Journal of Intelligent Manufacturing (2021) 32:2281–2299
https://doi.org/10.1007/s10845-021-01741-y

On the role of complexity in machining time estimation

Antonio Armillotta[1]

Received: 22 April 2020 / Accepted: 16 January 2021 / Published online: 7 February 2021
© The Author(s) 2021

Abstract

Early cost estimation of machined parts is difficult as it requires detailed process information that is not usually available during product design. Parametric methods address this issue by estimating machining time from predictors related to design choices. One of them is complexity, defined as a function of dimensions and tolerances from an analogy with information theory. However, complexity has only a limited correlation with machining time unless restrictive assumptions are made on part types and machining processes. The objective of the paper is to improve the estimation of machining time by combining complexity with additional parameters. For this purpose, it is first shown that three factors that influence machining time (part size, area of machined features, work material) are not fully captured by complexity alone. Then an optimal set of predictors is selected by regression analysis of time estimates made on sample parts using an existing feature-based method. The proposed parametric model is shown to predict machining time with an average percentage error of 25% compared to the baseline method, over a wide range of part geometries and machining processes. Therefore, the model is accurate enough to support comparison of design alternatives as well as bidding and make-or-buy decisions.

Keywords Design for manufacturing · Cost estimation · Machining · Cycle time · Tolerances · Complexity

Abbreviations

a	Machining allowance of a feature
A	Machined area of a feature
A_{tot}	Total area of machined features on a part
c_i	Parameter of a regression model
C	Information content-based complexity of a part
d	Depth of cut of a machining operation
D_i	Nominal value of a dimension
D_{tot}	Sum of the nominal dimensions of a part
f	Feed of a machining operation
H	Entropy of a set of events
I_i	Information content associated with either an event or a dimension
K_M	Material-related correction factor of cutting time
L	Length of a feature
n	Number of predictors in a regression model
N_D	Number of dimensions on a part
N_O	Number of machining operations required on a part
p_i	Probability of an event
r	Correlation coefficient between two variables
R^2	Coefficient of determination of a regression model
q	Batch size
Q_A	Removal rate based on machined area
Q_L	Removal rate based on feature length
Q_V	Removal rate based on removed volume
s	Standard error of a regression model
t	Machining time of a part
t'	Machining time of a part (estimated from a reduced set of predictors)
t_C	Cutting time of a part
t_H	Handling time of a part
t_N	Non-productive time of a part
t_S	Setup time of a part
T_i	Tolerance on a dimension
v	Cutting speed of a machining operation
V	Volume of a part
V_E	Envelope volume of a part
V_m	Machined volume of a feature
x_i	Predictor of a regression model

✉ Antonio Armillotta
antonio.armillotta@polimi.it

[1] Dipartimento di Meccanica, Politecnico di Milano, Milan, Italy

Introduction

The cost of a product can be accurately estimated only at the end of the development cycle, once detailed choices about the manufacturing process (equipment, tooling, operations) have been made. In earlier stages of development, cost would be valuable information to guide design choices in order to meet product specifications (target price, production volume, etc.). However, there is limited availability of cost estimation methods suited to the needs of product designers. These methods should be easy and fast, with a reasonable compromise in terms of accuracy.

Early estimation of manufacturing cost is especially difficult for precision machined parts. The machining cost is usually calculated by multiplying an appropriate shop rate (the hourly cost of equipment, labor, and indirect resources) by the machining time. Accurately estimating the machining time is not a trivial task, as it consists of planning the machining process and calculating the time taken for each operation. This requires the selection of cutting parameters (cutting speed, feed, depth of cut) taking into account specific requirements for the different types of operations (turning, drilling, milling, grinding, etc.). The estimation must also include non-productive and handling times, which depend on the size of the part and the types of machine tools to be used.

Beside requiring specialized manufacturing knowledge, these calculations would take too long during product design, when repeated cost estimates are needed to compare alternatives and evaluate the effects of changes to the product. A designer should be able to estimate the machining time more easily, using empirical functions of predictors related to design specifications. Procedures of this type, referred to as parametric methods, are used in rough order-of-magnitude estimates during product planning or conceptual design. In literature, attempts have been made to extend their use to the preliminary or detailed design stages. This objective requires a careful selection of the predictors in order to keep the estimation error within acceptable limits.

In the search for accurate predictors of machining time, one must obviously consider variables related to the amount of machining work needed, such as the area or length of the machined features. In addition, some studies have focused attention on the complexity of the machined part. This is often expressed in purely geometric terms, e.g. by counting the number of machined features or dimensions. Alternatively, an interesting formulation based on an analogy with information theory combines the dimensions with the tolerances, which are also likely to influence machining parameters. Previous studies (Muter 1993; Hoult and Meador 1996) have shown that a complexity measure defined by these variables is correlated with machining time for specific processes, although it probably does not allow reliable estimates over a wide range of part geometries.

This paper aims to propose a parametric procedure for machining time estimation from data available in the technical drawing of a part. The calculation is based on the regression analysis of detailed time estimates made on a sample of machined parts. The complexity measure is integrated in the regression model of machining time as one of the predictors. Unlike in previous studies, additional predictors are identified to better represent the effects of some design factors that are not fully represented by complexity alone. These include the overall part size, the material, and the types of machined features.

The motivation for the study is twofold. On a theoretical side, it tries to better understand the factors that influence machining time and cost, highlighting their relative importance. On a practical side, it aims to provide an estimation method that is simple yet accurate enough to guide design decisions and help improve the product in early development stages.

The remainder of the paper is organized as follows. Section 2 reviews related estimation methods from literature. Section 3 describes the time estimation problem to be solved, and introduces the basics of the proposed approach. Section 4 explains the method for the development of the parametric method, while Sect. 5 reports the results including the regression equation and the related accuracy tests. Section 6 demonstrates the use of the equation on an example. Section 7 discusses the advantages and limitations of the proposed method.

Literature review

The background of this work includes general methods for estimating manufacturing costs, specific methods for estimating machining cost and time, and approaches to measuring complexity for estimation purposes.

Estimation of manufacturing cost

Cost estimation supports several tasks throughout product development (AACE 2019). During preliminary design, it guides material and process selection as well as benchmarking and make-or-buy decisions. During detailed design, it helps compare design alternatives, evaluate the effect of redesigns, and verify compliance with cost budgets. Once the design is complete, it is useful for submitting bids, revising supplier quotations, and controlling manufacturing expenses.

These opportunities have been especially perceived in the development of highly complex products. Some introductions to cost estimation in the aerospace sector (Rush and Roy 2000; Roy and Sackett 2003) recommend the evolution from an approach based on expert knowledge to a formal data-driven process. Methods and tools for procurement activities at big companies and public agencies are discussed in reports and textbooks (GAO 2009; NASA 2015; Mislick and Nussbaum 2015). Methods for a wider range of applications are covered in Ehrlenspiel et al. (2007).

Cost estimation requirements may vary with the application. The selection of materials (Farag 2014) and manufacturing processes (Lenau and Haudrum 1994; Esawi and Ashby 2003) require streamlined methods with a trade-off on accuracy. The manufacturing cost is split into coarse elements (material, operating, and tooling costs), whose ranges are predicted by means of charts and graphs as a function of production volume; these are used to rank candidate process choices that meet product specifications (Ashby et al. 2007). More detailed estimates are needed for the comparison of design alternatives: as described in Weustink et al. (2000), the product is broken down into a relational structure with multiple levels (assembly, parts, features), each of which should be associated with process-dependent procedures for cost estimation. Such an analysis requires a correct cost structure, where indirect costs should be explicitly estimated to better identify cost reduction margins (Ulrich and Eppinger 2007).

Three main types of formal methods have been proposed for the estimation of manufacturing costs of whole products, assemblies or individual parts. *Analogy methods* are based on historical cost data, which are retrieved using group-technology coding (Ehrlenspiel et al. 2007) or case-based reasoning (Rehman and Guenov 1998; Duverlie and Castelain 1999). The accuracy of the estimate depends on the availability of similar cases, which requires considerable prior work for classification and normalization. To overcome this difficulty, some studies analyse the estimation process based on expert judgement, and translate it into procedures or rule-based systems (Rush and Roy 2001; Mauchand et al. 2008; Molcho et al. 2014).

Parametric methods are based on cost-estimating relationships (CER), which express the cost as a function of one or more variables (cost drivers). Different levels of accuracy can be achieved with this approach: in early development stages, the cost of a product may be estimated from either the mass or an appropriate functional parameter (Ehrlenspiel et al. 2007); in preliminary or detailed design, multiple cost drivers are usually preferred for more accurate estimates. Regression models are easy to use, and their uncertainty can be statistically evaluated within the range of available data (Foussier 2006); they are therefore recommended in procurement activities, where the importance of data normalization is again emphasized (Mislick and Nussbaum 2015). As an alternative parametric method, neural network models (Zhang and Fuh 1998) are said to be more accurate whenever costs cannot be modeled using common regression equations (Smith and Mason 1997; Cavalieri et al. 2004); their disadvantages include the lack of an interpretable equation and a statistical error magnitude. Advanced statistical methods proposed to improve regression accuracy include support vector regression, generalized additive models and gradient boosted trees (Huang 2007; Loyer et al. 2016).

Engineering build-up methods provide accurate cost estimates from a detailed description of manufacturing activities (work breakdown structure). For complex products, calculating the costs of individual activities requires gathering a large amount of information from many sources (Roy et al. 2011), and using detailed cost allocation criteria borrowed from accounting methods (Locascio 2000). These complications justify the development of supporting tools for engineering cost estimation; the requirements for software implementation are discussed in Roy and Sackett (2003), Nasr and Kamrani (2007), and NASA (2015). Proposed approaches include the integration of process planning methods in cost estimation tools (Grewal and Choi 2005) and the extraction of data from CAD product models (Liu and Basson 2001).

Estimation of machining time

Methods for machining cost estimation are reviewed in Niazi et al. (2006) and Garcia-Crespo et al. (2011). The problem involves an individual part (workpiece) which has a set of geometric features created through machining operations. As already mentioned, the problem usually comes down to estimating machining time. This can be done with several possible trade-offs between accuracy and calculation effort.

In the most accurate method, which is commonly used in downstream engineering work, the machining process is detailed into a sequence of cutting operations on one or more machine tools. The time for each operation is calculated by well-known shop formulas as a function of cutting parameters (Tanner 2006). Variations of the basic method differ in some assumptions or calculation details; for example, Creese et al. (1992) suggest two different expressions of cutting time: feed length divided by feed rate, and cutting path length divided by cutting speed. For turning operations, Jha (1996) calculates times and costs by means of optimization of cutting parameters with constraints on machining power and surface roughness. For drilling and milling, Maropoulos et al. (2000) correct default cutting parameters according to the specified roughness. A software tool described by Perera (2014) streamlines time estimation for different types of operations by recommending optimal values of the cutting parameters. In other software tools, feature dimensions and types of operations are automatically recognized from either

CAD models (Roberts et al. 1997; Germani et al. 2011) or G-codes for CNC machining (Ben-Arieh 2000).

To simplify the calculation, parametric methods estimate machining time without breaking it down into individual operations. Simple regression models find little use except for single machining processes: an equation mentioned by Creese et al. (1992) gives the cutting time of rotational parts as a function of mass, with corrections related to the material and to the fraction of volume removed. Better accuracy is achieved by multiple regression models. For rotational parts, Mahmoud (1979) proposes an equation for time estimation from part dimensions (length and average diameter) and coefficients related to material, type of lathe, specified tolerance, and machining complexity; the latter is expressed as the number of setups plus the number of discontinuities on the machined surface. Within a comprehensive method for cost-based process selection, Swift and Booker (2013) provide a cost equation for a wide range of part types; a coefficient related to geometric complexity is evaluated from a two-digit code depending on the general shape (rotational, prismatic, etc.) and on feature complexity attributes.

Feature-based methods allow a compromise in terms of accuracy and ease of use between engineering and parametric methods. They estimate machining time as a sum of contributions from individual part features. For each feature, the cutting time is calculated considering an appropriate removal rate depending on material and feature type. Lovejoy et al. (2005) always use a removal rate related to machined volume; Jung (2002) uses the same parameter for roughing operations, and a removal rate related to machined area for finishing operations; in addition, length-related removal rates are used for holes (Polgar 1996) and for end-milled features (Boothroyd et al. 2011). For prismatic parts, Ou-Yang and Lin (1997) estimate drilling and milling times from removal rates with feature extraction from CAD models. Rao et al. (2005), different removal rates are used to build metamodels (response surfaces) of machining cost to allow shape optimization on aerospace engine parts. Other feature-based methods directly estimate machining cost from variables related to feature type and complexity attributes (Feng et al. 1996), or use group-technology coding to select the cost drivers (Geiger and Dilts 1996).

Special or hybrid methods have also been proposed for specific applications. A neural network model is used in Atia et al. (2017) to estimate the machining time of rotational parts; the input variables cover a broad set of specifications including the main part dimensions and the types of machined features. Stockton and Wang (2004), a similar method with different cost drivers is compared for accuracy with a detailed method based on shop formulas. For prismatic parts, Shebab (2001) and Shebab and Abdalla (2001) apply fuzzy rules to the attributes of machined features (types, dimensions, tolerances, roughness) to evaluate machining time through membership functions (low, average, high). For procurement applications in the aerospace sector, Watson et al. (2006) propose a parametric method enhanced by analogy with historical data; the latter is used both to select a suitable regression model for the part, and to estimate an equivalent shop rate from supplier quotations. In Qian and Ben-Arieh (2008), the parametric method is combined with activity-based costing to include some indirect activities (design, CNC programming, prototyping, etc.) in the manufacturing cost of the part. Finally, a related problem consists of developing software tools for the estimation of CNC milling time on workpieces with free-form surfaces; the calculation is usually based on the toolpath length, which is evaluated from either the G-code (Heo et al. 2006; So et al. 2007; Liu et al. 2013; Shukla et al. 2015, 2016) or the analysis of surface shape (Siller et al. 2016).

Complexity measures

Time and cost represent the effort to achieve a result, and are thus intuitively related to the complexity of the system being analyzed. This observation is the basis of extensive research on the concept of complexity in manufacturing. A review of the topic (ElMaraghy et al. 2012) mentions various ways in which complexity arises in design, manufacturing and business activities, and a general trend towards increasing design complexity (number of parts, technology, size, geometry, variety). A complex system is characterized by many parts and many connections among them, which may lead to uncertain or even chaotic behavior.

Attempts to measure the complexity of a product have sprung up in response to the axiomatic design theory (Suh et al. 1978; Suh 1990), which recommends minimizing complexity as a design strategy. The measure suggested for complexity is based on an analogy with the information theory (Shannon 1948; ISO/IEC 1996). Design reduces uncertainty about product specifications by providing information; this is modelled by describing the product as a message and the manufacturing process as an information channel, which introduces noise in the message. The actual messages that can possibly be received are mutually exclusive events with probabilities p_i; each event has an information content I_i, i.e. the minimum amount of information that can be provided to determine its occurrence. This is measured by the following logarithmic expression:

$$I_i = \log_2 \frac{1}{p_i} \tag{1}$$

The whole set of events has an entropy H, i.e. an average information content, which is given by

$$H = \sum_i p_i \log_2 \frac{1}{p_i} \tag{2}$$

Entropy represents the expected value of complexity, while information content is the impact of the individual event on complexity. The former measure has been used to evaluate the complexity of a signal that may have several levels, as in recent applications of pattern recognition in areas related to machining such as surface metrology (Ullah et al. 2015) and tool condition monitoring (D'Addona et al. 2017). The latter measure is suitable for problems that focus on one event: in design, the event of interest is that the product meets its specifications.

When machining a part, the information content is the minimum amount of information that can be provided to satisfy a specification. For an individual machined feature, Wilson (1980) associates the probability p_i of such an event to the tolerance T_i divided by the nominal dimension D_i; this is the ratio of the favourable cases (the actual dimension is within the tolerance) to all the possible cases (the actual dimension is anywhere between zero and the nominal). Accordingly, the information content I_i of the specification for the feature is

$$I_i = \log_2 \frac{D_i}{T_i} \tag{3}$$

The information content-based complexity as a function of dimensions and tolerances is proposed in Muter (1993) and Hoult and Meador (1996) for the estimation of the cycle time for some manufacturing processes including turning and milling. This measure is shown to grasp the combined effects of part size and shape, which makes it a potentially better predictor of machining time than the purely geometric complexity measures proposed in different contexts such as mould/die cost estimation (dimension count, feature count, perimeter/area ratio, etc.). Tests on datasets of industrial cases reveal a good correlation with actual machining times; however, regression models based on complexity as the sole predictor are restricted to specific machining processes, and seem to neglect other factors that may influence machining time as will be discussed later in the paper.

An alternative product complexity measure is defined in ElMaraghy and Urbanic (2003); the expression consists of the logarithm of the number of features, corrected using coefficients related to feature patterns and other attributes (shape, tolerances, surface finish). A similar measure is also proposed for the complexity of the manufacturing process (ElMaraghy and Urbanic 2004) and demonstrated in design cases (Urbanic and ElMaraghy 2006). Budiono et al. (2014a, b), the above formulation is used for estimating machining time. For this purpose, product complexity is added to another machining-specific complexity index depending on the number of tools and machined sides on the workpiece; the resulting parameter is shown to be correlated with the cutting time for roughing operations.

Other manufacturing complexity indices are defined in Kerbrat et al. (2010) for CNC machining of injection moulding tools. They are evaluated from several mould component attributes such as the outside dimensions, the size of the required end mill, the blank volume, and the removed volume. Index values are displayed using colour maps on digital models of the mould components, but their possible use is also envisaged for estimating the machining hours required for the mould.

The measure of complexity as information content has been used in this work to develop a parametric model of machining time. A novel contribution compared to the above cited results is the attempt to combine the complexity with other predictors selected by statistical analysis of machining data spanning different processes (turning, milling, drilling, grinding). As described below, the analysis has led to the proposal of a multiple regression model with improved accuracy, while retaining reasonable convenience for the purposes of design evaluations.

Problem definition and assumptions

The problem to be solved is described below, specifying its objectives (input and output), and discussing some requirements and basic choices.

Input

It will be assumed that the machining time is to be estimated for a part specified in a technical drawing. The part may be simply an intermediate or candidate design, and its representation may be limited to a layout sketch, where the missing specifications can be retrieved from design notes or technical standards.

Different types of design data may possibly influence the machining time. General specifications for the part include the material, the outside dimensions, and a coarse description of shape (prismatic or rotational) and the machining blank (casting, forging, rolled stock). Detailed specifications for each machined feature include the type (flat, cylindrical, rotational with complex profile, thread, gearing, etc.), the associated dimensions (e.g. diameter and depth for a hole), and the tolerances.

Some assumptions will be made on the above specifications. The types of features are limited to those including a small number of dimensions, with the exclusion of free-form surfaces. Tolerances are only associated with dimensions, with the exclusion of geometric tolerances; as the only exceptions, position or profile tolerances are

converted into equivalent dimensional tolerances on the basic dimensions that define feature shape or location. Surface roughness will not be explicitly considered in time estimation, assuming that tolerances give enough information to understand the machining requirements.

Output

The result of the estimation is the cycle time of the machining process (floor-to-floor time), not broken down into operations as these are not necessarily known at the design stage. The estimate is reported as an expected value with a stated uncertainty. A designer should interpret the estimate as what could be achieved by means of a machining process that meets the productivity criteria of medium-to-high production volumes (at least thousands of units). The following assumptions are also made regarding the machining process:

- The workpiece has limited size and mass (within 300–400 mm and 15–20 kg).
- The equipment consists of CNC machine tools capable of complex sequences of operations with automatic tool changes. They include turning centres for rotational parts, and machining centres for prismatic parts. Additional machine tools may be needed for special operations such as gear hobbing, grinding or slotting.
- Machining operations use tools with a reasonable compromise between productivity and cost (generally with coated carbide inserts), and vendor-recommended cutting parameters.

The cycle time t includes several situations occurring during the machining of a workpiece. In more detailed estimation methods, these should be explicitly accounted for in an expression like the one below, adapted from Creese et al. (1992):

$$t = t_C + t_N + t_H + \frac{t_S}{q} \tag{4}$$

where

- t_C is the cutting time, during which the machine actually removes material from the workpiece.
- t_N is the non-productive time, mostly spent in operations on the cutting tools (tool engagement and return, indexing, tool change), including possible allowances for inspection and operator fatigue (neglected here, based on the above assumptions on process automation).
- t_H is the handling time, corresponding to loading and unloading the workpiece on/from the machine.
- t_S is the setup time, including programming and preparation of the machine (loading and unloading of fixtures and tools, first-article machining and inspection).
- q is the batch size, i.e. the number of workpieces produced between two consecutive setups.

According to the above assumptions on the machining process, the setup time t_S will not be considered as its contribution to cycle time would be negligible due to the large batch size q.

The uncertainty regarding machining time only includes the estimation error due to the limited set of variables used in the predictive model: parts with equal estimates could have different machining times in the actual process or when more detailed estimation methods are used. In practice, an additional uncertainty would be related to the machining choices made by a company. Among these, a prominent role is played by the selection of cutting parameters (cutting speed, feed, depth of cut), which may have a dramatic impact on cutting times. Actually, engineers in different companies choose different parameters for several reasons. Tool vendors recommend different choices of parameters based on their catalogues of tool materials and geometries; workshops may deviate from those recommendations in the attempt to control tool life according to their specific needs. Furthermore, the selection is subject to several constraints relating to available machines (power, torque, vibration control). The problem is coupled since each cutting parameter influences multiple selection requirements (reduced cost or environmental impact; increased productivity, accuracy, surface finish, etc.) unless selection strategies based on axiomatic design are adopted, e.g. Ullah et al. (2009).

A further uncertainty is related to any downtime or unpredictable delay that may occur at an operational level. While a contingency factor may have to be applied to the time estimate for some applications (e.g. bidding or quotations), these issues are of limited importance when the estimate is used to compare design alternatives.

Requirements and basic choices

Ease of use is the chief requirement for an estimation method to be used in product design. The machining time must be readily calculated from the input data described above, without the need for process planning or any complex reasoning. This would also apply if the estimation procedure were implemented in a software tool used to speed up the calculation or extract the data from a digital drawing or 3D model. Although such a tool could include process planning algorithms, the development effort would probably not be worth the objective of improving early cost estimation.

Keeping the procedure simple imposes a compromise in terms of accuracy. However, errors in the order of ± 30%

may be acceptable in preliminary design: the expected time could still allow comparisons, while the upper limit of the prediction interval would be a safer choice for bidding or make-or-buy decisions.

Based on these requirements, the machining time and its statistical uncertainty will be estimated through a parametric model. This raises the problem of choosing which variables to include as predictors in the model. A recent trend in similar tasks is the use of deep learning techniques on large datasets, with the aim of automatically identifying the input variables of a metamodel. In machining time estimation, such an approach might be applied to historical data collected in CNC machining workshops, where cycle times estimated by manufacturing engineers or simulation software (or possibly measured on the shop floor) are associated with design specifications for machined parts. Considering the difficulty of retrieving a sufficient amount of data for unsupervised learning, a standard linear regression analysis is preferred in this paper. Beside requiring fewer machining cases, this choice makes the evaluation of model uncertainty easier, and minimizes possible noise factors related to differences in estimation procedures and production settings at the companies providing the data.

The predictors will have to be calculated from the input data, and selected from those with the strongest influence on the response. For this purpose, it can be noted that the three main elements of machining time are influenced by different attributes of the part:

- The cutting time depends on the total extent of the machined features (which determine the average amount of work needed), and also on the types of features, the tolerances and the material (which determine the relative difficulty of the work).
- The non-productive time depends mainly on the number of machined features (which determine the number of individual operations required).
- The handling time depends almost exclusively on the outside dimensions of the part, regardless of its machined features.

These considerations help to understand the strengths and weaknesses of the complexity measure as a possible predictor of machining time. The complexity is the sum of logarithmic items (3) associated with the dimensions of machined features. Each item represents the size of a feature and the required machining precision, and therefore should have an influence on the cutting time. The sum of the logarithms reduces the size effect of the dimensions and gives more importance to their number, thus influencing the non-productive time. On the other hand, the complexity should have little influence on the handling time, especially for cast or forged workpieces that may be machined on a small fraction of their surface. Furthermore, the complexity does not take into account the types of machined features and the material, which can lead to significant variations in the cutting time.

Consequently, it is unlikely that the machining time can be estimated on the basis of complexity only, unless specific part families are considered as in previous studies. The analysis reported below selects additional predictors that can improve the accuracy of the estimation without requiring information not readily available at the design stage.

Method

The parametric method has been developed from a sample of cases. Each case consists of a machined part, for which a baseline machining time was estimated using a feature-based method from literature. A set of candidate predictors was also evaluated from the specifications of each part. The statistical analysis of the collected data made it possible to fit the sample data by means of a regression model with an optimal set of predictors.

Dataset

The sample includes 80 parts, the drawings for which were collected from various sources. They come from 11 mechanical assemblies including piping and fluid machinery (reciprocating compressor, gear pump, plug valve, globe valve), transmissions (gear reducer, universal joint), tools (circular saw, stamping die, screw vice), and various mechanisms (vehicle suspension, hydraulic jack).

General data for the parts is listed in Table 1. The sample covers size ranges in excess of the limits assumed for the estimation, as well as various materials and all types of basic shapes and blanks considered in the study. Although no statistical sampling was done on these properties, it is believed that the parts are sufficiently representative of a wider range of machine components.

Baseline estimate

For each part of the sample, the machining time was estimated using a baseline method that is thought to be more accurate than the one being developed. Due to the high number of parts, this method was also required to avoid calculations that are too detailed. This ruled out engineering build-up estimates based on process planning and selection of cutting parameters. The feature-based method described by Boothroyd et al. (2011) was chosen as it is based on simple cost parameters, the values of which can be evaluated for a wide range of feature types. Moreover, it is updated to state-of-the-art machining technology and is sufficiently

Table 1 Sample part data

No.	Names	Mat'l	Size, mm	Mass, kg	Type
1	Valve body	BS	90×75×66	0.5	R, C
2	Valve plug	BS	Ø27×89	0.4	R, C
3	Support	CI	355×100×50	7	P, C
4	Pulley	CI	Ø96×55	1.3	R, S
5	Flange	CI	Ø88×13	0.5	R, C
6	Flange	CI	Ø100×20	0.8	R, C
7	Spacer	MS	Ø36×20	0.05	R, S
8	Shaft	MS	Ø36×530	2.9	R, S
9	Casing	AA	150×80×150	1.5	P, C
10	Worm	CS	Ø34×130	0.4	R, S
11	Cover	AA	Ø52×19	0.05	R, S
12	Plug	AA	Ø18×18	0.01	R, S
13	Cover	AA	Ø113×21	0.2	R, C
14	Spacer	MS	Ø21×4	0.005	R, S
15	Shaft	CS	Ø16×110	0.15	R, S
16	Gear	BZ	Ø87×31.5	0.9	R, C
17	Housing	CI	90×30×105	1.6	P, C
18	Stud	MS	Ø4×19	0.002	R, S
19	Support	CI	84×60×65	0.8	P, C
20	Shaft	CS	Ø10×53	0.04	R, S
21	Flange	CI	84×60×20	0.4	P, C
22	Shaft	CS	Ø10×130	0.08	R, S
23	Stuffing box	BZ	35×22×14	0.12	R, S
24	Collar	MS	Ø22×4.5	0.07	R, S
25	Plug	MS	Ø30×12	0.04	R, S
26	Adj. screw	MS	Ø12×22	0.02	R, S
27	Valve	BZ	Ø12×20	0.01	R, S
28	Cover	MS	Ø64×26	1	R, S
29	Shaft sect	CS	Ø45×68	0.15	R, F
30	Pulley	AA	Ø76×15	0.15	R, S
31	Conn. rod	CS	80×17×10	0.03	P, F
32	Cylinder	CI	Ø63×63	0.3	P, C
33	Crankpin	CS	Ø5×28	0.005	R, S
34	Piston	AA	Ø30×32	0.03	R, C
35	Plate	CS	Ø63×4	0.1	R, S
36	Cyl. head	CS	Ø63×28	0.5	P, C
37	Bushing	BZ	Ø8×10	0.003	R, S
38	Plug	MS	Ø14×20	0.008	R, S
39	Pin	CS	Ø8×28	0.008	R, S
40	Bushing	BZ	Ø12×10	0.06	R, S
41	Shaft sect	CS	Ø45×22	0.12	R, S
42	Base	MS	130×68×80	4.5	P, C
43	Base	CI	115×108×56	1.8	P, C
44	Cylinder	MS	Ø40×160	0.8	R, S
45	Piston	CS	Ø28×185	0.7	R, S
46	Cylinder	CS	Ø25.4×65	0.1	R, S
47	Distr. valve	CS	Ø18×14	0.02	R, S
48	Piston	CS	Ø12×88	0.07	R, S
49	Lower shoe	CS	250×250×30	15	P, S
50	Punch	AS	Ø125×28	3	R, S
51	Stripper	CS	Ø155×14	0.7	R, S
52	Column	CS	Ø22×104	0.3	R, S
53	Upper shoe	CS	250×250×25	12	P, S
54	Bushing	CS	Ø45×38	0.3	R, S
55	Shank	MS	Ø36×80	0.5	R, S
56	Blankholder	CS	Ø125×14	1.4	R, S
57	Punch	AS	Ø180×32	2	R, S
58	Blankholder	CS	Ø210×8	1	R, S
59	Spindle	MS	Ø130×395	35	R, S
60	Bushing	MS	Ø175×128	7	R, C
61	Bushing	MS	Ø178×100	6	R, C
62	Bushing	MS	Ø120×65	2.5	R, S
63	Plug	MS	Ø90×40	2	R, S
64	Plug	MS	Ø32×25	0.2	R, S
65	Bushing	BZ	Ø150×240	14	R, S
66	Swivel_arm	AS	266×172×156	6	M, F
67	Hub	CI	Ø246×128	8	R, C
68	Vise body	CI	266×185×187	13	P, C
69	Base	CI	238×201×30	5	P, C
70	Mov. jaw	CI	354×112×101	7	M, C
71	Vise screw	CS	Ø47×475	2.5	R, S
72	Vise nut	CI	Ø48×200	0.8	R, C
73	Cap	CI	Ø92×63	0.8	R, C
74	Key	MS	200×16×10	0.3	P, S
75	Handle rod	MS	Ø17×303	0.5	R, S
76	Valve body	MS	Ø17×303	43	P, S
77	Flange	MS	130×130×45	6	P, S
78	Valve seat	SS	Ø49×140	2	R, S
79	Stuffing box	MS	Ø100×155	5	R, S
80	Valve stem	CS	Ø35×292	0.8	R, S

Material: CI: cast iron; MS: mild steel; CS: carbon steel; AS: alloy steel; SS: stainless steel; AA: aluminum alloy; BS: brass; BZ: bronze

Type: shape: P: prismatic; R: rotational; M: mixed; machined from: C: casting; F: forging; S: rolled stock

proven in real cases, thus allowing partial validation of the parametric method.

The feature-based method calculates the cutting time t_C [min] from removal rates defined in relation to volume, area or length depending on the type of feature and operation; the following criteria apply:

- For a finishing or grinding operation on a generic feature:

$$t_C = \frac{A}{Q_A} \tag{5}$$

where A [cm^2] is the machined area, and Q_A [cm^2/min] is the area-based removal rate.

- For a roughing operation on a generic feature:

$$t_C = \frac{V_m}{Q_V} = \frac{A \cdot a}{Q_V} \tag{6}$$

where V_m [cm^3] is the removed volume, and Q_V [cm^3/min] is the volume-based removal rate. The volume can be replaced by the machined area A if the machining allowance a [cm] is evaluated for each feature according to the difference between blank size and final feature dimensions.

- For an end-milling operation on a contour or groove:

$$t_C = \frac{L}{Q_L} \tag{7}$$

where L [cm] is the length of the machined feature, and Q_L [cm/min] is the length-based removal rate.

The removal rates provided in Boothroyd et al. (2011) do not explicitly consider the effect of the tolerances specified for the machined features. Therefore, a further level of detail was added by recalculating the removal rates for the most common types of operations. For each operation, the recommended ranges for the cutting parameters from tool catalogues were mapped to the allowable tolerance ranges. For example, the removal rate Q_V in a rough-turning operation is

$$Q_V = v \cdot d \cdot f \tag{8}$$

where v [m/min] is the cutting speed, d [mm] is the depth of cut, and f [mm/rev] is the feed. Recommended ranges of these three parameters for mild steel are $v = 110$–160 m/min, $d = 2$–4 mm, $f = 0.2$–0.4 mm/rev. They match the range of ISO tolerance grades between IT10 and IT13 that is commonly allowed for this type of operation. This results in a correspondence between removal rates and tolerance grades as shown in Table 2.

Similar mappings for other cases made it possible to get the list of parameters shown in Table 3. Each parameter is an approximate removal rate for mild steel under a given combination of feature, operation, and tolerance.

The machining time t [min] of a part made of any material is estimated as

$$t = K_M t_C + t_N + t_H \tag{9}$$

where

Table 2 Example of calculation of the removal rate

IT grade	v, m/min	d, mm	f, mm/rev	Q_V, cm^3/min
IT13	110	0.4	4	175
IT12	125	0.53	3.3	135
IT11	140	0.26	2.6	95
IT10	155	0.2	2	60

- The cutting time t_C is calculated from the above equations and parameters for mild steel.
- The correction factor K_M is related to the material. Rough values assumed here include 1 for mild steel, 1.3 for cast iron and medium-carbon steel, 1.5 for stainless steel, 2 for alloy steel, 0.5 for copper alloys, and 0.3 for aluminum alloys. More detailed evaluations could be made according to the machinability ratings available from various sources, e.g. Machinability Data Center (1980) and Drozda et al. (1983).
- The non-productive time t_N is calculated as a constant time per operation (0.1 min).
- The handling time t_H is calculated from the mass of the part (0.5–0.75–1–1.5 min for a part whose mass is less than 0.2–5–15–25 kg respectively).

Evaluation of predictors

The following candidate predictors were selected for the parametric model of machining time, and evaluated for each part of the sample:

- The part's volume V [dm^3], calculated as mass divided by material density.
- The envelope volume V_E [dm^3], i.e. the product of the outside dimensions along three reference axes.
- Identification of the material, equal to the K_M correction factor of the feature-based method.
- Two categorical variables related to the general shape and the type of blank (from Table 1).
- The number N_D of the dimensions D_i created through machining operations (i.e. excluding those possibly pre-existing on the initial casting or forging).
- The sum D_{tot} of the values of machined dimensions:

$$D_{tot} = \sum_{i=1}^{N_D} D_i \tag{10}$$

- The approximate number N_O of machining operations. As a detailed process plan is not available, this is calculated by counting at least one operation for each machined feature. Additional operations are counted considering special features on the drawing (threads, gearings), finishing operations for all features with a tolerance grade within a given limit (IT10 for cylindrical surfaces and profiles, IT9 for holes), and grinding operations for all features with a tolerance grade within a tighter limit (IT7 for planes, IT6 for cylindrical surfaces or profiles) or in the presence of machining notes on the drawing.
- The total area A_{tot} of the machined features.
- The complexity C of the part according to its definition (3) as information content [bits]:

Table 3 Parameters for feature-based estimation of machining time

Operation	IT grade	Q_V, cm^3/min	Q_A, cm^2/min	Q_L, cm/min	Additional parameters
Face milling (rough)	IT12	180			
	IT11	120			
Face milling (finish)	IT10		350		
	IT9		200		
	IT8		100		
Face grinding	IT7		80		
	IT6		60		
Ext./int./profile turning, boring (rough)	IT13	175			
	IT12	135			
	IT11	95			
	IT10	60			
Ext./int./profile turning, boring (finish)	IT10		300		
	IT9		250		
	IT8		180		
	IT7		100		
Ext. cyl. grinding	IT6		50		
	IT5		35		
Int. cyl. grinding	IT6		40		
	IT5		30		
Radial turning/boring			100		
Thread turning/boring			300		
Tapping (D = 2–30 mm)			20–150		
Hole drilling/boring/reaming	IT12		180		
	IT11		150		Multiply by $K_D \cdot K_L$
	IT10		120		K_D = 0.2–0.35–0.6–1–1.5
	IT9		90		(D = 3–6–12–25–50)
	IT8		75		K_L = 1–0.8–0.7–0.55–0.5
	IT7		60		(L/D = 2–3–4–5–6)
Contour/groove milling, slotting, gear cutting	IT10			25	
	IT9			20	

$$C = \sum_{i=1}^{N_D} \log_2 \frac{D_i}{T_i} \tag{11}$$

where the tolerance T_i associated with dimension D_i may be either explicitly specified in a dimension callout or calculated from the IT tolerance grade according to ISO standards (EN ISO 2010).

The parametric model is a linear regression equation that estimates the machining time t of a part as a function of one or more predictors to be selected from the above list. The equation is associated with an estimation error, which determines a prediction interval for the actual machining time for a new part.

Results

Table 4 shows the baseline time estimates and the predictor values calculated for the sample parts.

The error of a linear regression model should have a normal distribution with consistent parameters along the ranges of its variables. However, the machining time does not have a linear relationship with any of the predictors; its variation has a skewed distribution and increases with the values of the predictors. As an example, the graph in Fig. 1a shows the time as a function of complexity with a least-squared trend line and a 90% prediction interval. In Fig. 1b, a logarithmic transformation of both the response and the predictor yields a linear relationship and symmetric, uniform errors.

Table 4 Predictors and response for the parts of the sample

No.	V, dm^3	V_E, dm^3	N_D	D_{tot}, mm	No.	A_{tot}, cm^2	C, bits	t, min
1	0.056	0.446	18	531	29	151	153	6.7
2	0.044	0.065	6	178	10	63	48	3.9
3	0.972	1.775	38	1302	38	364	250	12.5
4	0.181	0.507	16	372	12	755	95	7.2
5	0.069	0.101	6	166	8	101	45	2.2
6	0.111	0.200	5	295	8	300	39	2.9
7	0.006	0.026	3	86	8	60	26	1.7
8	0.372	0.687	21	1009	18	1033	165	7.2
9	0.556	1.800	49	1883	52	255	383	8.0
10	0.051	0.150	20	385	25	110	153	8.1
11	0.019	0.051	17	279	16	116	106	2.2
12	0.004	0.006	8	83	10	21	44	1.5
13	0.074	0.268	16	228	16	86	99	2.9
14	0.001	0.002	3	41	6	16	20	1.2
15	0.019	0.028	15	314	15	54	121	6.4
16	0.100	0.242	15	470	14	125	125	5.9
17	0.222	0.284	78	1051	63	252	539	15.1
18	0.000	0.000	5	42	7	5	30	0.7
19	0.111	0.328	56	577	46	117	344	8.5
20	0.004	0.005	8	118	8	17	67	1.8
21	0.056	0.101	39	340	27	73	245	4.8
22	0.010	0.013	18	284	13	28	141	2.6
23	0.013	0.011	12	148	10	49	82	1.4
24	0.002	0.002	5	53	7	11	36	0.8
25	0.005	0.011	12	95	10	42	62	1.1
26	0.003	0.003	9	91	7	22	52	0.8
27	0.001	0.003	18	111	13	18	124	1.9
28	0.098	0.106	16	299	22	222	117	6.1
29	0.019	0.138	24	302	24	58	311	4.2
30	0.056	0.087	9	190	15	187	67	3.1
31	0.004	0.014	9	102	7	10	63	1.4
32	0.042	0.250	20	458	20	117	115	6.5
33	0.001	0.001	2	33	5	6	16	1.1
34	0.011	0.029	7	113	9	54	56	2.2
35	0.013	0.016	11	154	13	91	72	2.7
36	0.034	0.049	24	347	15	41	162	4.1
37	0.000	0.001	5	30	8	6	36	1.4
38	0.001	0.004	15	86	14	27	88	1.5
39	0.001	0.002	6	64	11	12	49	1.8
40	0.001	0.001	6	39	9	11	40	1.5
41	0.015	0.045	10	129	16	93	73	3.1
42	0.577	0.707	35	623	32	256	247	7.2
43	0.250	0.696	41	706	27	249	265	7.7
44	0.103	0.256	4	250	6	403	35	6.1
45	0.090	0.145	11	481	15	161	92	4.5
46	0.013	0.042	9	216	12	88	76	3.8
47	0.003	0.005	21	166	15	34	116	1.8
48	0.009	0.013	11	198	11	44	81	2.6
49	1.623	1.875	57	3103	34	2174	358	22.9
50	0.385	0.438	11	311	11	401	97	11.4
51	0.090	0.336	6	318	9	183	45	3.5

Table 4 (continued)

No.	V, dm^3	V_E, dm^3	N_D	D_{tot}, mm	No.	A_{tot}, cm^2	C, bits	t, min
52	0.038	0.050	4	213	7	131	40	4.2
53	1.508	1.563	57	3130	33	1776	433	19.5
54	0.038	0.077	6	165	7	146	47	2.0
55	0.064	0.104	10	231	13	185	72	2.8
56	0.179	0.219	11	220	15	336	80	6.4
57	0.256	1.037	13	744	15	1087	108	14.0
58	0.128	0.353	9	962	15	880	68	5.8
59	4.487	6.676	13	1540	13	2149	218	23.8
60	0.897	3.920	16	1304	18	1414	132	12.3
61	0.769	3.168	5	581	8	1673	43	14.2
62	0.321	0.936	8	488	13	883	62	8.1
63	0.256	0.324	5	208	7	298	34	3.6
64	0.026	0.026	8	162	9	77	63	1.9
65	1.556	5.400	3	510	8	2248	32	10.4
66	0.769	7.137	81	1930	56	820	383	24.9
67	1.111	7.746	52	1350	50	903	380	18.4
68	1.806	9.202	33	1412	36	1119	244	17.1
69	0.694	1.435	6	103	6	479	44	5.2
70	0.972	4.004	34	1752	31	721	146	15.6
71	0.321	1.049	12	1354	9	609	96	6.4
72	0.111	0.461	7	324	5	58	53	1.8
73	0.111	0.533	9	291	10	177	71	4.8
74	0.028	0.032	5	236	14	110	37	2.9
75	0.064	0.088	5	614	7	161	35	1.7
76	4.513	4.732	43	2181	32	2917	352	25.8
77	0.729	0.761	16	812	14	861	116	7.3
78	0.256	0.336	19	655	21	359	138	9.0
79	0.641	1.550	15	678	21	666	123	10.9
80	0.103	0.358	14	785	19	248	116	7.3

Therefore, a suitable choice for the model is

$$\log t = c_0 + c_1 \log x_1 + \cdots + c_n \log x_n \tag{12}$$

where x_i are the predictors, c_i are the regression parameters, and n is the number of predictors. Once the parameters have been estimated, the model can be back-transformed into product form:

$$t = k x_1^{c_1} \ldots x_n^{c_n},\ k = 10^{c_0} \tag{13}$$

A proper set of predictors must be selected for the model. It would be wrong to overfit the data by including all the candidate predictors regardless of their statistical significance; while keeping the error at a minimum on the sample, this would not guarantee the same accuracy on further cases. Moreover, an equation with too many variables would not be much help for understanding the individual effect of each of them.

In a first attempt to reduce the model to just one predictor, complexity would be an obvious candidate according to the results of previous studies. However, it is apparent from Fig. 1 that time has a limited correlation with complexity, as was somehow expected from the discussion in subsection 3.3. Other geometric predictors are better correlated with the response, as can be seen from the Pearson correlation coefficients r listed in Table 5. The variables related to part size (V and V_E) seem to be the most suitable choices for a single predictor, probably because they influence all the elements of machining time (cutting, non-productive, handling). The machined area A_{tot} has also a good correlation, which can be explained by its influence on cutting time. The remaining variables (N_D and N_O) are even less correlated than complexity.

For a multiple regression model, the predictors should ideally be independent from one another. If possible, each predictor should influence a different time element. To aid the choice, Table 6 shows the Pearson correlation coefficients between pairs of geometric predictors. Complexity C has little correlation with all the variables except the sum of dimensions D_{tot}; both have an influence on cutting

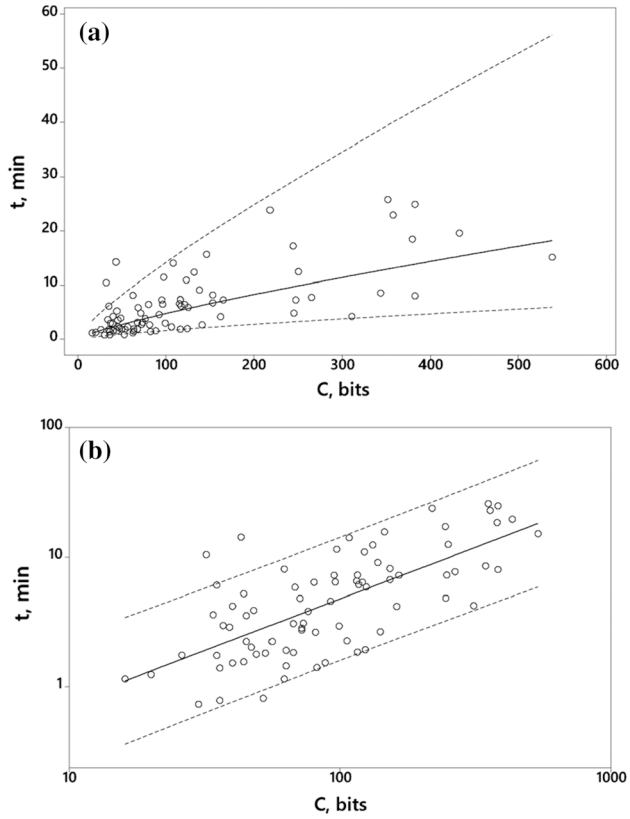

Fig. 1 Machining time as a function of complexity: **a** linear scale; **b** log–log scale

Table 5 Correlations of log t with the geometric predictors

Predictor	r
log V	0.880
log V_E	0.892
log N_D	0.622
log D_{tot}	0.887
log N_O	0.687
log A_{tot}	0.873
log C	0.696

Table 6 Pairwise correlations between geometric predictors

r	log A_{tot}	log V_E	log V	log D_{tot}
log C	0.460	0.557	0.546	0.750
log D_{tot}	0.852	0.883	0.874	
log V	0.935	0.963		
log V_E	0.932			

and non-productive times. The remaining variables are all highly correlated with one another but they influence different time elements: the machined area A_{tot} is related to the required amount of machining work (regardless of its difficulty), while the volumes V and V_E are especially related

Table 7 Analysis of variance in the regression model of log$t$

Source	DF	Adj SS	Adj MS	F-value	p-value
Regression	4	10.8619	2.7155	147.49	0.000
log C	1	0.8478	0.8478	46.05	0.000
log A_{tot}	1	0.3107	0.3107	16.88	0.000
log V_E	1	0.1038	0.1038	5.64	0.020
log K_M	1	0.0704	0.0704	3.83	0.054
Error	75	1.3753	0.0184		
Total	79				

to the handling time. Based on these considerations, a good choice of predictors could include C, V_E (or V), and A_{tot} if they proved to make a statistically significant contribution.

The final selection of the predictors was made by a stepwise regression procedure, which sequentially adds the predictors with the highest residual contribution, and removes those with non-significant contribution. The candidate set of predictors included the variables related to the size of the workpiece (V, V_E), the area of machined features (A_{tot}), the dimensions (C, D_{tot}), and the material (K_M), as well as the categorical predictors related to shape and blank. The best model expresses the machining time t [min] as

$$t = 0.29 \cdot C^{0.37} \cdot A_{tot}^{0.26} \cdot V_E^{0.10} \cdot K_M^{0.15} \tag{14}$$

after back-transformation from a logarithmic model with standard error $s = 0.136$ and coefficient of determination $R^2 = 88.7\%$. As confirmed by the analysis of variance in Table 7, the four predictors are all statistically significant and their relative contributions rank in the same order in which they appear in the equation. The residuals do not show significant deviations from the underlying assumptions of linear regression: they are normally distributed (p-value = 0.881 in the Anderson–Darling test), and have no systematic trends in relation to the fitted values, the predictor values, and the order of data collection.

For the sake of comparison, another regression model was fitted to the data using just two predictors related to part size and material:

$$t' = 8.5 \cdot V_E^{0.32} \cdot K_M^{0.20} \tag{15}$$

In its logarithmic form, the model has $s = 0.176$ and $R^2 = 80.6\%$; the residuals have a significant deviation from normality (p = 0.007) but no apparent trends in relation to fits and predictors. The predictive abilities of the two models can be compared considering their standard errors. Due to the high number of degrees of freedom of the error, the 90% prediction interval on log t is nearly equal to $\pm 1.67\ s$. Therefore t is estimated in an interval limited by the back-transformed regression value multiplied by factors $10^{-1.67\ s}$ and $10^{1.67\ s}$. The model including C and A_{tot} brings s from

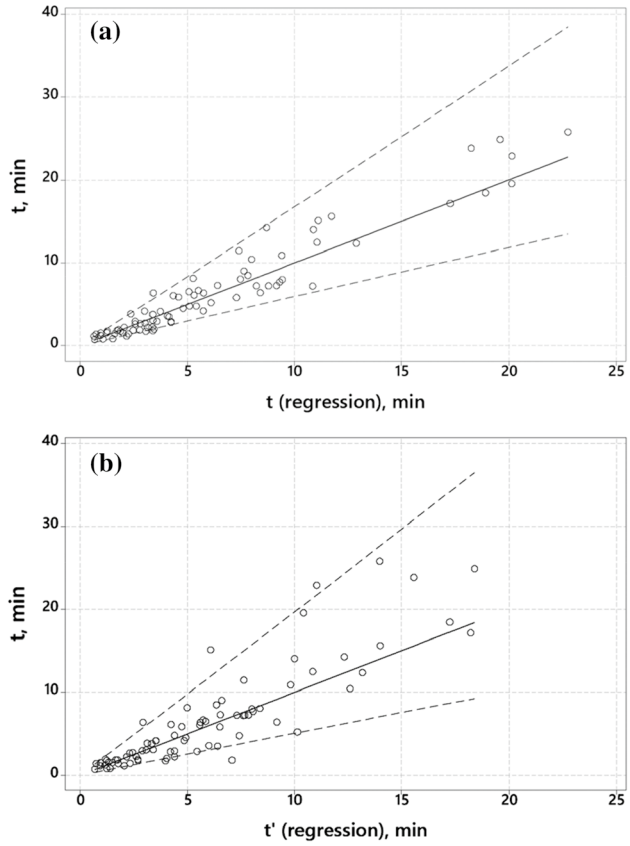

Fig. 2 Parametric vs feature-based estimates: **a** model $t(C, A_{tot}, V_E, K_M)$; **b** model $t'(V_E, K_M)$

0.176 to 0.136, which corresponds to a reduction of the uncertainty from about (−50/+100%) to about (−40/+70%) for the estimate of machining time.

The different prediction intervals are also shown in Fig. 2, which compares the feature-based estimates (9) to the parametric estimates with the two regression models (14) and (15). The full model (Fig. 2a) is visibly more accurate than the simplified one (Fig. 2b). While the absolute error of both models increases with the estimated time, it has been verified that the percentage error is fairly uniform: the mean and standard deviation of its absolute value are about 25 ± 20% for the full model and 40 ± 35% for the simplified one.

Discussion

The parametric model can be put in a more convenient form for practical use. It can be observed that a reference time of 10 min corresponds approximately to a reference set of predictor values: $C = 200$ bits, $A_{tot} = 500$ cm^2, $V_E = 1.5$ dm^3, $K_M = 1$. Accepting a further 5% uncertainty of the estimate, the regression Eqs. (14) and (15) become

$$t = 10\ \text{min} \cdot \left(\frac{C}{200}\right)^{0.37} \left(\frac{A_{tot}}{500}\right)^{0.26} \left(\frac{V_E}{1.5}\right)^{0.10} \left(\frac{K_M}{1}\right)^{0.15} \tag{16}$$

$$t' = 10\ \text{min} \cdot \left(\frac{V_E}{1.5}\right)^{0.32} \left(\frac{K_M}{1}\right)^{0.20} \tag{17}$$

The dimensionless factors in brackets may help explain why the time deviates from its reference value of 10 min. They can be traced back to design choices, and give a feel of what contributes most to the reduction of machining time and cost.

The example in Fig. 3 demonstrates the application of the proposed method. The part is a base of a rotary compressor, and is machined all over from a 1.5-kg grey iron casting. The machining process includes two setups on a CNC machining centre with milling, drilling, reaming and tapping operations.

Table 8 shows the estimation of the machining time of the part using the feature-based method. Each machined feature is associated with an area A and a possible machining allowance a, which are calculated from the dimensions of the feature and the precision expected for the casting. These pieces of data determine the cutting time of the feature by applying a removal rate that is set from the IT grade of the feature's main dimension (from either dimension callouts or assumed general tolerances). The baseline estimate (9) of machining time is 8.6 min, corresponding to an operating cost of 8–9 € for a shop rate of about 60 €/h.

The following predictors are evaluated for the part:

- Complexity: $C = 196$ bits $\rightarrow C/200 = 0.98$.
- Machined area: $A_{tot} = 320$ cm$^2 \rightarrow A_{tot}/500 = 0.64$.
- Envelope volume: $V_E = 0.43$ dm$^3 \rightarrow V_E/1.5 = 0.29$.
- Material factor: $K_M = 1.3 \rightarrow K_M/1 = 1.3$.

Table 9 details calculation of the complexity measure. The dimensions D_i, the tolerances T_i, and the numbers n_i of equal features are considered for all machined dimensions (underlined in Table 8). Considering possible feature patterns, the contribution of each dimension to the complexity is calculated as $C_i = n_i \log_2 D_i/T_i$.

The parametric estimate (16) of machining time is

$$\begin{aligned} t &= 10 \cdot 0.98^{0.37} \cdot 0.64^{0.26} \cdot 0.29^{0.10} \cdot 1.3^{0.15} \\ &= 10 \cdot 0.99 \cdot 0.89 \cdot 0.88 \cdot 1.04 = 8.1\ \text{min} \end{aligned}$$

and is fairly close to the baseline estimate (−6% error). The factors in the equation suggest that, compared to a 10-min standard, the machining time is 12% less due the small part size, and 11% less due to the small machined area. The material gives a 4% increase in machining time, while the complexity has a neutral influence (1% decrease). The model (17) based only on material and part size would give

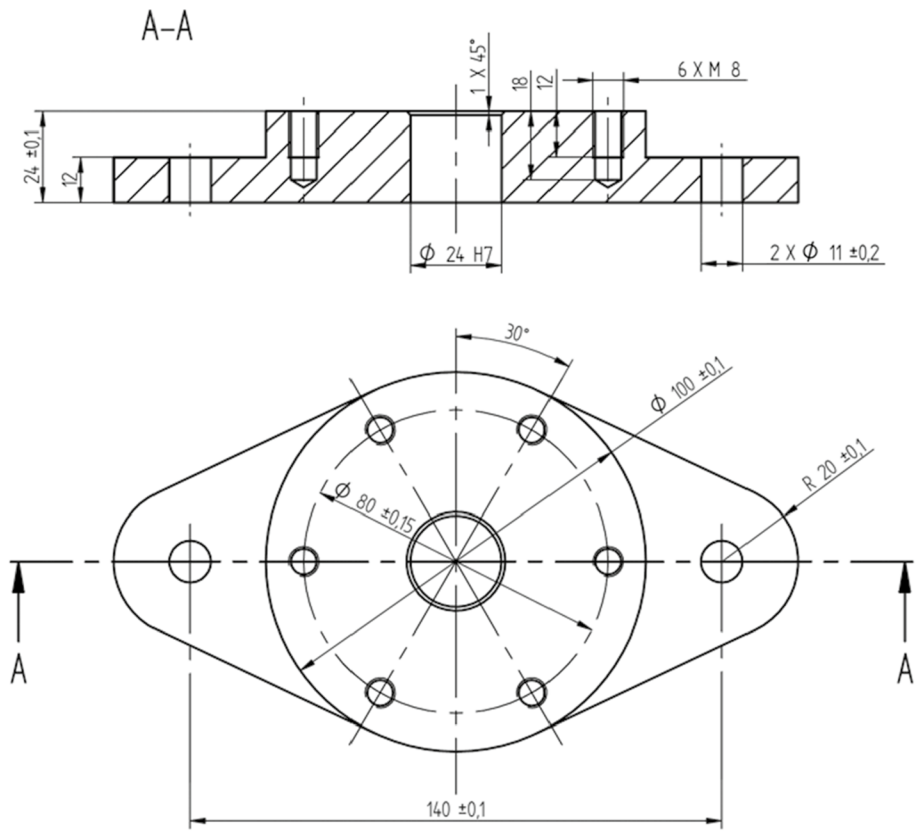

Fig. 3 Example part

Table 8 Baseline estimation of the machining time for the example

Operation	Time calculations	Notes
Face mill, bottom plane	Rough: 110.6 $cm^2 \times 0.3$ cm/180 cm^3/min = 0.18 min Finish: 110.6 cm^2/350 cm^2/min = 0.32 min	IT10, $A = 110.6$ cm^2, $a = 3$ mm
Face mill, top plane, 24 ± 0.1	Rough: 50.3 $cm^2 \times 0.3$ cm/180 cm^3/min = 0.08 min Finish: 50.3 cm^2/350 cm^2/min = 0.14 min	IT10, $A = 50.3$ cm^2, $a = 3$ mm
Drill + ream, Ø24 H7 × 24	10.1 cm^2/(1 × 1 × 60 cm^2/min) = 0.17 min	$A = 10.1$ cm^2
6 × Drill + tap M8 × 18–12, Ø80 ± 0.15	Drill: 6 × 4.5 cm^2/(0.45 × 0.8 × 120 cm^2/min) = 0.62 min Tap: 6 × 3.0 cm^2/50 cm^2/min = 0.36 min	IT10 (M8), IT12 (Ø80), IT13 (18, 12); $A = 4.5$ cm^2 (hole), 3.0 cm^2 (thread)
End mill, midplane, 12, Ø100 ± 0.1	60.3 $cm^2 \times 0.3$ cm/180 cm^3/min = 0.10 min	IT12, $A = 60.3$ cm^2, $a = 3$ mm
2 × Drill Ø11 ± 0.1 × 12, 140 ± 0.2	2 × 4.1 cm^2/(0.6 × 1 × 180 cm^2/min) = 0.08 min	IT12, $A = 4.1$ cm^2
End mill 2 × R20 ± 0.1, Ø100 ± 0.1	43.0 cm/25 cm/min = 1.72 min	IT10, $A = 51.6$ cm^2, $a = 3$ mm
→ Cutting time	3.76 min	
Non-productive time	22 × 0.1 min = 2.2 min	22 operations
Handling time	2 × 0.75 min = 1.5 min	2 setups on CNC mill, 1.5 kg
→ Machining time	1.5 + 2.2 + 1.3 × 3.76 = 8.6 min	$K_M = 1.3$ (cast iron)

$$t' = 10 \cdot 0.29^{0.32} \cdot 1.3^{0.20} = 7.1 \text{ min}$$

with a higher error than the full model (− 17%). It would seem that there is some advantage in using complexity as a predictor, although a single example could not obviously be taken as confirmation of this claim. After all, both models happen to do better than can be expected on average from the proposed parametric method. The real advantage of the full model is in its sensitivity in relation to design choices. This can be seen when evaluating the effects of two possible changes to part design:

- Redesign of the flange into a cylindrical shape, so that the part could be turned from mild-steel round stock:

Table 9 Calculation of complexity for the example

Callout	IT grade	D_i	T_i	n_i	C_i
24 ± 0.1		24	0.2	1	6.9
Ø24 H7		24	0.021	1	10.2
M8	IT10	8	0.058	6	42.6
18	IT13	18	0.27	6	36.3
12	IT13	12	0.27	6	32.8
Ø80 ± 0.15		80	0.3	1	8.1
12	IT12	12	0.2	1	5.9
Ø100 ± 0.1		100	0.2	1	9.0
Ø11 ± 0.1		11	0.2	2	13.6
140 ± 0.2		140	0.4	1	8.4
R20 ± 0.1		20	0.2	2	13.2
Ø100 ± 0.1		100	0.2	1	9.0
Total complexity					196.0

the baseline machining time is now 12.0 min due to the higher amount of material to be removed. The machined area and the envelope volume increase (A_{tot} = 963 cm^2, V_E = 0.78 dm^3), while the complexity and the material factor decrease slightly (C = 184 bits, K_M = 1). As a result, the parametric estimate goes to t = 10.8 min (− 10% error), compared to just t' = 8.1 min (− 32% error) if complexity and machined area are left out of the model.

- Specification of tighter tolerances on all profile and positional dimensions (IT grades reduced by two units): the machining operations require lower removal rates, which change the baseline estimate of machining time to 9.7 min. The complexity increases to C = 238 bits, and raises the parametric estimate to t = 8.7 min (− 10% error); the simplified model does not pick up any difference, and its estimate is again t' = 7.1 min (− 27% error).

The results suggest that complexity alone is not a sufficiently accurate predictor of machining time when a broader range of machining processes is involved. In fact, it is mainly correlated with the number and relative difficulty of machining operations, but not with the amount of cutting involved by the size of machined features; moreover, it is related to non-productive operations but not to part handling or material properties. The addition of three more predictors has made it possible to build a multiple regression model with average errors of around 25% of the estimated times compared to a feature-based method used as a baseline. It is believed that such an accuracy is sufficient to drive several types of design activities (benchmarking, make-or-buy, design for manufacturing), considering the reduced calculation effort compared to more detailed estimation methods.

In the attempt to avoid the need for detailed process information, the parametric method is obviously subject to an additional source of uncertainty related to shop-floor conditions, e.g. the flexibility and the performance of available machine tools. The estimates provided by the regression model should constantly be taken as referring to state-of-the-art equipment and tooling for medium-volume production. In perspective, a case in which the parametric method is found to be in strong disagreement with more accurate methods might even suggest that some redesign or alternative process routes may be advisable.

Conclusions

In summary, the parametric method proposed for the estimation of machining time can be justified with the following considerations:

- Compared to existing parametric methods, it is neither limited to particular types of products (e.g. rotational parts) nor bound to a common formulation for a wide variety of processes, while providing an adequate compromise in terms of estimation accuracy.
- Compared to feature-based methods, it does not need a separate estimation of machining time from removal rates for each part feature.
- Compared to engineering build-up estimation or hybrid methods, it does not need detailed process planning or selection of cutting parameters.
- Compared to advanced methods (e.g. neural networks), it needs smaller datasets in which the consistency in the estimation assumptions can be better controlled.

Future developments will try to overcome some limitations of this work. The primary objective will be to reduce the estimation error while keeping the procedure simple and fast enough. In the search for further predictors, the concept of entropy will possibly be investigated as an alternative to information content. The integration of machine learning methodologies could possibly lead to a reduction of random errors, while calibration to suit industrial cases will make it possible to reduce systematic errors for use in specific companies. This will also allow a real validation of the method, which is currently limited to a comparison with a more accurate method from literature. The applicability of the parametric method will also have to be improved by extending the complexity measure to modern standards of geometric tolerancing, as well as by software implementation that can automate the calculation of predictors and provide effective visualizations of the results.

Funding Open Access funding provided by Politecnico di Milano.

Availability of data and materials The author confirms that the data supporting the findings of this study are available within the paper.

Compliance with ethical standards

Conflict of interest The author declare that they have no conflict of interest.

Open Access This article is licensed under a Creative Commons Attribution 4.0 International License, which permits use, sharing, adaptation, distribution and reproduction in any medium or format, as long as you give appropriate credit to the original author(s) and the source, provide a link to the Creative Commons licence, and indicate if changes were made. The images or other third party material in this article are included in the article's Creative Commons licence, unless indicated otherwise in a credit line to the material. If material is not included in the article's Creative Commons licence and your intended use is not permitted by statutory regulation or exceeds the permitted use, you will need to obtain permission directly from the copyright holder. To view a copy of this licence, visit http://creativecommons.org/licenses/by/4.0/.

References

AACE International Recommended Practice 10S-90. (2019). *Cost engineering terminology*. Morgantown: AACE International.

Ashby, M., Shercliff, H., & Cebon, D. (2007). *Materials: engineering, science, processing and design*. Amsterdam: Butterworth-Heinemann.

Atia, M. R. A., Khalil, J., & Mokhtar, M. (2017). A cost estimation model for machining operations: An ANN parametric approach. *Journal of Al-Azhar University Engineering Sector, 12*(44), 878–885.

Ben-Arieh, D. (2000). Cost estimation system for machined parts. *International Journal of Production Research, 38*(17), 4481–4494.

Boothroyd, G., Dewhurst, P., & Knight, W. A. (2011). *Product design for manufacture and assembly* (3rd ed.). Boca Raton: CRC Press.

Budiono, H. D. S., Kiswanto, G., & Soemardi, T. P. (2014a). Method and model development for manufacturing cost estimation during the early design phase related to the complexity of the machining processes. *International Journal of Technology, 2,* 183–192.

Budiono, H. D. S., Sholeh, M., Kiswanto, G., & Soemardi, T. P. (2014b). Application of semi automatic model of product complexity index calculation by identification and recognition of geometric features information. *Applied Mechanics and Materials, 493,* 576–582.

Cavalieri, S., Maccarrone, P., & Pinto, R. (2004). Parametric vs. neural network models for the estimation of production costs: A case study in the automotive industry. *International Journal of Production Economics, 91,* 165–177.

Creese, R. C., Adithan, M., & Pabla, B. S. (1992). *Estimating and costing for the metal manufacturing industries*. New York: Marcel Dekker.

D'Addona, D. M., Ullah, A. M. M. S., & Matarazzo, D. (2017). Tool wear prediction and pattern recognition using artificial neural network and DNA-based computing. I. *Intelligent Manufacturing, 28,* 1285–1301.

Drozda, T. J., Wick, C., Mitchell, P., Bakerjian, R., Benedict, J. T., & Veilleux, R. F. (1983). *Tool and manufacturing engineers handbook, Vol. 6: Design for manufacturability*. Dearborn: Society of Manufacturing Engineers.

Duverlie, P., Castelain, J., & M. . (1999). Cost estimation during design step: Parametric method versus case based reasoning method. *The International Journal Advanced Manufacturing Technology, 15,* 895–906.

Ehrlenspiel, K., Kiewert, A., & Lindemann, U. (2007). *Cost-efficient design*. Berlin: Springer.

ElMaraghy, W. H., & Urbanic, R. J. (2003). Modeling of manufacturing systems complexity. *CIRP Annals Manufacturing Technology, 52*(1), 363–366.

ElMaraghy, W. H., & Urbanic, R. J. (2004). Assessment of manufacturing operational complexity. *CIRP Annals Manufacturing Technology, 53*(1), 401–406.

ElMaraghy, W., ElMaraghy, H., Tomiyama, T., & Monostori, L. (2012). Complexity in engineering design and manufacturing. *CIRP Annals Manufacturing Technology, 61,* 793–814.

EN ISO 286-1. (2010). *Geometrical product specifications (GPS): ISO code system for tolerances on linear sizes. Part 1: Basis of tolerances, deviations and fits*. Brussels: European Committee for Standardization.

Esawi, A. M. K., & Ashby, M. F. (2003). Cost estimates to guide preselection of processes. *Materials and Design, 24*(8), 605–616.

Farag, M. M. (2014). *Materials and process selection for engineering design* (3rd ed.). Boca Raton: CRC Press.

Feng, C. X., Kusiak, A., & Huang, C. C. (1996). Cost evaluation in design with form features. *Computer-Aided Design, 28*(11), 879–885.

Foussier, P. (2006). *From product description to cost: A practical approach, Vol. 1: The parametric approach*. Berlin: Springer.

GAO-09-3SP. (2009). *GAO Cost estimating and assessment guide: Best practices for developing and managing capital program costs*. Washington: US Government Accountability Office.

Garcia-Crespo, A., Ruiz-Mezcun, B., Lopez-Cuadrado, J. L., & Gonzalez-Carrasco, I. (2011). A review of conventional and knowledge-based systems for machining price quotation. *Journal of Intelligent Manufacturing, 22,* 823–841.

Geiger, T. S., & Dilts, D. M. (1996). Automated design-to-cost: Integrating costing into the design decision. *Computer-Aided Design, 28*(6/7), 423–438.

Germani, M., Mandolini, M., & Cicconi, P. (2011). Manufacturing cost estimation during early phases of machine design. In: *International conference on engineering design, Copenhagen*.

Grewal, S., & Choi, C. K. (2005). An integrated approach to manufacturing process design and costing. *Concurrent Engineering: Research and Applications, 13*(3), 199–207.

Heo, E. Y., Kim, D. W., Kim, B. H., & Chen, F. F. (2006). Estimation of NC machining time using NC block distribution for sculptured surface machining. *Robotics and Computer-Integrated Manufacturing, 22,* 437–466.

Hoult, D. P., Meador, C. L. (1996). Predicting product manufacturing costs from design attributes: A complexity theory approach. In: *SAE Transactions*, 960003. SAE International.

Huang, G. (2007). *Cost modeling based on support vector regression for complex products during the early design phases*. Ph.D. thesis, Virginia Polytechnic Institute and State University.

ISO/IEC 2382-16 (1996). *Information technology—Vocabulary. Part 16: Information theory*. Geneve: ISO/IEC.

Jha, N. K. (1996). Probabilistic cost estimation in advance of production in a computerized manufacturing system through stochastic geometric programming. *Computers and Industrial Engineering, 30*(4), 809–821.

Jung, J. Y. (2002). Manufacturing cost estimation for machined parts based on manufacturing features. *Journal of. Intelligent Manufacturing, 13,* 227–238.

Kerbrat, O., Mognol, P., & Hascoët, J. Y. (2010). Manufacturing complexity evaluation at the design stage for both machining and

layered manufacturing. *CIRP Journal of Manufacturing Science and Technology, 2*(3), 208–215.

Lenau, T., & Haudrum, J. (1994). Cost evaluation of alternative production methods. *Materials and Design, 15*(15), 235–247.

Liu, Y., Basson, A. H. (2001). Interfacing CAD and manufacturing cost estimation software using COM/OLE. In: *International CIRP design seminar, Stockholm* (pp. 289–294).

Liu, C., Li, Y., Wang, W., & Shen, W. (2013). A feature-based method for NC machining time estimation. *Robotics and Computer-Integrated Manufacturing, 29,* 8–14.

Locascio, A. (2000). Manufacturing cost modeling for product design. *International Journal of Flexible Manufacturing Systems, 12,* 207–217.

Lovejoy, W., Fixson, S., & Jackson, S. (2005). *Product costing guidelines, integrated product development*. Ann Arbor: University of Michigan.

Loyer, J. L., Henriques, E., Fontul, M., & Wiseall, S. (2016). Comparison of machine learning methods applied to the estimation of manufacturing cost of jet engine components. *International Journal of Production Economics, 178,* 109–119.

Mahmoud, M. A. M. (1979). *The costing of turned components at the design stage*. Ph.D thesis, Loughborough University.

Maropoulos, P. G., Baker, R. P., & Paramor, K. Y. G. (2000). Integration of tool selection with design, Part 2: Aggregate machining time estimation. *Journal of Materials Processing Technology, 107,* 135–142.

Mauchand, M., Siadat, A., Bernard, A., & Perry, N. (2008). Proposal for tool-based method of product cost estimation during conceptual design. *Journal of Engineering Design, 19*(2), 159–172.

Machinability Data Center. (1980). *Machining data handbook* (3rd ed.). Cincinnati: Metcut Research Association.

Mislick, G. K., & Nussbaum, D. A. (2015). *Cost estimation: Methods and tools*. Hoboken: Wiley.

Molcho, G., Cristal, A., & Shpitalni, M. (2014). Part cost estimation at early design phase. *CIRP Annals Manufacturing Technology, 63,* 153–156.

Muter, S. (1993). *Cost comparison of alternate designs: an information based model*. M.Sc thesis, Massachusetts Institute of Technology.

NASA (2015). *Cost estimating handbook*, version 4, NP-2015-03-1473-HQ, Washington.

Nasr, E. A., & Kamrani, A. K. (2007). *Computer based design and manufacturing*. Berlin: Springer.

Niazi, A., Dai, J. S., Balabani, S., & Seneviratne, L. (2006). Product cost estimation: technique classification and methodology review. *ASME Journal of Manufacturing Science and Engineering, 128,* 563–575.

Ou-Yang, C., & Lin, T. S. (1997). Developing an integrated framework for feature-based early manufacturing cost estimation. *International Journal of Advanced Manufacturing Technology, 13,* 618–629.

Perera, G. I. P. (2014). Operation based cost assessment of machined parts in the early design stages. *Engineering, 8,* 57–65.

Polgar, K. C. (1996). *Simplified time estimation for basic machining operations*. M.Sc Thesis, Massachusetts Institute of Technology.

Qian, L., & Ben-Arieh, D. (2008). Parametric cost estimation based on activity-based costing: A case study for design and development of rotational parts. *International Journal of Production Economics, 113,* 805–818.

Rao, A. R., Keane, A. J., Scanlan, J. P. (2005). A cost based methodology for design optimization. In: *ASME int. design engineering technical conf., Long Beach, CA, DETC2005-84685*.

Rehman, S., & Guenov, M. D. (1998). A methodology for modelling manufacturing costs at conceptual design. *Computers and Industrial Engineering, 35*(3/4), 623–626.

Roberts, C., Hubele, N., Henderson, M., Stage, R., & Perez, J. (1997). Manufacturing evaluation using resource-based, template-free features. *Journal of Intelligent Manufacturing, 8,* 323–331.

Rodriguez-Alabanda, O., Guerrero-Varcas, G., & Romero, P. E. (2019). Machining time estimation using the geometric features of the 2.5D pocket contour. *Procedia Manufacturing, 41,* 508–515.

Roy, R., & Sackett, P. (2003). *Cost engineering: The practice and the future, CASA/SME Blue Books Series*. Southfield: Society of Manufacturing Engineers.

Roy, R., Souchoroukov, P., & Shehab, E. (2011). Detailed cost estimating in the automotive industry: Data and information requirements. *International Journal of Production Economics, 133,* 694–707.

Rush, C., Roy, R. (2000). Analysis of cost estimating processes used within a concurrent engineering environment throughout a product life cycle. In: *ISPE int. conf. on concurrent engineering: Research and applications, Lyon* (pp. 58–67).

Rush, C., & Roy, R. (2001). Expert judgement in cost estimating: modelling the reasoning process. *Concurrent Engineering: Research and Applications, 9*(4), 271–284.

Shannon, C. E. (1948). A Mathematical Theory of Communication. *Bell System Technical Journal, 27*(3), 21623–21656.

Shebab, E. M. (2001). *An intelligent knowledge-based cost modelling system for innovative product development*. Ph.D thesis, De Montfort University.

Shebab, E. M., & Abdalla, H. S. (2001). Manufacturing cost modelling for concurrent product development. *Robotics and Computer-Integrated Manufacturing, 17,* 341–353.

Shukla, A., Jani, S. S., & Acharya, G. D. (2015). Feature based time estimation: a review. *Journal of Production Research and Management, 5*(3), 1–6.

Shukla, A., Jani, S. S., & Acharya, G. D. (2016). Developing a CAM estimator for feature based machining time estimation in VB environment. *International Journal of Advance Research and Innovative Ideas in Education, 2*(3), 527–535.

Siller, H., Rodriguez, C. A., & Ahuett, H. (2006). Cycle time prediction in high-speed milling operations for sculptured surface finishing. *Journal of Materials Processing Technology, 174,* 355–362.

Smith, A. E., & Mason, A. K. (1997). Cost estimation predictive modeling: Regression versus neural network. *Engineering Economist, 42*(2), 137–161.

So, B. S., Jung, Y. H., Park, J. W., & Lee, D. W. (2007). Five-axis machining time estimation algorithm based on machine characteristics. *Journal of Materials Processing Technology, 187–188*(37), 40.

Stockton, D., & Wang, Q. (2004). Developing cost models by advanced modelling technology. *Proceedings of the Institution of Mechanical Engineers, Part B: Journal of Engineering Manufacture, 218,* 213–224.

Suh, N. P., Bell, A. C., & Gossard, D. C. (1978). On an axiomatic approach to manufacturing and manufacturing systems. *ASME Journal of Engineering for Industry, 100*(2), 127–130.

Suh, N. P. (1990). *The principles of design*. Oxford: Oxford University Press.

Swift, K. G., & Booker, J. D. (2013). *Manufacturing process selection handbook*. Amsterdam: Butterworth-Heinemann.

Tanner, J. P. (2006). Practical cost estimating for manufacturing. In R. Crowson (Ed.), *Factory operations: Planning and instructional methods*. Boca Raton: CRC Press.

Ullah, A. M. M. S., Fuji, A., Kubo, A., Tamaki, J., & Kimura, M. (2015). On the surface metrology of bimetallic components. *Machining Science and Technology, 19*(2), 339–359.

Ullah A. M. M. S., Harib K. H., Awda M., Zenieh S. S. (2009). Application of axiomatic design principles for process planning in milling. SAE Paper 2009-01-0792, SAE International.

Ulrich, K. T., & Eppinger, S. D. (2007). *Product design and development* (4th ed.). New York: McGraw-Hill.

Urbanic, R. J., & ElMaraghy, W. H. (2006). Modeling of manufacturing process complexity. In H. A. ElMaraghy & W. H. ElMaraghy (Eds.), *Advances in design, Springer Series in Advanced Manufacturing*. London: Springer.

Watson, P., Curran, R., Murphy, A., & Cowan, S. (2006). Cost estimation of machined parts within an aerospace supply network. *Concurrent Engineering: Research and Applications., 14*(1), 17–26.

Weustink, I. F., ten Brinke, E., Streppel, A. H., & Kals, H. J. J. (2000). A generic framework for cost estimation and cost control in product design. *Journal of Materials Processing Technology, 103,* 141–148.

Wilson, D. R. (1980). *An exploratory study of complexity in axiomatic design*. Ph.D Thesis, Massachusetts Institute of Technology.

Zhang, Y. F., & Fuh, J. Y. H. (1998). A neural network approach for early cost estimation of packaging products. *Computers and Industrial Engineering, 34*(2), 433–450.