Investigation of the relationship between production time and features of product (through four groups of independent variables) can give as result a regression equation. All elements of the sample are records for the created database (Fig. 2).

For establishing potential high quality relationship between features of the drawing and production time we had to execute two actions.

One action can be explained as exploring measures for the reduction of the number of independent variables for regression analysis. The method of analysis of variance (ANOVA procedure) and stepwise multiple linear regression (Excel, MatLab) were very helpful in the process of reduction of the number of independent variables.

The other action was the process of sample homogenization, for example, elimination of too big or too small values of the members of the sample. Product shape (dependent variable), as the most important criterion was established for 8 different product types. As result, we developed several regression equations. The size of the samples is the result of sample homogenization and query of logical operators (classifiers) for 12 basic technological operations (OTP), (Fig. 3).

6 DEVELOPMENT OF STEPWISE LINEAR MULTIPLE REGRESSIONS

A desirable level of generalization in regression analysis will be an important indicator for the quality of regression equation. One of the most important problems was the process of homogenization of the sample of products. Adequate method for this action was one of the methods of group technology.

Logical operators during query process in database Access (Fig. 2) were very helpful in the process of homogenization of the sample of products. As the result of the previous research, sample homogenization, classifier selection and stepwise multiple linear regression we obtained (Fig. 3): 7 independent selected variables, basic sample of 320 parts, constraints for data parts, 8 regression equations, percentage of explained effects, relative error (7 to 30%), etc. (Table 1).

The lowest relative error -7.20% (Fig. 3, for grinded discs, No. 5) and the highest index of determination r^2 = 0.9851 for the grinded discs group are the consequence of the simultaneous action of logical operators (round bars, discs and fine machining – i.e. diameter tolerance better than IT7). Thus, with the simultaneous action of several operators, a lower scattering of production time values has been achieved, i.e. better homogeneity of the created group.

The other approaches included critical analysis of the values of dependent variables and excluding of those with extreme values. As an example of multiple linear regression (seven variables), after classification actions and stepwise multiple regression, a group of 221 parts in a sample was selected.

7 AUTOMATIC RETREIVING OF THE DATA FROM 3D MODEL OF THE CONSIDERED PART

The second phase in the research was the investigation of the possibility for easy automatic, direct retrieving of 3D features of the considered axial symmetric product to the regression model. The defined requirement resulted in the development of the process for the transfer of parameters from 3D models with a low level of subjectivity. It is a very fast and reliable process via CAD report to the regression model [11]. As the first step some dimensions were considered as parameters (outside product diameter, inside diameter, product length and width as well as product thickness).

For the development of IT application (PhP), Pro/ENGINEER Wildfire3.0), Mozilla Firefox, Ultra Edit 32, WAMP5 (the server runs PhP code) and Excel 2003 & MatLAB 2006 software was used.

The first version of the application was based on retrieving in *html format* and the last version of the report was in *txt format*.

As the second step of our research, an application was developed for recognition of the important parts of the source code and retrieving from CAD application (Pro/E) to output table report. As a result we obtained a clearly defined table of possible independent variables (number of features, number of dimensions with tolerances, the total number of dimensions, additional parameters such as mass, volume, superficial area).

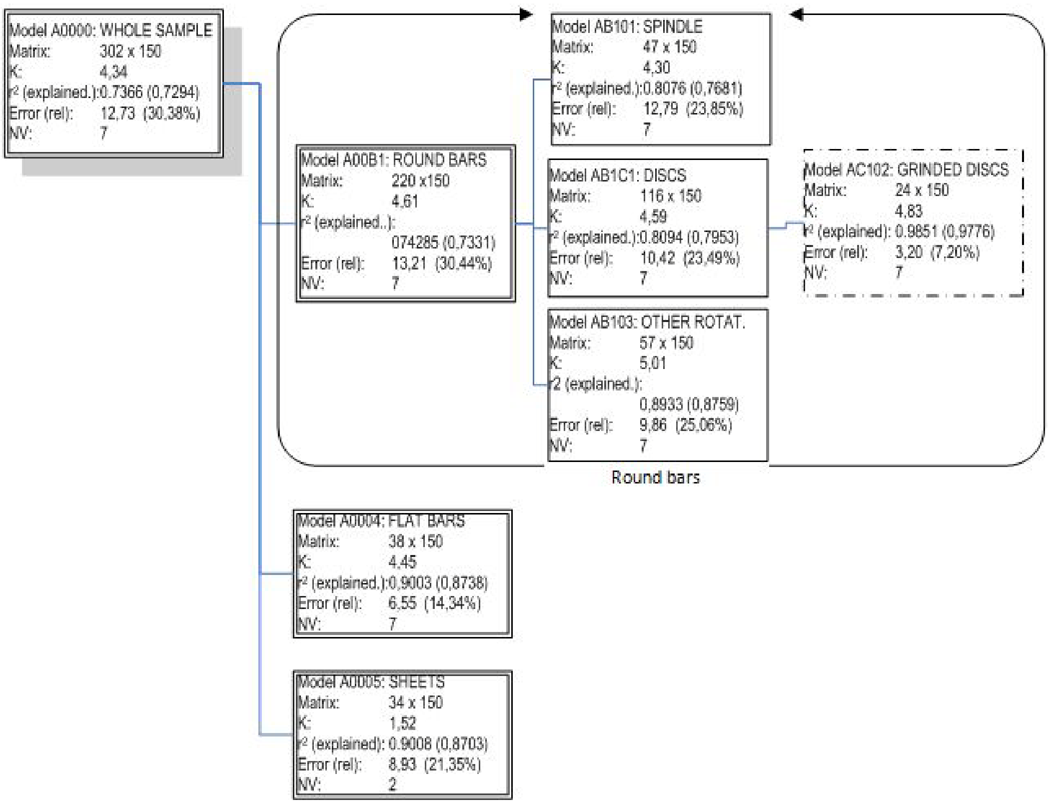

Fig. 3. *Result of classification actions in technological knowledge base [9]*

Using these additional new independent variables it is possible to compare precision and reliability of the ‘classical’ and ‘automatic’ estimation of production times, estimate competitiveness of different competitors and suppliers and develop some aspects of decision support for management through the use of optimization tools (genetic algorithm). With these ‘tools’, we can use some aspects of competitive intelligence and meet some requirements of virtual manufacturing.

8 DESIGN OF THE 3D MODEL OF THE PRODUCT

During modelling process it is necessary to use some rules for mathematical definition of the geometrical model rigid body, for example: rigidity, homogenous three-dimensional body, finality, closed condition of basic transformations and some Boolean operations, finality of description, definiteness of the superficial area [11].

For easier data transfer (variables) from 3D model, it is desirable to define unambiguously characteristic dimensions of the model of the product. The simplest manner is assigning for the important dimension lines (naming the important dimension lines). These names will be variable names for later data sharing: outside diameter of the product, inside diameter of the product, product breadth, thickness, and length.

Some of them were very complicated to detect in the report generated from the application. It was necessary to develop algorithm which had to anticipate many various product shapes. After all the mentioned conditions of the model stability had been satisfied (mathematical definition of geometrical model rigid body), it was necessary to define all characteristics of the surfaces, shape and position tolerance from the technical drawing. Defined characteristics after finished modelling are: tolerance of dimension lines, roughness of surface, tolerance field, position and shape tolerance.

9 ADVANTAGES AND DISADVANTAGES OF AUTOMATIC RETRIEVAL OF FEATURES

"Migration" from a classical technical drawing to a 3D model of a product for automatic retrieval of features has important advantages:

- automatic retrieval of features and use in other applications,
- it is a good manner to avoid subjective features from the drawing which generally depend of the designer (scale, number of views, number of dimension lines, etc),
- elimination of the possibility for loss or wrong assumption of features (elimination of operator’s influence),
- possibility for adoption of additional variables that are not observable from the 2D drawing,
- better perception of the real product appearance,
- possibility to observe geometrically non-definite models – during report generation it is visible whether the model is fully geometrically defined.

Some of the most important disadvantages of "migration" from a classical technical drawing to a 3D model of a product for automatic retrieval of features are:

- need to have 3D CAD application and necessary skills for using it,
- making of a 3D model in the case that it is not supplied by the client,
- it is necessary to keep to the rules for mathematical definition of geometrical model in order to make possible retrieving of data (features),
- unclarity of presentation of defined dimension lines, tolerances and surface characteristics - if all the necessary features are not defined (surface roughness, bore tolerance field, tolerances of dimension lines), it is hard to detect at first sight their incompleteness compared to a 2 drawing where this can be easily seen,
- impossibility of retrieving the features related to surface quality (roughness), shape and position tolerances, radial and axial run-out.

10 DEVELOPMENT OF THE WEB APPLICATION FOR AUTOMATIC RETRIEVAL OF FEATURES FROM CAD SOFTWARE

In Pro/Engineer - CAD application there is a possibility for adjusting format for writing of reports. The *html format* is better for easier source code (one colour, size and font style). Therefore, code application is simpler for writing and modification.

Reports generated in *txt format* are easier for presentation of generated data (letter size, font style, letter colours).

11 RUNNING OF APPLICATION FOR RETRIEVAL OF FEATURES FROM CAD SOFTWARE

As an example of multiple linear regression, after classifiers actions and stepwise multiple regression, a group of rotational parts with 9 parts in a sample for 2D paper drawing was selected. Observed multiple linear regression for 4 independent variables $Y = f(x_{16}, x_{21}, x_{31}, x_{44})$ has the index of determination $r^2 = 0.998651$ and regression equation [11]:

$$Y = 757.472 + 2.52562x_{16} - 0.333048x_{21} - 62.3869x_{31} + 0.300431x_{44} \quad (1)$$

$0 < x_{16} < 3$– number of descriptions
$10 < x_{21} < 36$– number of usual dimension lines
$11.21 < x_{31} < 11.71$– mass strength of material
$43.18 < x_{44} < 290.45$– volume of material
$45.00 < Y < 111.00$ – production time.

Observed multiple linear regression for 6 independent variables, for the same sample and 3D model is $Y = f(x_{19}, K_s, f_{ea}, x_{45}, V, x_{43})$ (2), it has the index of determination $r^2 = 0.9918$ and regression equation:

$$Y = 28.77308 + 8.277896x_{19} - 0.16359K_s - 1.46341f_{ea} - 50.8704x_{45} + 0.000324x_{44} + 0.002462x_{43} \quad (2)$$

$2.00 < x_{19} < 8.00$– tolerance of dimension line of the part
$13.00 < K_s < 46.00$ – all dimension lines
$9.00 < f_{ea} < 25.00$ – features of 3D
$0.174 < x_{45} < 0.584$– mass of the part
$4{,}063.80 < x_{44} < 74{,}724.50$ – volume of the part
$6{,}660.70 < x_{43} < 28{,}131.30$ – superficial area
$45.00 < Y < 111.00$ – production time.

Error between estimation by regression and calculated production time for each part (-5.64%;+ 4.32%).

12 CONCLUSION

Using the "manual" retrieval of features from a 2D model, 8 regression equations were obtained for the estimation of production time with the index of determination r^2 (0.736552 to 0.985057).

So, it can be very useful as decision support for the consideration and accepting of offers. Using automatic retrieval of features from a 3D model, values of variables were obtained. Retrieval of features from a 3D model, compared to the classical ‘’manual’’ retrieval from a 2D paper drawing introduces new variables which are important in the process of 3D modelling.

In the near future further research will be oriented to different multi-objective optimization models for the estimation of dependent variables: production time, size of profit, delivery deadlines, and distribution of product manufacturing per working places/machines.

ACKNOWLEDGEMENTS

This research is a part of the scientific projects (2007-2009) titled *Process Production Impacts to the Competitive and Sustainable Development* 120-1521781-3116 (main researcher Prof. Predrag Cosic Ph.D.) financed by the Ministry of Science, Education and Sports of the Republic of Croatia and Slovenian-Croatian project of bilateral cooperation in science and technology, (2007-2008) named *Virtual Manufacturing – Step to Competitive and Sustainable Development* (main researcher Prof. Predrag Cosic Ph.D.), financed by the Ministry of Science, Education and Sports of the Republic of Croatia and Slovenian Research Agency (ARSS), main researcher in Slovenia Prof. Slavko Dolinsek Ph.D. We express gratitude for the financial support to these projects.

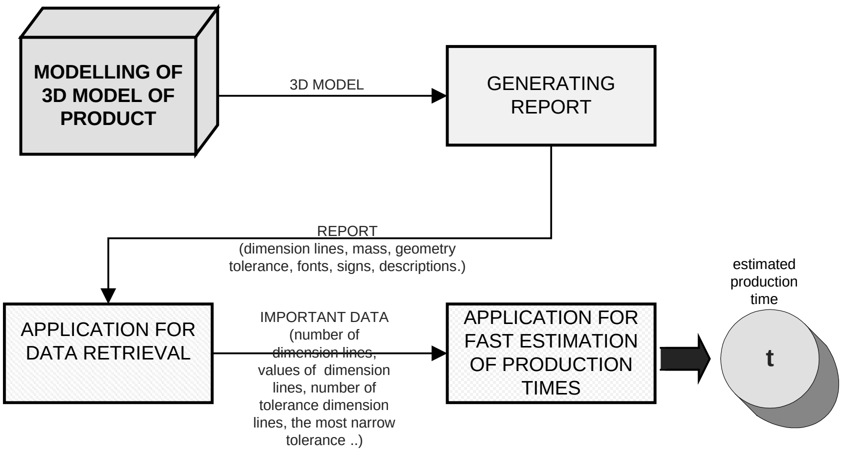

Fig. 4. *Shematic overview for the process of fast estimation of production times [11]*

No	Shape of product - representative of product group	Regression equations	Index of determination r^2	Relative error [%]	Comment on regression equation
1	Whole sample	$t = -11.69 + 16.95x_{45} + 1.22\,x_{40} + 0.54\,x_{47} + 127.47x_{22} – 3.24x_{18} + 0.15x_{32} + 0.03x_{6}$	*0.736552*	30.74	Model is developed with procedure in advance. Three independent variables are omitted x_8, x_{19} and x_{33}.
2	Round bars	$t = 55.47 + 22.43x_{45} + 1.162x_{40} + 0.43x_{11} + 1.61x_{50} – 5.41x_{8} – 3.26x_{18} + 1.78x_{42}$	0.74285	30.95	Model is developed with procedure in advance.Two independent variables are omitted x_1 and x_{26}.
3	Shafts	$t = 6.13 +0.83_{x2} +1.27_{x39} – 3.30_{x8} +5.51x_{46} – 6.86x_{18} +0.09\ _{x6} + 124.33x_{22}$	0.807626	25.90	Model covers more narrow field of rotational parts. It gives better results than No.2.
4	Discs	$t = -5.17 + 0.73x_{47} + 0.93_{x40} + 5.25_{x20} + 0.52x_{24} + 139.11x_{30} + 0.23x_{32} – 0.51x_{33}$	0.809405	24.24	Simillar results as in No.3.
5	Discs-with fine machining	$t = -60.78 + 0.59x_{47} +047_{x9} +0.74_{x1} + 0.25x_{10} + 0.84x_{39} + 291.07x_{25} + 5.9x_{15}$	0.985057	8.01	Model covers more narrow field of rotational parts. It gives better results than all the previous models.
6	Rotational parts	$t = -37.11 + 0.94x_{40} +0.03x_{29} +319.22x_{26} + 0.13x_{23} + 114.67x_{43} – 80.98x_{45} - 0.46x_{6}$	0.893321	27.06	Model is better than No. 2 as a result of higher degree of homogenization of data. Solution is better with omitted variables x_2 and included variables x_6, x_{23}, x_{43} and x_{45}.
7	Flat bars	$t = -10.96 + 0.58x_{40\,+}34.50x_{45} +218.42x_{22} – 5.48x_{50} + 185.03x_{26} +0.39x_{9} -0.50x_{49}$	0.900332	15.92	Constraints are greater for all variables so results are better. Narrow field of homogenization.
8	Sheet metals	$t = 0.47 +1.27x_{40} +137.45x_{45} – 13.23x_{43} – 0.70x_{43} + 0.28x4 + 0.05x_{6} +3.91x_{16}$	0.900823	24.04	Model is characterized with the presence of complex variables x_{40}, x_{43}, x_{45}

Table 1. *Presentation of created regression equations*

Table 2. *Meaning of symbols*

Symbol	Physical unit	Meaning of symbols
f_{ea}	-	Features of 3D
K	-	Coefficient of time
K_s	-	All dimension lines
r^2	-	Index of determination
t	[min]	Machining time
x_1	[IT]	Order of tolerance outside diameter
x_2	[mm]	Outside diameter of material
x_4	[mm]	Width of material
x_6	[mm]	Length of material
x_8	[Class h]	Roughness of open areas
x_9	[HRc]	Hardness of product
x_{10}	[mm]	Outside diameter of product
x_{11}	[mm]	Inside diameter of product
x_{15}	[-]	Number of product perspectives
x_{16}	[-]	Number of descriptions of product
x_{18}	[-]	Number of location marks (geometry)
x_{19}	[-]	Number of dimension line tolerances
x_{20}	[-]	Number of special dimension lines
x_{21}	[-]	Number of usual dimension lines
x_{22}	[1/class]	Roughness request Ra
x_{23}	[1/mm]	Location request (geometry)
x_{24}	[1/mm]	Dimension request
x_{25}	[1/IT]	Diameter request
x_{26}	[mm^2]	Area of sketch
x_{29}	[N/mm^2]	Ultimate tensile strength of material
x_{30}	[m^2]	Requested area of sketch
x_{31}	[-]	Mass strength of material
x_{32}	[mm]	Thickness wall of products
x_{33}	[-]	Ratio of diameter and length
x_{39}	[-]	Number of all dimension lines
x_{40}	[-]	Product complexity
x_{42}	[Class h]	Difference of roughness
x_{43}	[dm^2]	Difference superficial areas of material
x_{44}	[cm^3]	Volume of material
x_{45}	[kg]	Mass of material
x_{46}	[mm]	Difference of outside diameters
x_{47}	[mm]	Difference of outside diameter of products
x_{49}	[mm]	Difference of thicknesses
x_{50}	[mm]	Difference of lengths
Y	[min]	Machining time

REFERENCES

[1] Han, J., Kang, M., Choi, H. (2001). STEP – based recognition for manufacturing cost optimization. *Computer-Aided Design* vol. 33, p. 671-686.

[2] Culler, D.E., Burd, W. (2007). A framework for extending computer aided process planning to include business activities and computer aided design and manufacturing (CAD/CAM) data retrieval. *Robotics and Computer-Integrated Manufacturing*, vol. 23, p. 339-350.

[3] El-Melahalawi, M., Miller, R.A. (2003), A database system of mechanical components based on geometric and topological similarity. Part I: representation. *Computer-Aided Design*, vol. 35, p. 83-94.

[4] El-Melahalawi, M., Miller, R.A. (2003). A database system of mechanical components based on geometric and topological similarity. Part II: indexing, retrieval, matching, and similarity assessment. *Computer-Aided Design*, vol. 35, p. 95-105.

[5] H'mida, F., Martin, P., Vernadat, F. (2006). Cost estimation in mechanical production: The cost entity approach applied to integrated product engineering. *Int. J. Production Economics*, vol. 103, p. 17-35.

[6] de Cos, J., Sanchez, F., Ortega, F., Montequin, V. (2007). Rapid cost estimation of metallic components for the aerospace industry. *Int. J. Production Economics*. doi: 10.1016/jijpe. 2007.05.016.

[7] Simunovic, G., Saric, T., Lujic, R. (2008). Application of neural networks in evaluation of technological time, *Strojniški Vestnik-Journal of Mechanical Engineering*, vol. 54, no. 3, p. 179-188.

[8] Kostanjevec, T., Balic, J., Polajnar, A. (2009). Multiple criteria analysis of a large-scale product with the DFx *method.*

Advances in Production Engineering & Management, vol. 4, no. 3, p. 105-114.

[9] Cosic, P., Antolic, D., Milic, I. (2007). Web oriented sequence operations. *19th International Conference on Production Research, ICPR-19,* , Valparaiso, on CD.

[10] Antolic, D. (2007). Estimation of production times by regression models, *Master's thesis, University of Zagreb, Faculty of Mechanical Engineering and Naval Architecture* (FSB). (in Croatian)

[11] Kovacic, I. (2007). An overview of fast estimation of production times and delivery deadlines. *Graduation thesis, University of Zagreb, Faculty of Mechanical Engineering and Naval Architecture* FSB.