Advances in Production Engineering & Management

Identification Statement

APEM Editorial

Advances in Production Engineering & Management

Contents

Scope and topics

Advances in Production Engineering & Management (*APEM journal*) is an interdisciplinary refereed international academic journal published quarterly by the *Production Engineering Institute* at the *University of Maribor*. The main goal of the *APEM journal* is to present original, high quality, theoretical and application-oriented research developments in all areas of production engineering and production management to a broad audience of academics and practitioners. In order to bridge the gap between theory and practice, applications based on advanced theory and case studies are particularly welcome. For theoretical papers, their originality and research contributions are the main factors in the evaluation process. General approaches, formalisms, algorithms or techniques should be illustrated with significant applications that demonstrate their applicability to real-world problems. Although the *APEM journal* main goal is to publish original research papers, review articles and professional papers are occasionally published.

Fields of interest include, but are not limited to:

Additive Manufacturing Processes
Advanced Production Technologies
Artificial Intelligence in Production
Assembly Systems
Automation
Big Data in Production
Computer-Integrated Manufacturing
Cutting and Forming Processes
Decision Support Systems
Discrete Systems and Methodology
e-Manufacturing
Evolutionary Computation in Production
Fuzzy Systems
Human Factor Engineering, Ergonomics
Industrial Engineering
Industrial Processes
Industrial Robotics
Intelligent Manufacturing Systems
Joining Processes
Knowledge Management
Logistics in Production

Machine Learning in Production
Machine Tools
Machining Systems
Manufacturing Systems
Materials Science, Multidisciplinary
Mechanical Engineering
Mechatronics
Metrology in Production
Modelling and Simulation
Numerical Techniques
Operations Research
Operations Planning, Scheduling and Control
Optimisation Techniques
Project Management
Quality Management
Risk and Uncertainty
Self-Organizing Systems
Statistical Methods
Supply Chain Management
Virtual Reality in Production

Production lot-sizing decision making considering bottleneck drift in multi-stage manufacturing system

Zhou, F.L.[a,*], **Wang, X.**[a,b,*], **He, Y.D.**[a], **Goh, M.**[c]

[a]Department of Industrial Engineering, Chongqing University (CQU), Chongqing, P.R. China
[b]State Key Laboratory of Mechanical Transmission, Chongqing University (CQU), Chongqing, P.R. China
[c]Department of Analytics and Operations, NUS Business School and The Logistics Institute-Asia Pacific, National University of Singapore (NUS), Singapore

ABSTRACT

Under a capacity constrained multi-product manufacturing system, the products are usually prepared and produced in lots. As a lot-sizing strategy is critical for effective production and high productivity, this encourages practical and research interest in the strategic batch sizing decision for a minimum procedure time in an order-to-delivery (OTD) operating environment. While the lot-sizing plan can be formed by studying the manufacturing parameters of the established bottleneck procedure, for a multi-stage manufacturing system, the bottleneck is not fixed and fluctuates with the production rate or batch size. This paper proposes a lot-sizing strategy to determine the optimal lot-size for each class of products taking bottleneck drifting into consideration. A queuing network analyser (QNA) method is employed to deal with the non-linear mixed integer programming model targeting at the total flow time minimization of the system. A practical case is presented and solved using the proposed method, and the results are validated with Flexsim, a simulation model.

1. Introduction

In the modern manufacturing era, enterprises are increasingly focusing on optimizing the total flow time under an order-to-delivery (OTD) environment [1]. With the proliferation in product variety, many products are usually processed and produced in the same production system with a view to improving the production efficiency and reducing cost. In the multi-product production system, the productivity and lead time of the manufacturing workshop is directly affected by the batch size of the product [2, 3]. With mass production, the productivity can be increased albeit with a longer lead time for the lot production; with low-volume production, the effect is reversed. In the literature, this phenomenon of large lot sizes will cause long lead times known as the batching effect. As the lot size decreases (the jobshop context), the lead time will also decrease, but once a minimal lot size is reached a further reduction in lot size will cause high traffic intensities resulting in longer lead times, called the saturation effect [4]. Therefore, to meet the shortest delivery lead time possible, it is often necessary to determine the right type and produce products in an appropriate quantity.

There are many models, in the literature, focusing on the complex relationship between batch, lead time, and work-in-process (WIP) [5-7]. Karmarkar [5] had studied the influences of batch (lot-size) on manufacturing lead times, and reported that the best lot sizes, high capacity

utilization exact a high price in lead time and WIP. The relationship between batch size and lead-time variability was investigated by Kuik and Tielemans [6], who concluded that a minimization of the average queue delay or the average time in system performance measure, would not result in minimum lead-time variability. Vaughan [7] developed a comprehensive process lot sizing/order point model, and found that it was more efficient to adjust the safety stock and lead time than to change the shortage penalty parameter solely through adjusting the safety stock. Recently, some scholars argued that the lead time and productivity levels of the production system were determined by the bottleneck equipment, and the optimal processing of a single batch was studied and analysed [8-12]. Koo and Koh [8] proposed a batch decision optimization model that included a variety of product lines with the same preparation time by targeting maximum profit as the objective. Similarly, an extended model for optimizing the lot size in various production lines with different process and preparation times was formulated [9]. Liu [13] analysed the possibility of resources becoming the bottleneck through the bottleneck drift index, and calculated the optimal lot size and lead time of the bottleneck resource targeting the non-value added time and the difference in the manufacturing unit processing rate minimum as the optimization goal. However, they did not consider the bottleneck drifting caused by lot sizes.

In reality, in a relatively balanced production system with a large variety of products, the fluctuation of the product portfolio and the lot sizes can alter the load level of the resources, which might lead to bottleneck drifting. As a result, the lot size and lead time found in previous studies may not be optimal. Adacher and Cassandras [4] studied the optimal lot sizing problem with the example of two products and two processing operations, using the substitution method and the random comparison algorithm. Glock [14] solved the optimal lot size problem for a multi-stage manufacturing system by studying the relationship between the quantity and the total cost. Amy [15] proposed a mixed nonlinear integer programming model to solve the lot-sizing optimization problem with multiple suppliers in multiple periods considering quantity discounts, and a Genetic Algorithm (GA) is developed to minimize the total related cost. Therefore, this paper investigates the lot-sizing problem with bottleneck drifting. As the lot-sizing decision making is related to productivity, which influences the efficient output of the manufacturing system [1], this paper seeks to examine the lot-sizing problem for different experimental productivity scenarios.

In this paper, we apply QNA (queuing network analysis) [16] to establish a lot-size optimization model on a multi process production system, to decide the optimal lot sizes in the production of multiple products. The objective of the batch optimization model is to produce products with the shortest lead time according to the demand. QNA is an accurate queuing network analysis model developed by Bell Laboratories in the US [16]. QNA is used to analyze the queuing network by estimating the distribution and variation coefficient of each arrival process and each service time. QNA is suitable for solving complex queuing network problems, given its low computational complexity. Given the relationship between the processes of a multi operation system, we apply the QNA method to determine the waiting time of the tandem process queue, with the total process time as the shortest objective function, such that the changes caused by the bottleneck of the product batch will not affect the optimization objective.

The next section presents the studied problem and the assumptions thereof. In Section 3, the non-linear programing model considering bottleneck drifting is formulated and a four-step algorithm is designed by a traversal operation and implemented to deal with the lot-sizing optimization model. A practical case is presented in Section 4 to validate the proposed model. Lastly, we conclude with some remarks.

2. Problem description and assumptions

As shown in Fig. 1, the production line is a system with many work procedures. Each service station represents a processing or assembly operation. All kinds of products enter and leave independently with a minimum unit. Each batch of products follows a first-come-first-served (FCFS) queuing strategy.

The assumptions are as follows:

- The equipment utilization rate is less than 1, and productivity cannot exceed the production capacity;
- Each service station shows the same product capacity for the same products;
- The proportion of each product is known, and each class of products has a fixed manufacturing procedure.
- The processing times of various products are different and independent;
- Each lot of products arrives as a Poisson process;
- The preparation time before lot-size production, and the preparation time of each product is different and independent.
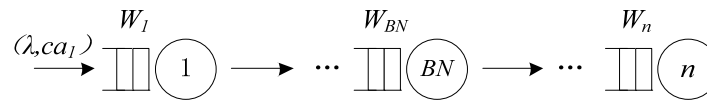
Fig. 1 Production system with multi-stage procedures

The parameters are denoted and described as follows:

Notation	Interpretation
F	Total processing time, units/period
$\bar{F}_j$	Average processing time in process j, units/period
$\bar{w}_j$	Average queuing time on equipment j, units/period
$\bar{s}_j$	Average processing time on equipment j , units/period
x_i	Productivity of product item i, ($i = 1,2,3,...,m$); $X = \sum_i x_i$
q_i	Lot size of product item i
r_i	Proportion of product item i, ($\sum_i r_i = 1$)
p_{ij}	Processing time on equipment j of product item i
τ_{ij}	Setup time on equipment j of product item i
s_{ij}	Total time required for product item i ($s_{ij} = q_i p_{ij} + \tau_{ij}$)
μ_{ij}	Service rate of a certain batch item i ($\mu_{ij} = 1/s_{ij}$)
ρ_j	Logistics intensity of equipment j
λ_i	Average number of arrivals of product i per time unit (arrival rate)
cs_j	Variation coefficient of processing time on equipment j
ca_j	Variation coefficient of product inter-arrival time on equipment j
cd_j	Variation coefficient of product inter-left time on equipment j

3. Model and algorithm

The non-linear mixed integer programming (MIP) model is formulated and constructed by minimizing the total flow time with the QNA technique. The algorithm procedures through the traversal operation are designed and implemented in this section.

3.1 Objective function

The optimization objective of the described problem is targeted as the total production time minimization considering production capacity and other common constraints. The model is formulated as follows.

$$\min F = \sum_{j=1}^{n} \bar{F}_j = \sum_{j=1}^{n} (\bar{w}_j + \bar{s}_j) \tag{1}$$

$$s.t. \sum_{i=1}^{m} (x_i p_{ij} + x_i \tau_{ij}/q_i) < 1 \tag{2}$$

$$u_j = \sum_i u_{ij}, i = 1, 2, \dots, m \tag{3}$$

$$u_{ij} = x_i p_{ij} \tag{4}$$

$$v_j = \sum_i u_{ij}, i = 1, 2, \dots, m \tag{5}$$

$$v_{ij} = x_i \tau_{ij}/q_i \tag{6}$$

$$q_i \leq \bar{q}_\iota, q_i \epsilon Z^+ \tag{7}$$

$$x_i, q_i > 0 \tag{8}$$

where parameters u and v are the processing utilization rate and setup utilization rate respectively. Eq. 1 is the objective function for the total production time minimization. Eqs. 2 to 8 are the relevant constraints on productivity, order, and practical production.

3.2 Estimation of service time

In the production system, the arrival of a variety of products can be regarded as a Poisson process. It is reasonable to simulate the multi-process production system with the tandem queuing model, where $\bar{F}_j$ denotes the mean flow time of a product in processing j, which includes the waiting and processing times. The mean processing time of a unit product in process j is found from Eq. 9.

$$\bar{s}_j = \sum_i r_i \left(p_{ij} q_i + \tau_{ij} \right) \tag{9}$$

In a serial queuing system, the arrival processes are determined by the output of the previous process in addition to the first process. In this paper, the QNA method is used to estimate the arrival process variation coefficient of each node, which is regarded as an update process. Therefore, the queuing time in a steady state can be obtained [17]. Suresh and Whitt (1990) proposed the estimation method for the arrival process variation coefficient, and improved the accuracy of the estimation of the queuing time when the traffic intensity is 0.9 [18]. The improved coefficient of the arrival process is presented in Eq. 10.

$$cd_j^2 = \left(1 - \rho_j^2 \left(1 - \rho_{j+1}^{10} \right) \right) ca_j^2 + \rho_j^2 \left(1 - \rho_{j+1}^{10} \right) cs_j^2$$
$$ca_j^2 = cd_{j-1}^2 \tag{10}$$

The queuing time by the QNA technique is estimated from Eq. 11 and Eq. 12.

If $ca_j^2 < 1$

$$\bar{w}_j = \frac{\bar{s}_j(ca_j^2 + cs_j^2)\sum_i x_i(p_{ij} + \tau_{ij}/q_i)}{2[1 - \sum_i x_i(p_{ij} + x_i\tau_{ij}/q_i)]} \exp\left\{ -\frac{2[1 - \sum_i x_i(p_{ij} + \tau_{ij}/q_i)](1 - ca_j^2)}{3\sum_i x_i(p_{ij} + \tau_{ij}/q_i)(ca_j^2 + cs_j^2)} \right\} \tag{11}$$

If $ca_j^2 \geq 1$

$$\bar{w}_j = \frac{\bar{s}_j p_j(ca_j^2 + cs_j^2)}{2(1 - p_j)} = \frac{\bar{s}_j(ca_j^2 + cs_j^2)\sum_i x_i(p_{ij} + \tau_{ij}/q_i)}{2[1 - \sum_i x_i(p_{ij} + x_i\tau_{ij}/q_i)]} \tag{12}$$

3.3 Productivity analysis

As the proportion of products r_i is known, the productivity x_i of an item i can be found from the total productivity X. The capacity constraints are illustrated in Eq. 13. From the upper bound on the lot-size of each class of products $\bar{q}_\iota$, we deduce the upper bound of X as shown in Eq. 14.

$$X \sum_i (r_i p_{ij} + r_i \tau_{ij}/q_i) < 1 \tag{13}$$

$$\bar{X} = \left\lfloor 1/\sum_i (r_i p_{ij} + r_i \tau_{ij}/q_i) \right\rfloor \tag{14}$$

where $\lfloor X \rfloor$ is the largest integer less than or equal to X.

3.4 Algorithm

To establish the optimized lot-sizing of each kind of product with different productivity levels, we targeted the total production time as the objective function and formulated a non-linear MIP model. To determine the optimal lot size, we introduce the algorithm for dealing with this issue. First, the upper bound on productivity is analysed based on the previous section. Then, the optimal solution is searched by a traversal operation. Fig. 2 shows the steps of the algorithm. As Fig. 2 shows, there are four steps in the proposed algorithm.
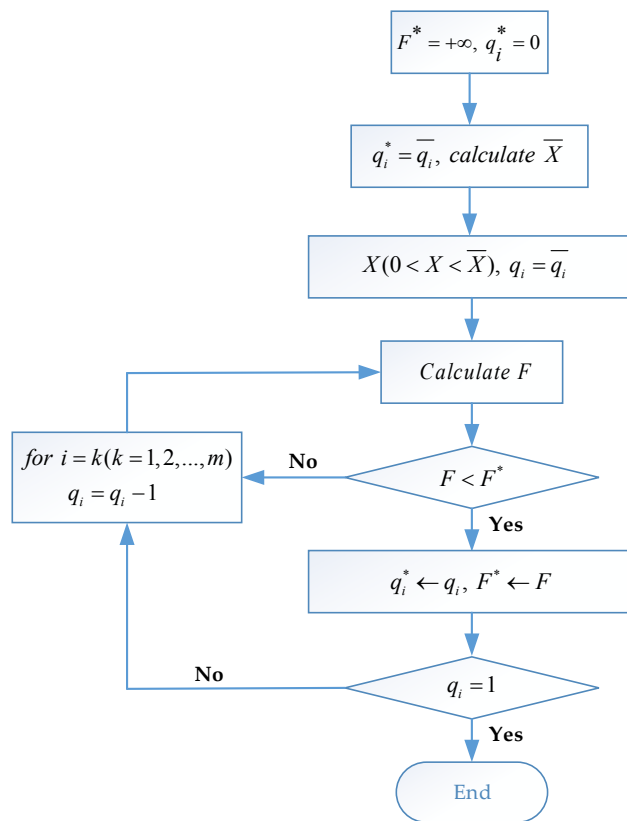
Four stages

Satege 1: Initialization. Initialize objective value setting: $F^* = +\infty, q_i^* = 0$. We obtain the upper bound of total productivity $\bar{X}$ for different kinds of products by Eq. 14.

Stage 2: Total production time F (including processing and waiting) computed by Eqs. 9 to 12.

Stage 3: If $F < F^*$, then $q_i^* = q_i$. If $F > F^*$, then for $i = k$ ($k = 1, 2, \ldots, m$), $q_i = q_i - 1$, and return to stage 2.

Stage 4: If $q_i = 1$, stop. The best lot-sizing solution is q_i^*, and the minimum total production time is denoted as F^*. If $q_i \neq 1$, then return to stage 3.

Fig. 2 Two-step optimization algorithm for multi-stage production system

4. Case study and simulation comparison

4.1 Case solutions

We consider a production system with five main procedures, and four classes of products need to be processed (Fig. 3). The information on the different parameters are presented in Table 1 such as processing time, upper bound of each class of products and product ratio. Suppose the arrival times of all products are regarded as Poisson, then $ca_1 = 1$, $cs_j^2 = 0.5$. The best solution of the lot-sizing for each class of products is calculated using the proposed algorithm procedure.
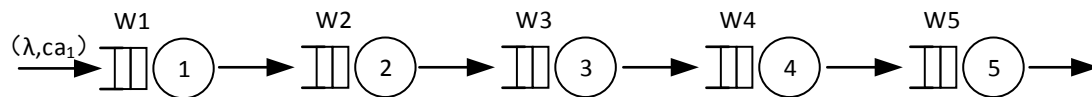
Fig. 3 Serial production system with five procedures

Table 1 Parameter values of four products in serial manufacturing system

Parameters		Product 1	Product 2	Product 3	Product 4
Waiting time	τ_{i1}	0.018	0.012	0.015	0.01
	τ_{i2}	0.01	0.02	0.02	0.01
	τ_{i3}	0.02	0.01	0.02	0.01
	τ_{i4}	0.015	0.005	0.015	0.01
	τ_{i5}	0.012	0.012	0.015	0.01
Processing time	p_{i1}	0.008	0.01	0.006	0.006
	p_{i2}	0.008	0.01	0.006	0.006
	p_{i3}	0.01	0.008	0.005	0.005
	p_{i4}	0.01	0.01	0.006	0.005
	p_{i5}	0.01	0.009	0.005	0.006
Product ratio	r_i	0.3	0.2	0.1	0.4
Upper bound	$\bar{q_i}$	30	30	50	50

From the algorithm, we obtain the best lot sizes under different productivity conditions. For instance, with a productivity of 90 %, the best solution of the four products is 7, 6, 10 and 7 units respectively, and with minimum total production time of 1.230 units. The fifth procedure is the bottleneck with the highest utilization rates of 84.02 %, as shown by Table 2. The results are illustrated in Table 2 for the various productivity situations.

From Table 2, the best lot-sizing solution for each class of products obtained from our algorithm suggests the following findings:

- As the productivity rate declines, the best lot-size for each class of products and the total production time decrease.
- In a relatively balanced manufacturing system, with a variation in lot-sizing, the bottleneck of the manufacturing system drifts as we imagined.
- As the productivity rate declines, the utilization rate of the machine fluctuates slightly, rather than declining. The reason for the utilization rate increases is that the waiting time has been increased with the much more preparation operations.

Table 2 Results for different productivity situations

Produc-tivity (%)	Utilization rate of the machine (%)					Total time (days)	Lot size (piece)			
	Machine 1	Machine 2	Machine 3	Machine 4	Machine 5		Prod 1	Prod 2	Prod 3	Prod 4
90	83.64	83.40	81.56	82.18	84.02	1.230	7	6	10	7
89	82.85	82.67	80.85	81.41	83.24	1.173	7	6	9	7
88	81.92	81.74	79.94	80.50	82.30	1.121	7	6	9	7
87	82.11	81.43	80.27	80.52	82.11	1.073	6	6	9	7
86	81.35	80.74	79.59	79.77	81.35	1.028	6	6	8	7
85	81.89	81.74	80.04	79.94	81.89	0.983	6	5	8	6
84	80.93	80.78	79.10	79.00	80.93	0.941	6	5	8	6
83	79.96	79.82	78.16	78.05	79.96	0.903	6	5	8	6
82	79.22	79.15	77.51	77.33	79.22	0.869	6	5	7	6
81	79.71	78.99	78.18	77.61	79.23	0.835	5	5	7	6
80	79.79	79.09	78.29	77.71	79.31	0.803	5	5	7	5
79	78.80	78.10	77.31	76.74	78.32	0.772	5	5	7	5
78	78.74	78.67	77.11	76.16	78.27	0.743	5	4	7	5

4.2 Simulation verification

A simulation test is implemented to validate the effectiveness and validity of the proposed model. When the variation coefficient of the processing time is $cs_j^2 > 1$, we supposed the processing time complies with the H_2 distribution. When $cs_j^2 = 1$, we obtain the exponential distribution. When $cs_j^2 > 1$, the processing time can be treated as an Erlang distribution [18]. The manufacturing process is simulated by the Flexsim software with the original data information in Table 1, and the simulation is run for ten thousand days (10 times). Fig. 4 compares the simulated total production time and the proposed model.

From Fig. 4, the deviation result between the simulation model and programming model is about 12%. From Wu's research [19], when $cs^2 < 1$, the estimation variation of waiting time for multi-stage queuing system by the QNA method is 6.3%. As there are five procedures in the manufacturing system, the variation of the serial system is much more than the single machine (6.3%) due to the deviation accumulation. As the waiting time estimation method focuses on a single machine, the accuracy of the estimation of waiting time directly influences the performance of the programming model. As for the multi-stage serial manufacturing system with limited procedures, the best solution accuracy of the proposed model is satisfactory and tolerable based on the above comparison analysis.
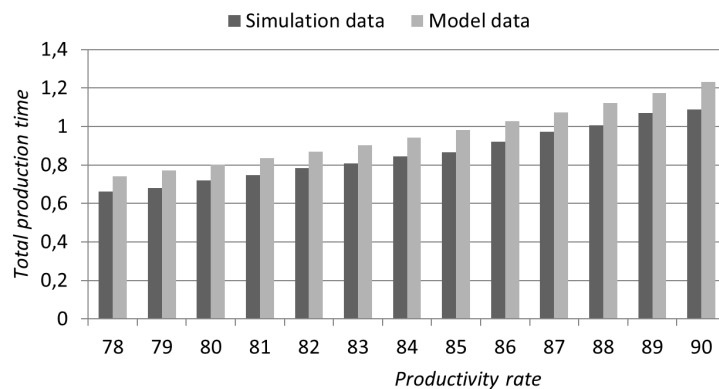
Fig. 4 Comparison of total time for simulation model and programming model

5. Conclusion

In this paper, the non-linear MIP model by the QNA method is proposed to deal with the lot-sizing problem for a multi-stage manufacturing system with multiple classes of products. Not only can the model be applied into a lot-sizing optimization problem of a manufacturing system with fixed bottlenecks, but also this model is applicable to a sensitive production system whose bottleneck fluctuates with the variation of the product combination and lot-sizing. The model is validated against a simulation model run by the *Flexsim* software, and it demonstrates excellent performance. However, from the comparison results, the model shows tolerable variation, which stimulates us to improve the accuracy of the model by revising the waiting time parameters. In future, the lot-sizing decision making problem which jointly considers bottleneck drift and unpredictable events (machine malfunction or failure, and stochastic occurrences) can be studied.

Acknowledgements

References

[1] Zhou, F.-L., Wang, X., Lin, Y. (2016). Production effectiveness-based system reliability calculation of serial manufacturing with checking machine, *Journal of Computers*, Vol. 27, No. 3, 201-211, doi: 10.3966/199115592016102703017.

[2] Potts, C.N., Van Wassenhove, L.N. (1992). Integrating scheduling with batching and lot-sizing: A review of algorithms and complexity, *Journal of the Operational Research Society*, Vol. 43, No. 5, 395-406, doi: 10.1057/jors.1992.66.

[3] Karimi, B., Fatemi Ghomi, S.M.T., Wilson, J.M. (2003). The capacitated lot sizing problem: A review of models and algorithms, *Omega*, Vol. 31, No. 5, 365-378, doi: 10.1016/s0305-0483(03)00059-8.

[4] Adacher, L., Cassandras, C.G. (2014). Lot size optimization in manufacturing systems: The surrogate method, *International Journal of Production Economics*, Vol. 155, 418-426, doi: 10.1016/j.ijpe.2013.07.026.

[5] Karmarkar, U.S. (1987). Lot sizes, lead times and in-process inventories, *Management Science*, Vol. 33, No. 3, 409-418, doi: 10.1287/mnsc.33.3.409.

[6] Kuik, R., Tielemans, P.F.J. (1999). Lead-time variability in a homogeneous queueing model of batching, *International Journal of Production Economics*, Vol. 59, No. 1-3, 435-441, doi: 10.1016/s0925-5273(98)00029-2.

[7] Vaughan, T.S. (2006). Lot size effects on process lead time, lead time demand, and safety stock, *International Journal of Production Economics*, Vol. 100, No. 1, 1-9, doi: 10.1016/j.ijpe.2004.09.012.

[8] Koo, P.-H., Bulfin, R., Koh, S.-G. (2007). Determination of batch size at a bottleneck machine in manufacturing systems, *International Journal of Production Research*, Vol. 45, No. 5, 1215-1231, doi: 10.1080/00207540600675793.

[9] Koo, P.-H., Koh, S.-G., Lee, W.-S. (2011). Simultaneous determination of lot size and production rate at capacity-constrained multiple-product systems, *Flexible Services and Manufacturing Journal*, Vol. 23, No. 3, 346-359, doi: 10.1007/s10696-011-9091-6.

[10] Kuik, R., Tielemans, P.F.J. (2004). Expected time in system analysis of a single-machine multi-item processing center, *European Journal of Operational Research*, Vol. 156, No. 2, 287-304, doi: 10.1016/s0377-2217(03)00022-5.

[11] Glock, C.H. (2010). Batch sizing with controllable production rates, *International Journal of Production Research*, Vol. 48, No. 20, 5925-5942, doi: 10.1080/00207540903170906.

[12] Maddah, B., Moussawi, L., Jaber, M.Y. (2010). Lot sizing with a Markov production process and imperfect items scrapped, *International Journal of Production Economics*, Vol. 124, No. 2, 340-347, doi: 10.1016/j.ijpe.2009.11.029.

[13] Liu, M., Ling, L., Tang, J. (2013). Research on production lot-size and lead time based on logistics shifting bottleneck in manufacturing shop, *Chinese Mechanical Engineering*, Vol. 24, No. 2, 220-225, doi: 10.3969/j.issn.1004-132X.2013.02.015.

[14] Glock, C.H. (2011). Batch sizing with controllable production rates in a multi-stage production system, *International Journal of Production Research*, Vol. 49, No. 20, 6017-6039, doi: 10.1080/00207543.2010.528058.

[15] Lee, A.H.I., Kang, H.-Y., Lai, C.-M., Hong, W.-Y. (2013). An integrated model for lot sizing with supplier selection and quantity discounts, *Applied Mathematical Modelling*, Vol. 37, No. 7, 4733-4746, doi: 10.1016/j.apm.2012.09.056.

[16] Hopp, W.J., Spearman, M.L., Chayet, S., Donohue, K.L., Gel, E.S. (2002). Using an optimized queueing network model to support wafer fab design, *IIE Transactions*, Vol. 34, No. 2, 119-130, doi: 10.1080/07408170208928855.

[17] Whitt, W. (1983). The queueing network analyzer, *The Bell System Technical Journal,* Vol. 62, No. 9, 2779-2815, doi: 10.1002/j.1538-7305.1983.tb03204.x.

[18] Suresh, S., Whitt, W. (1990). Arranging queues in series: A simulation experiment, *Management Science*, Vol. 36, No. 9, 1080-1091, doi: 10.1287/mnsc.36.9.1080.

[19] Wu, K., McGinnis, L. (2013). Interpolation approximations for queues in series, *IIE Transactions*, Vol. 45, No. 3, 273-290, doi: 10.1080/0740817x.2012.682699.

Diagnostic of peripheral longitudinal grinding by using acoustic emission signal

Zylka, L.[a,*], Burek, J.[a], Mazur, D.[b]

[a]Department of Manufacturing Techniques and Automation, Rzeszow University of Technology, Poland
[b]Department of Electrical and Computer Engineering Fundamentals, Rzeszow University of Technology, Poland

A B S T R A C T

Grinding burn is one of the well-known problems in grinding processes. The phenomenon of burns causes permanent damage to the ground surface. Therefore, there is a need of monitoring the grinding processes in order to prevent surface damage of a workpiece. This paper presents a method of diagnosing grinding wheel wear with the use of acoustic emission signal generated during grinding. The method aims to detect the occurrence of burn in the surface layer of ground workpieces, and, thus, to replace costly and troublesome surface layer control methods performed after grinding. Experimental research of the grinding process together with the control of surface layer condition was conducted by means of the nital etching method. A band analysis of acoustic emission signal was completed and the influence of the grinding burns phenomenon on the signal amplitude in the range of low frequencies was presented. A boundary value of the *AE* describing the appearance of grinding burns was determined. Moreover, *RMS* value of acoustic emission signal was analysed, and the influence of grinding wheel wear on the signal variations was determined. A new parameter was proposed in order to determine the end of grinding wheel life-time. A boundary value of this *AE* parameter, which indicates the excessive wear of grinding wheel was determined.

1. Introduction

Grinding is classified as after-machining and, therefore, it is usually the last step of a technological process. Therefore, grinding and its result determine final parameters of the manufactured parts, which have a direct influence on their operational properties [1]. A very important aspect is ensuring the correct grinding process, namely acquiring the assumed quality parameters of the manufactured products, which include, among others, the condition of the surface layer after machining [2]. Currently, in industrial production the control of surface layer damage of the ground surfaces is done by means of many control methods, often very time-consuming or harmful for the natural environment [3, 4]. Additionally, the use of such methods as fluorescent, magnetic or nital etching ones requires high qualifications from the employees. These measurements are done after machining, thus, approving or rejecting the results of machining.

One type of surface layer damage that completely excludes a workpiece from usage are grinding burns. The term grinding burn refers usually to the altered structure of the ground surface layer of a workpiece, generated in a result of external, thermal influence of contact area of a grinding wheel with a workpiece [5-7]. Irrespectively from the kind of the resultant grinding

burn it negatively influences usage properties of the external layer, causing first and foremost lowering of the fatigue strength. The grinding burns should be detected during the grinding process so as not to accept a thermally damaged product to usage [8].

Performing control during the grinding process enables correction of its parameters and possible prevention of defects during grinding [9, 10]. Therefore, the search for such control methods is an extremely important task [11, 12]. In many papers, an attempt was made to supervise the machining process (the surface layer), that comprised in the measurement and proper analysis of acoustic emission signal (*AE*) generated during grinding [13-17]. Much experimental research was done in the area of controlling the generation of grinding burns during the grinding process, but in none of the studies comprehensive solution to the problem was presented.

An *AE* signal was subjected to various processing methods in order to isolate its parameters that would indicate its relation to the burns. In the last years, a lot of research of the acoustic emission in the grinding was made. For example Dotto, Aguiar and others developed a method for grinding burns detection with the application of acoustic emission and spindle motor power as signals [5, 6], and the efficiency of this method amounted to 82 %. In 2006 Dotto, Aguiar, Serni and Bianchi presented a method for detecting grinding burns that was based on detailed analysis of *AE* static parameters in categories of time and frequency [15], such as: effective *RMS* value, skewness, kurtosis, autocorrelation as well as *CFAR* parameters (*Constant False Alarm Rate*), *MVD* (*Mean Value Dispersion Statistic*) and *ROP* (*Ratio of Power*). Similar studies were done by Wang, Willett, Aguiar and Webster. They came to similar conclusions, but they, however, chose *CFAR* and *MVD* parameters and the ones best correlating with grinding burns. Additionally, they proposed an artificial neural network as the tool for assessing grinding burn occurrence [18]. Qiang Liu, Xun Chen and Nabil Gindy, on the other hand, did studies in which analysis was done with regard to signal distribution in the category of time [19, 20]. Jae-Seob Kwak and Ji-Bok Song performed studies of shafts grinding process with the registration of acoustic emission signal [21] and proposed recording *AE* signal parameters, such as: standard deviation, *RMS* peak value, *FFT* peak value and exceeding the threshold value. These parameters were assumed as input values to the neural network.

Analysing the existing works it should be stated that whereas a burn is distinguished by an increase in the amplitude of the *AE* signal in the range of selected frequencies [18, 22-25]. Up to now, however, the problem of on-line detection of occurring grinding burns has not been comprehensively solved, and the acquired results were often contradictory.

2. Materials and methods

2.1 Experimental details

The test stand, that was created to conduct experimental studies, was built on the basis of the Geibel&Hotz FS 640Z CNC grinder, which can be seen in Fig. 1, together with test equipment.
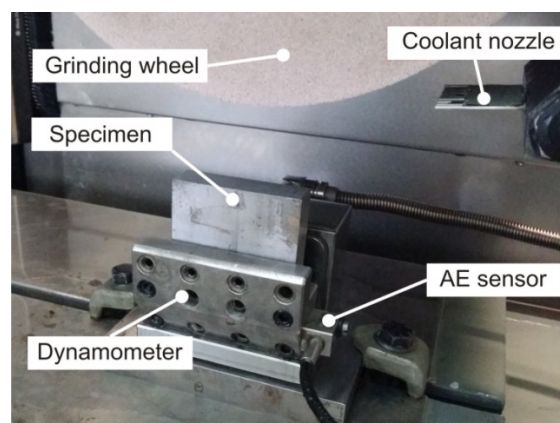
Fig. 1 The view of the test stand

The grinder was equipped with the Kistler acoustic emission measurement system, type 8152B2, with charge amplifier and *RMS* converter, type 5125B1. In the conducted research the value of acoustic emission signal amplitude was registered. It was measured by the *AE* sensor and filtered by a band-pass filter that formed a part of the 5125B1 converter of the Kistler company, and then it was recorded by means of a 12-bit A/D Acquitek measurement card, with sampling frequency of 2 MHz (the sampling frequency was selected with consideration of Nyquist criterion of sensor operation frequency and the range of filtering). All the measurement data were recorded in computer memory in the binary form, and then the data were converted in the MATLAB programme to the form that enabled their further processing and analysis.

Samples the size of $100 \times 50 \times 15$ mm, made of 20MnCr5 steel and subjected to case hardening were ground. The material hardness amounted to 57 HRC. The machining process was performed by a fused alumina grinding wheel of vitrified bond with 4 % solution of semisynthetic coolant. All the samples were initially ground with reduced grinding parameters in order to ensure constant value of ground surface layer cross-section. Then, every sample was one-pass ground and subjected to the process of nital etching in order to identify grinding burns. Detailed test conditions are presented in Table 1.

The initial test was performed and on its basis grinding wheel lifetime was approximately defined, which amounted to V_w = 960 mm³ for the considered conditions. It means that after 64 passes of the grinding wheel the first grinding burn occurred. On the basis of this study, 35 research samples were prepared, which provided the opportunity to make 70 grinding tests in the period of grinding wheel lifetime. Each sample was ground on both sides.

Table 1 Test conditions

Technological parameters		
Parameter	Symbol	Value
Grinding velocity	v_s	30 m/s
Workpiece length	l_D	100 mm
Workpiece width	b_D	15 mm
Grinding infeed	a_e	0.01 mm
Grinding feed	v_w	2000 mm/min

2.2 Research methodology

Grinding burns develop due to excessive temperature in the grinding zone. Additionally, it is assumed that the amount of heat generated during the grinding process is proportional to the volume of material removed by the grinding wheel, in a given contact surface of a grinding wheel with a workpiece in a time unit [14]. Such an assumption implies that the density distribution of grinding power on the contact surface of a grinding wheel with a workpiece is connected to the size of a ground layer. Additionally, based on a presented relationship it can be stated, that the amount of heat stream in time also depends on the tangential component of the grinding force:

$$q_t = \frac{F_t v_s R}{A_k} \tag{1}$$

where: R – coefficient determining the amount of heat getting into the workpiece, F_t – tangential component of the grinding forces, v_s – grinding speed, A_k – cross section of the ground layer.

In the presented equation the quantities that may change values are the cross sectional area of the ground surface A_k as well as the tangential component of grinding force F_t. The cross-sectional area A_k may undergo uncontrolled changes due to object deformation after heat or chemical-heat treatment. The changes (increase) of force F_t are usually caused by progressive wear of an abrasive wheel. As a result of a wheel blunting, the increase of stress between a grinding wheel and a machined object occurs, as well as the increase of the amount of friction and decrease of machining. This is subsequently followed by the increase of power and grinding forces, and in consequence, generation of an excessive amount of heat in the grinding zone (thermal damage of the ground surface) [26].
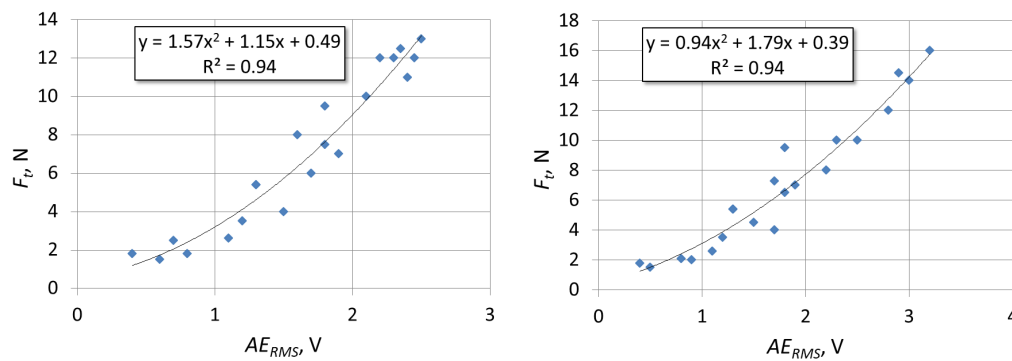
Fig. 2 Correlation of F_t and AE_{RMS} signals in the beginning (left) and in the end (right) of grinding wheel lifecycle

In order to control the grinding process by means of the AE signal and considering the afore-mentioned relationship it was necessary to determine the correlation ratio of AE_{RMS} and F_t. In order to enable replacement of force signal with the signal of acoustic emission, experimental studies were performed to study the functional relationship between these signals. The studies were conducted during the period of grinding wheel lifecycle in order to determine the type of changes of the correlation coefficient and functional relationship. The Fig. 2 presents correlation F_t and AE signals in the grinding wheel lifetime.

The presented research results of correlation between F_t and AE_{RMS} signals imply, that in the period of grinding wheel lifecycle, the signals of tangential force and the effective value of acoustic emission are characterised by high correlation ratio $R^2 > 0.9$. It was furthermore observed, that in the grinding wheel lifecycle the functional relationship between the studied signals also undergoes a change. It may be explained by the fact that the level of AE signal has an influence on many factors and phenomena taking place in the machining zone, which have an impact on the abrasive wheel wear [16]. It was also observed that the acoustic emission signal in the range of small machining forces is characterised by greater value sensitivity than tangential force. The general relationship of signals $F_t = f(AE_{RMS})^2$ confirms that. Due to the high value of the coefficient of correlation between the effective value of AE signal and the tangential force, it was assumed that the AE_{RMS} signal would be useful in controlling the grinding process.

3. Results and discussion

3.1 Analysis of the *AE* signal

On the basis of the bibliographical research one may conclude, that in grinding burns diagnostics, the spectral analysis of acoustic emission signal may also be useful. A raw AE signal, however, must be subjected to complex processing process in order to extract signal features or parameters that have a relationship with the occurrence of grinding burns. Fig. 3 presents the applied algorithm of AE processing.

The data for analysis were acquired from a measurement card in a form of files in binary format, which were transformed in order to acquire one result matrix for all measurements. The measurement was done 20 times in each abrasive wheel pass, once in 5 mm.

Then, just before the transition to the frequency range, signal windowing was performed, applying for this purpose the Hamming window. It enabled to avoid signal deformations by the subsequent Fourier transform *(FFT)*. From such prepared data, a frequency spectrum was defined by means of the *FFT* transform. Because the acquired distribution is noisy it is necessary to perform data smoothing (filtering). A filter that gives small approximation error is the Savitzky and Golay filter that functions on the basis of smoothing data with an orthogonal polynomial. Filtration parameters were selected in such a manner as to acquire maximal reduction of random errors and to minimize the introduction of systematic errors. On the basis of bibliographical data, the values of filtering parameters were chosen: degree of an approximating polynomial $j = 3$ and the width of averaging window 201.
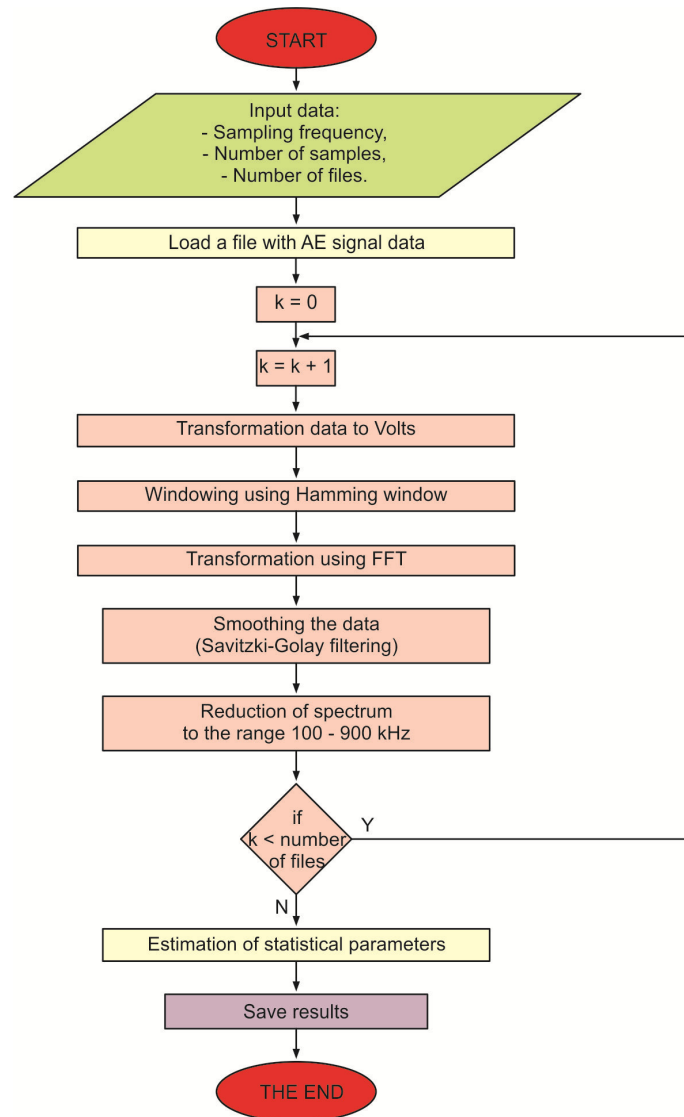
Fig. 3 Algorithm of the *AE* signal processing

The spectral distribution of the *AE* signal filtered this way was subjected to normalization, in order to bring all the frequency spectrums developing in subsequent grinding passes to one level. Each frequency spectrum was normalised by its average value. This method of amplitude spectrum normalization was applied so as to make the values of signal amplitudes independent from uncontrolled and unpredictable changes of the cross-sectional area of the machined layer, which may occur in production, and are a result of workpiece deformations generated in heat treatment. Due to the fact that the *AE* signal was hardware filtered in the range of 100-900 kHz, the spectrum from beyond that range was not taken into consideration. The last stage was the record of all the spectrums to one cumulative matrix. The data processed this way were subjected to further analysis.

During grinding of first samples, the grinding process proceeded correctly and in the test of nital etching no occurrence of grinding burns was observed. Fig. 4(a) shows an exemplary distribution of *AE* spectrums for a signal from one selected grinding pass, in which no thermal damage of the ground surface was observed. The presented spectrums imply, that in the range of all the signal samples no significant changes in the *AE* spectrum occur. Additionally, in fig 2a it can be observed that the *AE* signal has the greatest amplitude in low frequencies, up to 600 Hz, where two ranges of increased levels can be distinguished: 100-300 kHz and 350-550 kHz. In all the cases in which no grinding burns were observed, the signal spectrum was similar.

a)

b)
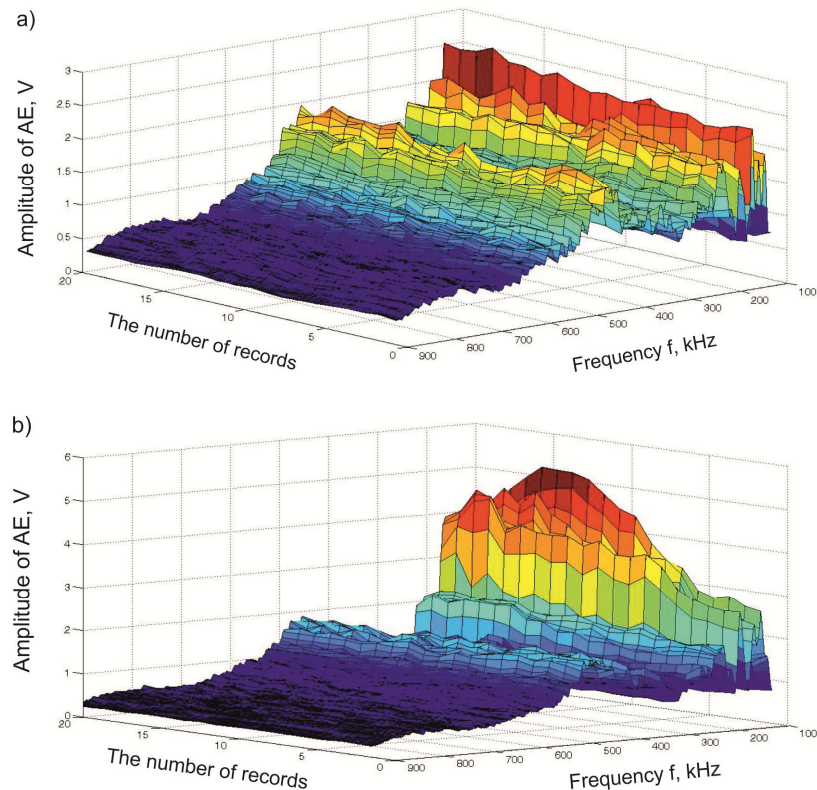
Fig. 4 The *AE* signal spectrum: a) without burns, b) with a visible grinding burn

By the end of the grinding wheel lifetime, the grinding burns started to occur. The first grinding burn was observed in the form of material tempering in the surface layer. The grinding burn was detected at the end of the sample, in the range from 70 to 100 mm of the grinding pass (sample length). Fig. 4(b) presents spectral distribution recorded while grinding the sample, together with a visible increase of signal amplitude in low frequencies, beginning from the 8-th spectrum. Comparing the spectral distribution in Fig. 4(a) with the distribution in Fig. 4(b) one can observe a significant growth of signal amplitude in the range of 100-200 kHz, while in the remaining part of the spectrum the increase is unnoticeable. This increase is observed for all spectral distributions generated from a signal recorded in the cases in which a grinding burn occurred.

The distributions of the *AE* signal recorded during further grinding of samples and the occurrence of subsequent thermal damage, such as tempering or secondary hardening, imply that irrespectively of the kind of the grinding burn, the same type of change of the *AE* signal spectrum occurs. A significant increase in the amplitude of the *AE* signal is observed in the range of frequencies of 100-200 kHz. The same changes in frequency spectrum were observed also in the case of the occurrence of grinding burns, which were acquired by increasing the infeed to the value of a_e = 0.2 mm. Thermal damage was not a result of the grinding wheel wear, but of the increased machining allowance. It is confirmed by the fact, that the increase of amplitude in low frequencies has a relationship with the occurrence of grinding burns, irrespectively of their cause.

In all cases of the grinding burn occurrence, an increase of energy in low frequencies was observed, but the maximal amplitudes were observed for different frequencies. Due to this fact, the frequency spectrum was analysed by bands, every 100 kHz, analysing the maximal amplitudes. Fig. 5 compiles maximal values of amplitudes in particular bands of the *AE* signal spectrums, recorded during grinding tests, in which a grinding burn occurred, and in these in which no grinding burn was observed.
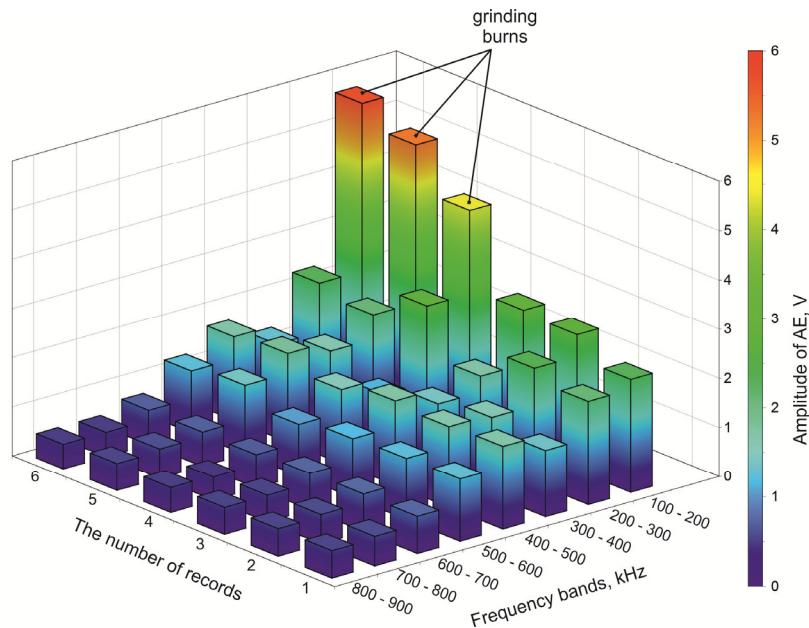
Fig. 5 Maximal values of the *AE* signal amplitudes in bands, every 100 kHz

It may be concluded from the presented research results of the changes of the *AE* signal spectrum in particular frequency bands, that the best criterion for the assessment of grinding burns occurrence is the value of the maximal amplitude of the signal in the 100-200 kHz band, because the amplitude of *AE* signal in the case of grinding burns occurrence is almost two-fold higher than in the process without visible burns.

In order to specify the criterion of grinding burns occurrence, a statistical analysis of computed maximal values of amplitudes *AE* was conducted in the band of 100-200 kHz. 30 maximal values of the *AE* amplitudes were chosen at random, for the cases in which grinding burns occurred, as well as 30 amplitudes for the cases without burns. For each group confidence interval was computed, assuming the significance level α = 0.05. Due to the fact that confidence intervals do not overlap, an unambiguous criterion for the occurrence of grinding burns was proposed in the form of the P_p parameter, that is defined by the relation:

$$P_p > \frac{\mu_{g\,min} - \mu_{d\,max}}{2}$$

(2)

where: $\mu_{g\,min}$ – the bottom value of the confidence interval of the *AE* signal amplitudes for the occurrence of grinding burns, $\mu_{d\,max}$ – the top value of the confidence interval of the *AE* signal amplitudes without the occurrence of grinding burns.

It can be assumed, that for the case under study, when P_p > 3.5 then a grinding burn occurs with the probability of p = 0.95.

3.2 The AE_{RMS} analysis

The effective value AE_{RMS} that was hardware generated by the *RMS* converter, type 5125B1, had been recorded. Then, signal values were subjected to normalization in order to acquire average values of the AE_{RMS} signal with reference to the cross-sectional area of the machined layer A_k. This procedure was performed in order to eliminate the influence of uncontrolled changes of machining allowance, resulting e.g. from object deformation during heat treatment, on the *AE* signal. For the AE'_{RMS} normalised signal that was acquired in the stated above procedure the average value for one grinding pass was computed. The recorded waveforms imply that the average value AE'_{RMS} changes in the period of grinding wheel lifetime and has a rising tendency. Fig. 6 shows changes of the processed signal value AE'_{RMS} for all grinding tests in the period of grinding wheel lifetime.
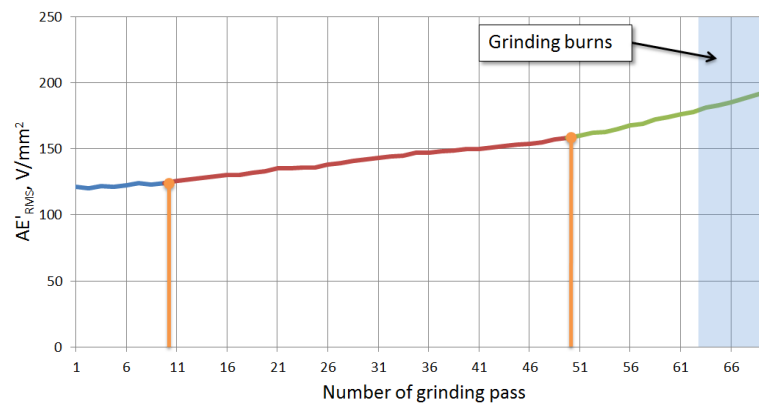
Fig. 6 Change of the AE'_{RMS} signal in the period of grinding wheel lifetime

The presented waveform implies that the AE'_{RMS} signal is correlated with the progressive wear of the grinding wheel in the period of its lifetime. Analogically to the correlation studies of F_t and AE_{RMS} signals one may distinguish three stages of grinding wheel utilization in the period of its lifetime. The first stage is connected to a grinding wheel active surface forming and appears in the beginning of the lifetime. It may be observed that the AE'_{RMS} signal presents slightly rising tendency, significantly smaller than the one occurring during further period of grinding wheel utilization and is characterised by relatively great dispersion of values. It can be explained by the phenomenon of chipping of abrasive grains that were not removed in the dressing process, and their connection with the bond. Together with crumbling of the abrasive grain and blunting their edges, the AE'_{RMS} stabilizes and starts to rise. Then, the second stage of grinding wheel utilization takes place, in which glazing of abrasive grains and partial pore-filling dominates. It can be observed, that at this stage the AE'_{RMS} signal slightly rises, which is probably caused by the progressive wear of the grinding wheel. In the end of the lifetime, the third stage occurs, in which a significant increase in the value of the AE'_{RMS} signal can be observed. The abrasive wheel is glazed and has filled pores, resulting in losing the machining properties. The third stage of the grinding wheel utilization is the end of the lifetime of a grinding wheel.

By the end of the grinding wheel lifetime, the grinding burns occur. The area of grinding burns is marked in blue in Fig. 6. However, on the basis of the signal value changes AE'_{RMS} it is not possible to specify the moment, when first burns occur. Nonetheless, it can be observed, that by the end of the grinding wheel lifetime period, the AE'_{RMS} signal shows a greater increasing tendency in comparison to the middle period of its utilization.

Taking into consideration the nature of the AE'_{RMS} signal changes, the following method for assessing the end of the grinding wheel lifetime is presented. Due to the fact, that the AE'_{RMS} signal shows a linear increase with the progressive wear of the grinding wheel, it is possible to determine the equation of lines that approximate changes of the AE'_{RMS} signal in particular stages of the grinding wheel utilization (Fig. 7).
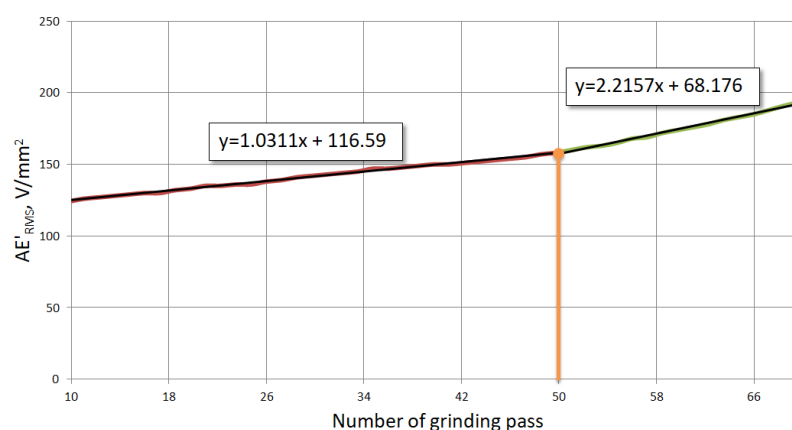
Fig. 7 Approximation of the AE'_{RMS} signal change in the period of the grinding wheel lifetime

In order to control the wear of the grinding wheel in on-line mode and stop the grinding process by the end of the grinding wheel lifetime, a proper criterion, that clearly defines the moment of the end of the grinding process and dressing the abrasive wheel should be selected to avoid the occurrence of grinding burns. Developing the criterion of the end of the grinding wheel lifetime, one should take into consideration the fact, that analysis of the AE signals acquired from the insufficient number of grinding passes will cause the criterion instability. However, too high a number of the AE signals will cause that the criterion will react too late. A compromise solution was developed by means of the P_g parameter, expressed by the following relationship:

$$P_g = \frac{\frac{1}{3}\sum_{k}^{k-2} AE'_{RMS(k)} - \frac{1}{3}\sum_{k-3}^{k-5} AE'_{RMS(k)}}{3} \qquad (3)$$

where: $AE'_{RMS(k)}$ – the average value of the AE'_{RMS} signal for the k-th grinding pass. The value of the P_g parameter for the k-th grinding pass is obtained from the previous six passes. The difference between average values of signals obtained from grinding passes, k-1 and k-2 as well as k-3, k-4, k-5 is divided by 3; thus the value of the P_g parameter can be compared to the directional coefficient of lines that approximate the change of the AE'_{RMS} signal in time.

Due to the fact, that by the end of the grinding wheel lifetime the directional coefficient increases over twofold, the following criterion that determines the end of the grinding wheel lifetime was calculated:

$$P_g > \frac{a_1 + a_2}{2} \qquad (4)$$

where: a_1 = 1.031, a_2 = 2.216 – directional coefficients of lines. Fig. 8 shows the change of the P_g parameter in the grinding wheel lifetime.

The end of the grinding wheel lifetime occurs when the difference between average values of the AE'_{RMS} signal is higher than the average value of the directional coefficients of lines, which describe the signal change in the period of the grinding wheel lifetime $P_g > 1.62$. Achieving this value implies, that the grinding process should be stopped. The criterion of the end of the grinding wheel lifetime was met by the 60th grinding pass, before the first occurrence of grinding burns.

The analysis conducted for the AE_{RMS} signal in the grinding wheel lifetime implies, that by applying proper method of processing, the signal and the method of assessing the change of its value, it is possible to define the end of the grinding wheel lifetime in the peripheral longitudinal grinding process even before the first occurrence of grinding burns.
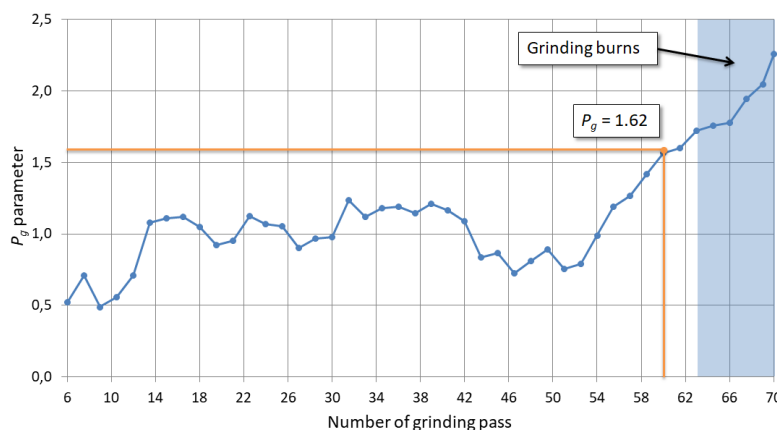
Fig. 8 Change of the P_g parameter in the grinding wheel lifetime

3.3 Diagnostic methodology

In practice, grinding burns usually occur due to two factors:

- Loss of grinding wheel's machinability and reaching the end of its lifetime,
- Uncontrolled local variations of machining layer, generally due to workpiece deformation during heat or chemical-heat treatment.

Due to the existence of two separate sources of grinding burns, developing of two diagnostic methods, that are able to react to grinding burns has been necessary. Thus, a separate spectral analysis of raw AE signal as well as its RMS value was performed. The analysis allowed to formulate two criteria of grinding process diagnostics (Fig. 9):

- Criterion I, based on measurement and analysis of AE_{RMS}, if parameter $P_g > 1.62$, it means, that a grinding wheel reached the end of its lifetime and grinding process should be stopped in order to avoid grinding burns,
- Criterion II, based on a spectral analysis of AE signal, if parameter $P_p > 3.5$, it means, that a grinding burn on a workpiece surface occurred, however, the wear of a grinding wheel had not been the source of a burn.

In Fig. 9 an idea of complex diagnostics system of grinding process based on the measurement and analysis of AE signal is presented. It depends on continuous control of the value of two parameters: P_p and P_g. If parameter P_g exceeds established threshold value, it means, that the grinding process should be stopped and the grinding wheel dressed. Continuing the grinding process may lead to grinding burns. Whereas, when parameter P_p exceeds established value, it means that a thermal damage to the workpiece occurred. Application of presented in Fig. 9 diagnostics system allows to have constant control of grinding process as well as prevents grinding burns occurring due to grinding wheel wear.
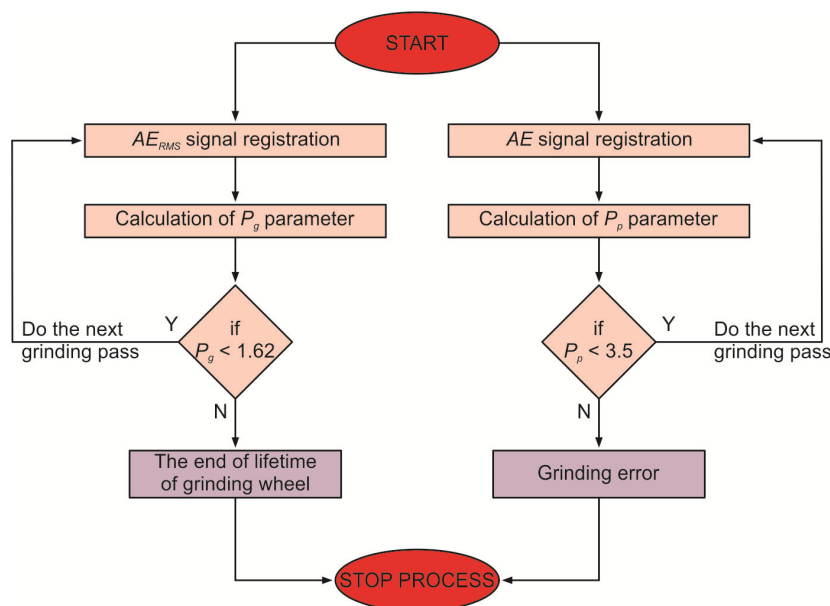
Fig. 9 General operation algorithm of the grinding diagnostics system

4. Conclusion

The above paper discusses a new method for the surface layer control during the peripheral longitudinal grinding, using acoustic emission as the diagnostic signal. The method solves comprehensively the problem of detecting and preventing the occurrence of grinding burns. The two-passes analysis of the acoustic emission signal enables to stop the grinding process before first grinding burns occur, as well as to determine a probable cause of the burns.

The most important scientific accomplishment of the paper is presenting a new, complex method of grinding burns detection and its prevention. In the diagnostic method, two parameters were used: P_p and P_g, for which threshold values were determined, allowing to prevent thermal damage to the ground surface.

The developed method of the grinding process diagnostics is based on measurement, processing and analysis of the *AE* signal. The root mean square of *AE* in a grinding wheel lifetime was analysed. The criterion, that defines the moment of the end of the grinding wheel lifetime and allows to avoid the occurrence of grinding burns was developed. The new parameter P_g had been proposed. When its value exceeds 1.62, the grinding process should be stopped.

Due to the fact, that it is not possible to detect the occurrence of grinding burns by means the AE_{RMS} signal, the *AE* signal was also analysed in terms of frequency, searching for the connection between grinding burns occurrence and the *AE* signal spectral parameters. A suitable band analysis of the acoustic emission signal enables to detect grinding burns irrespectively of their cause. The new method of calculating the parameter P_p was proposed, which may be applied in order to detect thermal damages of grinding surfaces. When $P_p > 3.5$, it indicates that grinding burns will occur during grinding.

The method does not require almost any structural changes to the grinder and the *AE* signal sensors are not very susceptible to external factors. The application of this method requires developing a proper procedure for processing and analysing the *AE* signal in the range of frequency. The analysis is necessary to define signal parameters and to specify the conditions determining the occurrence of thermal damage to the workpiece during the machining process. It should be noted, that the boundary values of the P_p and P_g parameters were determined for the adopted study conditions and for other types of ground material and different abrasive wheel they may assume different values.

References

[1] Nadolny, K. (2013). A review on single-pass grinding processes, *Journal of Central South University,* Vol. 20, No. 6, 1502-1509, doi: 10.1007/s11771-013-1641-5.

[2] Demir, H., Gullu, A., Ciftci, I., Seker, U. (2010). An investigation into the influences of grain size and grinding parameters on surface roughness and grinding forces when grinding, *Strojniški vestnik – Journal of Mechanical Engineering*, Vol. 56, No. 7-8, 447-454.

[3] Klocke, F., König, W. (2005). *Fertigungsverfahren – Schleifen, Honen, Läppen,* 4th edition, Springer, Berlin, Heidelberg, Germany, doi: 10.1007/3-540-27699-8.

[4] Malkin, S., Guo, C. (2008). *Grinding technology: Theory and applications of machining with abrasives,* Industrial Press, New York, USA.

[5] De Aguiar, P.R., Bianchi, E.C., Serni, P.J.A., Lanconi, P.N. (2002). Control of thermal damage in grinding by digital signal processing of raw acoustic emission, In: *Seventh International Conference on Control, Automation, Robotics and Vision (ICARCV'02)*, Singapore, 1392-1397, doi: 10.1109/ICARCV.2002.1234976.

[6] Dotto, F.R.L., De Aguiar, P.R., Bianchi, E.C., Flauzino, R.A., Castelhano, G.O., Pansanato, L. (2003). Automatic detection of thermal damage in grinding process by artificial neural network, *Rem: Revista Escola de Minas*, Vol. 56, No. 4, 295-300, doi: 10.1590/S0370-44672003000400013.

[7] Sinha, M.K., Setti, D., Ghosh, S., Rao, P.V. (2016). An investigation on surface burn during grinding of Inconel 718, *Journal of Manufacturing Processes*, Vol. 21, 124-133, doi: 10.1016/j.jmapro.2015.12.004.

[8] Azarhoushang, B., Daneshi, A., Lee, D.H. (2017). Evaluation of thermal damages and residual stresses in dry grinding by structured wheels, *Journal of Cleaner Production*, Vol. 142, Part 4, 1922-1930, doi: 10.1016/j.jclepro.2016.11.091.

[9] Sutowski, P., Nadolny, K., Kaplonek, W. (2012). Monitoring of cylindrical grinding processes by use of a non-contact AE system, *International Journal of Precision Engineering and Manufacturing*, Vol. 13, No. 10, 1737-1743, doi: 10.1007/s12541-012-0228-7.

[10] Kruszyński, B.W., Lajmert, P. (2005). An intelligent supervision system for cylindrical traverse grinding, *CIRP Annals – Manufacturing Technology,* Vol. 54, No. 1, 305-308, doi: 10.1016/S0007-8506(07)60109-7.

[11] Mičietová, A., Neslušan, M., Čep, R., Ochodek, V., Mičieta, B., Pagáč, M. (2017). Detection of grinding burn through the high and low frequency Barkhausen noise, *Tehnički vjesnik – Technical Gazette*, Vol. 24, No. 1, 71-77, doi:10.17559/TV-20140203083223.

[12] Neslušan, M., Čížek, J., Kolařík, K., Minárik, P., Čilliková, M., Melikhova, O. (2017). Monitoring of grinding burn via Barkhausen noise emission in case-hardened steel in large-bearing production, *Journal of Materials Processing Technology*, Vol. 240, 104-117, doi: 10.1016/j.jmatprotec.2016.09.015.

[13] Kruszyński, B.W., Lajmert, P. (2006). An intelligent system for online optimization of the cylindrical traverse grinding operation, *Proceedings of the Institution of Mechanical Engineers, Part B: Journal of Engineering Manufacture*, Vol. 220, No. 3, 355-363, doi: 10.1243/095440506X77607.

[14] Sutowski, P., Plichta, J. (2006). An investigation of the grinding wheel wear with the use of root-mean-square value of acoustic emission, *Archives of Civil and Mechanical Engineering*, Vol. 6, No. 1, 87-98, doi: 10.1016/S1644-9665(12)60078-8.

[15] Aguiar, P.R., Serni, P.J.A., Dotto, F.R.L., Bianchi, E.C. (2006). In-process grinding monitoring though acoustic emission, *Journal of the Brazilian Society of Mechanical Science and Engineering*, Vol. 28, No. 1, 118-124, doi: 10.1590/S1678-58782006000100014.

[16] Lv, C., Li, H. (2010). Acoustic emission signal processing of grinding monitor, In: *2010 3rd International Congress on Image and Signal Processing*, Yantai, China, 3836-3838, doi: 10.1109/CISP.2010.5646807.

[17] Kwak, J.-S., Ha, M.-K. (2004). Neural network approach for diagnosis of grinding operation by acoustic emission and power signals, *Journal of Materials Processing Technology*, Vol. 147, No. 1, 65-71, doi: 10.1016/j.jmatprotec.2003.11.016.

[18] Wang, Z., Willett, P., DeAguiar, P.R., Webster, J. (2001). Neural network detection of grinding burn from acoustic emission, *International Journal of Machine Tools and Manufacture*, Vol. 41, No. 2, 283-309, doi: 10.1016/S0890-6955(00)00057-2.

[19] Liu, Q., Chen, X., Gindy, N. (2005). Fuzzy pattern recognition of AE signals for grinding burn, *International Journal of Machine Tools and Manufacture*, Vol. 45, No. 7-8, 811-818, doi: 10.1016/j.ijmachtools.2004.11.002.

[20] Liu, Q., Chen, X., Gindy, N. (2006). Investigation of acoustic emission signals under a simulative environment of grinding burn, *International Journal of Machine Tools and Manufacture,* Vol. 46, No. 3-4, 284-292, doi: 10.1016/j.ijmachtools.2005.05.017.

[21] Kwak, J.-S., Song, J.-B. (2001). Trouble diagnosis of the grinding process by using acoustic emission signals, *International Journal of Machine Tools and Manufacture*, Vol. 41, No. 6, 899-913, doi: 10.1016/S0890-6955(00)00082-1.

[22] Saravanapriyan, S.N.A., Vijayaraghavan, L., Krishnamurthy, R. (2001). On-line detection of grinding burn by integrated sensing, In: *SIcon'01 Sensors for Industry Conference*, *Proceedings of the First ISA/IEEE Conference*, Illinois, USA, 89-94, doi: 10.1109/SFICON.2001.968505.

[23] Griffin, J.M., Chen, X. (2009). Multiple classification of the acoustic emission signals extracted during burn and chatter anomalies using genetic programming, *International Journal of Advanced Manufacturing Technology,* Vol. 45, No. 11-12, 1152-1168, doi: 10.1007/s00170-009-2026-7.

[24] Kwak, J-S., Ha, M.-K. (2004). Neural network approach for diagnosis of grinding operation by acoustic emission and power signals, *Journal of Materials Processing Technology*, Vol. 147, No. 1, 65-71, doi: 10.1016/j.jmatprotec.2003.11.016.

[25] Yang, Z., Yu, Z., Xie, C., Huang, Y. (2014). Application of Hilbert-Huang transform to acoustic emission signal for burn feature extraction in surface grinding process, *Measurement*, Vol. 47, 14-21, doi: 10.1016/j.measurement.2013.08.036.

[26] Liao, T.W., Tang, F., Qu, J., Blau, P.J. (2008). Grinding wheel condition monitoring with boosted minimum distance classifiers, *Mechanical Systems and Signal Processing,* Vol. 22, No. 1, 217-232, doi: 10.1016/j.ymssp.2007.06.005.

Determination of accuracy contour and optimization of workpiece positioning for robot milling

Gotlih, J.[a,*], Brezocnik, M.[a], Balic, J.[a], Karner, T.[a], Razborsek, B.[a], Gotlih, K.[a]

[a]Faculty of mechanical engineering, University of Maribor, Maribor, Slovenia

ABSTRACT

Workpiece positioning into a machine's workspace has become a simple task. Advanced CNC machines are equipped with standardized clamping systems, allowed workpiece dimensions are listed in the machine's documentation and tolerance levels of the end produced parts are known. This gives users plenty of information and good confidence that they are choosing the best machine for a specific task. For more universal machines like industrial robots this is not the case. Due to their flexibility industrial robots can be an alternative to specialized CNC machines, but when a specific task should be executed, important information is missing. For a standard industrial robot the mechanisms layout, its dimensions and its reachable workspace is known, but accuracy levels over the robot's workspace are not. If a workpiece should be milled within certain accuracy limits the robot's documentation offers no information on how close it can be located to the borders of the robot's workspace. This article deals with the mentioned problem with a novel methodology. Based on experimental data we found that a standard 6 DOF industrial robot's reachable workspace can be divided into two regions, one with suitable milling accuracy and another with rapidly decreasing milling accuracy. To isolate the suitable accuracy region a regional non-dominated sorting algorithm was developed and an accuracy contour separating the regions was extracted. In the second part of the article a genetic search algorithm based on regional non-dominated sorting was applied to find the biggest arbitrary shaped workpiece's size, position and orientation in the suitable milling accuracy region of the robot's workspace.

1. Introduction

Using industrial robots for less common tasks is becoming increasingly interesting among researchers, as due to their flexibility, machining could become a more regular application for industrial robots. To reach the level where robot machining becomes an acceptable alternative to CNC machining, robot machining accuracy needs to be analysed, understood and improved. The main reasons for inaccuracies in robot machining were classified and divided into three groups by Kubela et al. [1] as environmental errors, robot dependent errors and process dependent errors. Robot dependent errors were further divided into geometrical, non-geometrical and system errors. Their study on wood milling showed that the most significant errors are the result of robots reverse motion, causing end product shape inaccuracies in range from 0.3–0.5 mm. Geometric error compensation has been studied by Wu et al. [2], who added a fifth step to the usual robot calibration approach, used to achieve the desired robot positioning accuracy. The proposed steps include modelling, design of experiments, measurement, identification, and imple-

mentation. Their goal was to improve the standard approach, so that a minimum number of experiments is required for robot's calibration. With the proposed calibration technique an improved positioning accuracy was achieved in absolute range of up to 0.17 mm. Greater positioning accuracy was measured by Józwik et al. [3], who used a high speed camera to investigate the effect of direction of approach on positioning accuracy. In their experiment the authors found that positioning inaccuracy is raising with robot's continuous operation. Accuracy improvement in continuous motion of a welding robot was achieved by Marônek et al. [4] by insertion of additional points into the robot's trajectory. Due to technological reasons constant speed needs to be maintained during welding. With help of experiments the authors found better trajectory matching at higher movement speeds. Jezeršek et al. [5] developed a system for automatic trajectory generation for robotic welding based on robotic laser scanning. Accuracy levels comparable to traditional visually based trajectory generation methods were achieved, but teaching times were shortened more than 30-times. A system for optical accuracy evaluation of robots in static and dynamic operation mode was set up by Papa and Torkar [6]. In their experiment the authors found the same subpixel accuracy in both modes.

To predict a mechanism's accuracy, different dexterity related performance measures have been introduced in the past. Manipulating ability of end-effector of robotic mechanisms was studied by Yoshikawa [7] and a manipulability index was introduced. Manipulability, defined as determinant of the Jacobian matrix of a robotic mechanism is proposed as a robot's performance indicator. Other dexterity measures based on determinant of the Jacobian matrix like condition number and monotone manipulability were proposed, as summarized by Gotlih et al. [8]. In their study authors demonstrated application of velocity anisotropy on a real industrial robot and introduced a four dimensional robotic workspace for visualisation of robot's manipulability. Problems of linking performance measures to a robot's accuracy were exposed by Merlet [9]. In his experiment only the robot's manipulability index has shown the same trend as its measured accuracy, but no connection between manipulability and accuracy was established. To overcome unit based inconsistency, transition from kinematic to dynamic performance measures was proposed and a power manipulability index was introduced by Mansouri and Ouali [10]. The power manipulability index is a combination of kinematic and static parameters and includes translational and rotational components. The authors showed that power manipulability is a homogenous tensor insensitive to physical units change.

Optimization of robotic systems with algorithms based on performance measures was studied by many authors. Dynamic performance measures were found to be more significant than kinematic by Vosniakos and Matsas [11]. The authors used genetic algorithm for optimal positioning of a predefined milling path into a robot's workspace. They found out that with application of genetic algorithm an improvement in robot's manipulability and a reduction in joint torques can be achieved. Genetic algorithm was also used for robot's path planning to avoid collision of robot's structure with obstacles in its workspace [12]. Another approach of robot path planning considering a mechanical power index and an obstacle avoidance index in a multi-criteria optimization problem was suggested by R. dos Santos et al. [13]. To overcome local minima authors have used a two-phase optimization process and the sequential linear programming method. Results showed good agreement with Pareto optimality as better obstacle avoidance caused an increase in power consumption. A multi-criteria path planning algorithm for robotic laser cutting was developed by Dolgui and Pashkevich [14] and implemented in an industrial software package, yielding significant reduction in process planning and changeover time for end users. A three step optimization procedure for workpiece positioning into a robot's workspace was proposed by Lin et al. [15]. The authors developed an approach based on manipulability and stiffness indexes and introduced a deformation index to refine the final workspace maps, where contour lines dividing the robot's workspace into sub regions indicate optimal workpiece positioning locations.

To position a workpiece into a robot's refined workspace precisely, it is necessary to extract individual contour lines. Techniques for contour extraction are mainly associated with computer vision and shape detection and are used in various engineering fields. An approach for workpiece localization and its shape deviations from nominal geometry in robotic deburring was

proposed by Kuss et al. [16]. A point cloud of a CAD model was compared with a point cloud representing a real workpiece in a robot's workspace. The best matching point clouds were used to generate workpiece contours and trajectories for the deburring process. In their experiment the authors used a milling robot equipped with a stereo camera for shape sensing and achieved a significant improvement in deburring quality. A robotic cell for optical workpiece detection was developed by Klancnik et al. [17]. The authors used neural networks for system calibration where the relation between captured pictures and the robot's coordinate system was learnt. Numerous tests proved the system adequate for use in practice. Karimi and Nategh [18] introduced contour maps to graphically estimate kinematic errors in hexapod machines for an optimal workpiece setup. The authors developed an analytical model for kinematic error estimation with an average 11 % discrepancy from experimental results. Contour line generation from short line segments with a genetic algorithm based method was proposed by Wei et al. [19]. An automatic octree based algorithm for façade and opening contour detection from point clouds obtained with terrestrial laser scanning was developed by Truong-Hong and Laefer [20] where a knowledge database was used for contour validation. Relative errors of maximum 3 % were found between measured drawings and the reconstructed models. A method for building boundary identification based the Delaunay triangulation was developed by Awrangjeb and Guojun [21]. By defining a maximum point to point distance the algorithm proved capable of detecting concave edges and holes of any point cloud shape.

2. Material and methods

An experiment to evaluate the effect of different performance measures on robot milling accuracy was set up by Veber [22] in his master's thesis. A standard 6 DOF industrial milling robot equipped with a milling tool and a clamping table was selected for the experiment and the system was calibrated. To reduce influence of human errors each experiment was repeated three times. Workpieces of polymer composite material with low hardness were selected to reduce inaccuracies caused by cutting forces. According to the robot's manufacturer, workpiece hardness should not exceed 35 N/mm² and deviations in range of ± 0.12 mm should be expected due to the robot's rigidity. With all inaccuracy causes controlled, measured inaccuracies were expected to be influenced mainly by the robot's dexterity.

In the experiment five milling locations were selected and at each location three workpieces were milled in three sequential runs. The positioning of the workpieces is presented in Table 1, where x, y, z represent distances in X, Y and Z direction from the robot's base coordinate system and d is the resulting absolute distance.

Edge and circular pocket milling operations were performed on each workpiece. End shapes were then measured with a coordinate measuring machine. Measurement results compared to nominal values are presented in Table 2.

Table 1 Initial positions of the measured workpieces

Pos.	x (mm)	y (mm)	z (mm)	d (mm)
1	1940	1053	610	2207
2	2060	-498	610	2060
3	1283	0	610	1376
4	1515	821	610	1723
5	1811	-709	610	1945

Table 2 Robot milling measurement results

Pos.	Experiment 1 Nominal value (mm)		Experiment 2 Nominal value (mm)		Experiment 3 Nominal value (mm)	
	20	80	20	80	20	80
1	20.134	80.122	20.136	80.122	20.134	80.124
2	20.166	80.143	20.167	80.142	20.167	80.143
3	20.477	80.522	20.476	80.523	20.476	80.523
4	20.165	80.134	20.166	80.134	20.165	80.136
5	20.155	80.134	20.155	80.134	20.155	80.134

3. Results and discussion

For result analysis absolute errors for measured diameters and edge lengths were calculated and a normalized distance parameter (NDS) expressed in percent was introduced. NDS was calculated as the distance of a workpiece from the robot's base coordinate system divided by the maximum distance the robot could reach with a vertical tool orientation.

Fig. 1 reveals highest absolute errors at milling position 3, located close to 50 % NDS. For workpieces located at 63 % NDS and above, absolute errors are almost constant and do not exceed 0.2 mm, which is close to the inaccuracy caused by the robot's rigidity. In range between 51 % and 63 % NDS, milling inaccuracy quickly raises. Due to high measurement consistency indicated by low result variation, human and cutting force based inaccuracies can be excluded and the quick decrease of accuracy at milling position 3 can be attributed to the robot's dexterity.

Fig. 2 shows result point deviations from a normal distribution. Measurement results are represented by sample quantiles, while theoretical quantiles represent a standard normal distribution. For both measured quantities the highest deviations are found at highest inaccuracies. This indicates a nonlinear dependency between workpiece distance from the robot's base and the robot's milling accuracy, based on which a sharp boundary separating the robot's workspace into two milling accuracy regions is expected.
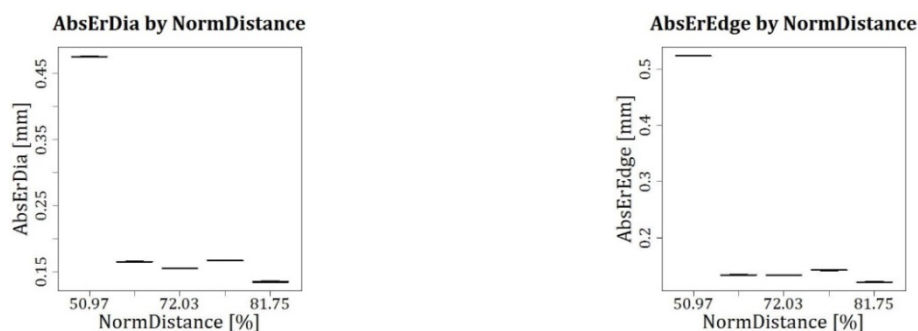
Fig. 1 Absolute errors for diameter (left) and edge length (right) versus normalized distance

Fig. 2 Absolute error deviations for diameter (left) and for edge length (right)

3.1 Accuracy contour for robot milling

In order to extract the suitable milling accuracy region from the robot's workspace, a connection between accuracy and dexterity was established. Because of the unit based inconsistency of manipulability [10] we did not intend to find a model connecting accuracy with dexterity over its complete range, instead we only equated accuracy and manipulability at a constant value. The dexterity measure defined by Yoshikawa [7] was found to be the most reliable and was therefore used in our analysis. By assuming that same manipulability causes same milling accuracy, we were able to extrapolate a milling accuracy boundary value to the robot's three dimensional workspace to create a spatial accuracy contour, separating the region of suitable milling accuracy from the region of quickly reducing milling accuracy.

As an example a standard 6 DOF robot, described by Denavit-Hartenberg (DH) notation, given in Table 3 was chosen.

With inverse kinematics the robot's joint angles and dexterity values at all five milling positions were found. Tool orientation during milling was parallel to the robot's main Z axis, therefore $\Theta4$ and $\Theta6$ joints were considered constant. Additionally, all workpieces were projected to the Y = 0 plane, so that rotation about robot's first axis $\Theta1$ was also considered constant. As a result only $\Theta2$, $\Theta3$ and $\Theta5$ contributed to dexterity calculation.

Normalized dexterity (ND) was introduced as actual dexterity divided by the highest dexterity of the mechanism. ND is expressed in percent (Eq. 1).

$$ND = \frac{detJ}{\max(detJ)} \cdot 100 \tag{1}$$

Table 4 shows relative joint angles and ND values of the robot at all five milling positions. Robot's initial angle positions were $\Theta1 = \Theta2 = \Theta3 = \Theta6 = 0°$, $\Theta4 = 180°$ and $\Theta5 = 90°$.

Results presented in Tables 2 and 4 reveal that the rapid decrease in accuracy happens between 22 % and 45 % normalized dexterity. Based on our conclusion that one region with suitable milling accuracy exists and that the transition between the suitable accuracy region and the rapidly decreasing accuracy region is sharp, 40 % normalized dexterity was chosen as the boundary value for accuracy contour generation.

Fig. 3 shows the clamping table covered by normalized dexterity contours for vertical tool orientation. Workpieces at actual milling positions are represented by black squares scaled by 5.

Table 3 DH notation of the discussed robot

Link	d (mm)	a (mm)	α (°)	θ (°)	Range (°)
0	0	0	0	0	0
1	750	350	90	0	+/-185
2	0	1250	0	0	0/-146
3	0	-55	90	90	+155/-119
4	1100	0	90	180	+/-350
5	0	0	90	90	+/-125
6	0	0	0	0	+/-350
Tool tip	A	B	C	D	E

Table 4 Joint angles and ND at milling positions

Pos.	$\Theta1$ (°)	$\Theta2$ (°)	$\Theta3$ (°)	$\Theta4$ (°)	$\Theta5$ (°)	$\Theta6$ (°)	ND (%)
1	0	37.8	17.2	180	125	0	81.8
2	0	43.0	6.0	180	130	0	72.6
3	0	62.0	-35.7	180	153.8	0	22.2
4	0	53.0	-15.8	180	142.8	0	46.1
5	0	46.6	-1.8	180	135.2	0	63.9
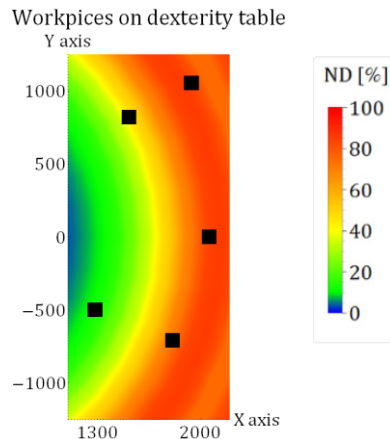
Fig. 3 Top view of workpieces on clamping table with dexterity contours

To generate the accuracy contour, regions in the robot's workspace with at last 40 % normalized dexterity were isolated. As revolution about the robot's first axis does not affect its dexterity it was possible to consider only the workspace cross section and revolve results to three-dimensional space afterwards. According to this simplification a two- dimensional restricted workspace was created, only containing points with normalized dexterity values 40 % and above. For a more general solution tool orientation was considered free.

To isolate contour points of the restricted workspace a regional non-dominated sorting algorithm was developed. The algorithm searches for non-dominated minimum points defined by x and z in the restricted workspace point cloud, which represent to the robot's tool tip positions in X and Z direction relative to the robot's base coordinate system. Extracted two- dimensional contour points form the outer boundary of the subspace in the robot's complete workspace, inside which milling with suitable accuracy is possible. Pseudocode of the algorithm is presented in Fig. 4. *IC* stands for iteration count, *RC* stands for region, *LE* stands for number of local extremes and *Q1-Q4* are bool variables representing quadrant 1-4. MA stands for mirror axis, NnD stands for non-dominated and AC stands for accuracy contour. Initialization values are presented in Table 5.

To find the first non-dominated front in the restricted workspace, which corresponds to the first iteration in Fig. 4, regular non-dominated search algorithm was applied. In the following iterations the restricted workspace was mirrored over each axis to expose outer regions and isolate front points contributing to the final accuracy contour. Due to the shape of the restricted workspace shown in Fig. 5, local extreme points in X and Z direction had to be found in order to generate an equally dense contour. In our case one local extreme point was defined in initialization dividing the concave region into two refined regions, but an arbitrary refinement level would be possible if necessary. The refined regions were restricted by squares ranging from the existing regions contour's end point to the enclosed local extreme point. The relevant mirror axis for each refined region was determined based on the adjacent region's mirror axis and the transformation Table 6, where *Q1-Q4* represent square quadrants in XZ plane (Fig. 5) sequenced clockwise starting at the top right corner.

1	Initialization.
2	If *IC* < 5:
3	If *IC* = 1 ∨ *Q3* = 1: MA = None, NnD sort, *Q3* = 0, *IC* + 1, go to step 2.
4	If *IC* = 2 ∨ *Q4* = 1: MA = X, NnD sort, *Q4* = 0, *IC* + 1, go to step 2.
5	If *IC* = 3 ∨ *Q2* = 1: MA = Z, NnD sort, *Q2* = 0, *IC* + 1, go to step 2.
6	If *IC* = 4 ∨ *Q1* = 1: MA = X ∧ MA = Z, NnD sort, *Q1* = 0, *IC* + 1, go to step 2.
7	Else:
8	If *RC* = 0:
9	If *LE* > 0: *RC* = 1, go to step 7.
10	If *LE* = 0: Go to step 13.
11	If 0 < *RC* < *LE* + 1: *RC* boundaries. Find mirror axes. *Q (1-4)* = 1, *RC* = *RC* + 1, go to step 2.
12	If *RC* = *LE* + 1: Go to step 13.
13	Join NnD regions into suitable accuracy contour.

Fig. 4 Pseudocode for accuracy contour generation

Table 5 Regional non-dominated sorting algorithm variables and initial values

Regional NnD – variables	Unit	Value
Iteration count (*IC*)	-	1
Region count (*RC*)	-	0
Local extremes (*LE*)	-	1
Quadrants (*Q1-Q4*)	-	0

Table 6 Transformation table for refined region definition

Quadrant	Change direction	Quadrant
Q1	⇔	Q2
Q2	⇔	Q1
Q3	⇔	Q4
Q4	⇔	Q3
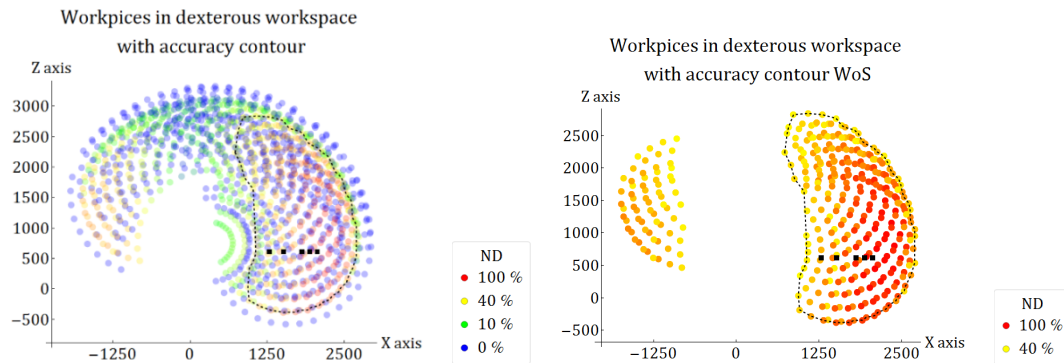
Fig. 5 Cross section of robot's complete (left) and restricted dexterous workspace (right) with projected workpiece positions and accuracy contour

After a refined region was mirrored over the relevant axis, the non-dominated sorting algorithm was applied to extract the regions contour points. Finally, all partial results were joined into the accuracy contour.

On the left side of Fig. 5 complete dexterous workspace cross section with accuracy contour and five workpiece positions is shown. Blue dots represent singularity points, which are regions with lowest dexterity, green and yellow dots represent transition regions with medium dexterity and red dots represent regions with highest dexterity. Accuracy contour is marked with a dashed black line. Workpieces are presented in the projected plane and scaled by 5. The plot on the right side shows the suitable milling accuracy region with normalized dexterity values 40 % and above, the accuracy contour and all five projected workpiece positions. For accuracy contour generation only the bigger point cloud was considered as the revolution about the robot's main Z axis covers complete 360 °.

Now, the tool orientation effect becomes visible. Dexterity contours corresponding to a fixed tool axis shown in Fig. 3 are different from dexterity contours shown in Fig. 5, where tool orientation was considered free. Free tool orientation enables another degree of freedom, which increases the size of the suitable region. By allowing free tool orientation, also the workpiece closest to robot's base is located inside the suitable milling accuracy region, meaning that all workpieces could be milled with suitable accuracy, regarding their size and assuming correct tool and technology selection.

3.2 Optimization of workpiece positioning

With the separation of the robot's workspace into two milling quality regions, it is interesting to find the size, position and orientation of the biggest workpiece that can be milled with suitable accuracy. As the three dimensional workspace of the robot in consideration can be created by revolving its cross section about its main Z axis, it is obvious that the biggest workpiece with eight corner points is a parallelepiped.

Genetic algorithm was applied as the search tool for the biggest parallelepiped. The algorithm was set up as an extension of the regional non-dominated sorting algorithm used to create the accuracy contour. Pseudocode of the algorithm is presented in Fig. 6 where *IC* stands for iteration count, *MI* for maximum number of iterations, *PC* for population count, *PS* for population size and *CC* for convergence criteria. GA stands for genetic algorithm, RP for reference point and NnD stands for non-dominated.

The algorithm requires correct initialization as shown in Fig. 6. Maximum allowed iteration number, population size, reproduction, crossover and mutation frequency, mutation strength and convergence criteria were declared initialization variables. Maximum number of iterations defines the maximum allowed loop repetitions, population size defines the quantity of parallelepipeds generated in one iteration, reproduction frequency defines the quantity of individuals to be transferred to the next generation, crossover frequency defines the fraction of individuals to participate in gene exchange and mutation frequency defines the number of individuals to mutate with a magnitude of geometrical change defined by mutation strength. Convergence criteria is an early exit criteria. Values used for initialization are presented in Table 7.

1	Initialization.
2	If $IC = 1$: Boundaries for first generation, go to step 5.
3	If $1 < IC < MI$: Boundaries for non-first generation, GA mechanics, go to step 5.
4	If $IC = MI$: Go to step 12.
5	Generate parallelepipeds, add RP to restricted workspace list, NnD sort.
6	If RP is on front: Discard parallelepiped, go to step 2.
7	Else: Add parallelepiped to population.
8	If $PC < PS$: Go to step 2.
9	Else:
10	$IC = IC + 1$,
10	If CC = True: Go to step 12.
11	Else: Go to step 2.
12	Find global best.

Fig. 6 Pseudocode of the genetic search algorithm

Table 7 Search algorithm variables and their initial values

GA - variables	Unit	Value
Max. iteration number (MI)	-	100
Population size (PS)	-	20
Reproduction frequency (RF)	%	10
Crossover frequency (CF)	%	60
Mutation frequency (MF)	%	30
Mutation strength (MS)	-	0.1
Convergence criteria (CC)	-	10

To generate a parallelepiped a random corner point and two edge lengths describing a parallelogram in the Y = 0 plane were created. The third dimension was added by extrusion of the parallelogram in Y direction. For the first iteration different boundary conditions to restrict parallelepiped's sizes closely to the allowed region were used as for the following iterations. In the first iteration parallelepiped's sizes, positions and orientations were generated by Eqs. 2-8:

$$x = randomReal\{x_{min} + 500, x_{max} - 500\} \quad (2)$$

$$y = 0 \quad (3)$$

$$z = randomReal\{z_{min} + 500, z_{max} - 500\} \quad (4)$$

$$l_a = randomReal\{0, 1\} \cdot \frac{x_{max} - x_{min}}{2} \quad (5)$$

$$l_b = randomReal\{0, 1\} \cdot \frac{z_{max} - z_{min}}{2} \quad (6)$$

$$l_c = randomReal\{0, 1\} \cdot \frac{x_{max} - x_{min}}{2} \quad (7)$$

$$\varphi = randomReal\{0, \pi\} \quad (8)$$

where x, y, z are the biggest parallelogram's corner point coordinates, φ is its orientation about the Y axis and l_a, l_b and l_c are the associated parallelepiped's edge lengths. $x_{min}, x_{max}, z_{min}$ and z_{max} are the extreme points of the accuracy contour in X, Y and Z direction. As the workspace is rotationally symmetrical x_{min} and x_{max} are used as extreme points in Y direction and only one orientation has to be considered variable.

After a parallelepiped was created and positioned in the robot's workspace, its reference points were added to the restricted workspace point cloud, whereby corner points of extruded corners were projected to the Y = 0 plane. Because of the concave shape of the accuracy contour also edge points had to be checked by the contour criteria, therefore three equidistant reference points were created on each edge and added to the restricted workspace point cloud. Then regional non-dominated sorting was applied on the extended point cloud. If a reference point was

found non-dominated it was located outside of the allowed space and a workpiece containing such a point would fall outside of the robot's suitable milling region. To avoid bad individuals proceeding to the next generation, every parallelepiped that failed the above criteria was discarded and a new one was created instead. After as many acceptable parallelepipeds as defined by population size were created, they were sorted by relevance, where bigger volume was considered better.

For the following iterations GA mechanisms selection, reproduction, crossover and mutation were applied to create new generations of parallelepipeds. For reproduction, only the best individuals from the last generation were used. If multiple individuals were tied, a random selection among them was made. In crossover each individual from the last generation could participate multiple times. To avoid reproduction an individual could never be crossed with itself and at last one gene was exchanged by the parents. Eqs. 9-15 were used to generate new individuals by mutation, where index old defines the selected variables old value and MS defines mutation strength. Each individual from the last generation could participate in mutation with same probability.

$$x = x_{old} + randomReal\{-100, 100\} \cdot \frac{x_{max}}{x_{min}} \cdot MS \qquad (9)$$

$$Y = 0 \qquad (10)$$

$$z = z_{old} + randomReal\{-100, 100\} \cdot \frac{z_{max}}{z_{min}} \cdot MS \qquad (11)$$

$$l_a = l_{a,old} + randomReal\{-100, 100\} \cdot \frac{x_{max}}{x_{min}} \cdot MS \qquad (12)$$

$$l_b = l_{b,old} + randomReal\{-100, 100\} \cdot \frac{z_{max}}{z_{min}} \cdot MS \qquad (13)$$

$$l_c = l_{c,old} + randomReal\{-100, 100\} \cdot \frac{x_{max}}{x_{min}} \cdot MS \qquad (14)$$

$$\varphi = \varphi_{old} + randomReal\{0, \pi\} \cdot MS \qquad (15)$$

The optimization loop was repeated until a parallelepiped was found to be the best for same iteration count as defined by convergence criteria or until maximum iteration number was reached. On completion the algorithm returned the parallelepiped with the biggest volume as the optimization target. The biggest parallelepiped represents the size, position and orientation of the biggest workpiece that can be milled with the chosen accuracy with the discussed robot.

In the first run the algorithm did not converge with the defined settings. Workpieces still increased in size from generation to generation, indicating space for improvement. For this reason an efficiency mechanism was built into the code, generating a restart file if maximum iteration number was reached before convergence criteria. From the restart file a new optimization run could be started with the individuals from the last generation of the previous run adopted for the initial generation of the new optimization run. In our case, six optimization runs were performed to find the target workpiece.

Locations of two most opposite corner points of the biggest workpiece found by the applied genetic algorithm are A_x = 2000.23 mm, A_y = -1162.99 mm, A_z = 1963.87 mm and G_x = 1354.13 mm, G_y = 1162.99 mm, G_z = 292.81 mm. Edge lengths l_a, l_b and l_c are l_a = 1002.44 mm, l_b = 1486.24 mm and l_c = 2325.98 mm, where l_a is the distance between two corner points on the same edge of the parallelepiped in X direction and it is rotated by angle φ about the Y axis in counter clockwise direction in front view, l_b is the distance between two corner points on the same edge of the parallelepiped in Z direction and it is rotated by angle φ about the Y axis in counter clockwise direction in front view and l_c is the distance between two corner points on the same edge of the parallelepiped in Y direction. Workpiece rotation about the Y axis is φ = 3.37°. The volume of the workpiece, calculated as the product of its edge lengths is 3.47×10^9 mm^3.

The workpiece is presented in Fig. 7. Graphics on the left hand side show workpiece size, position and orientation in the robot's complete workspace, while graphics on the right hand side show the same workpiece in the robot's workspace without singularities. Accuracy contour is plotted in Y = 0 plane only. Dexterity in blend colours is used for workspace presentation, but for size, position and orientation search only the accuracy contour marked in black was relevant. For each case figures are presented in default, front and in top view.
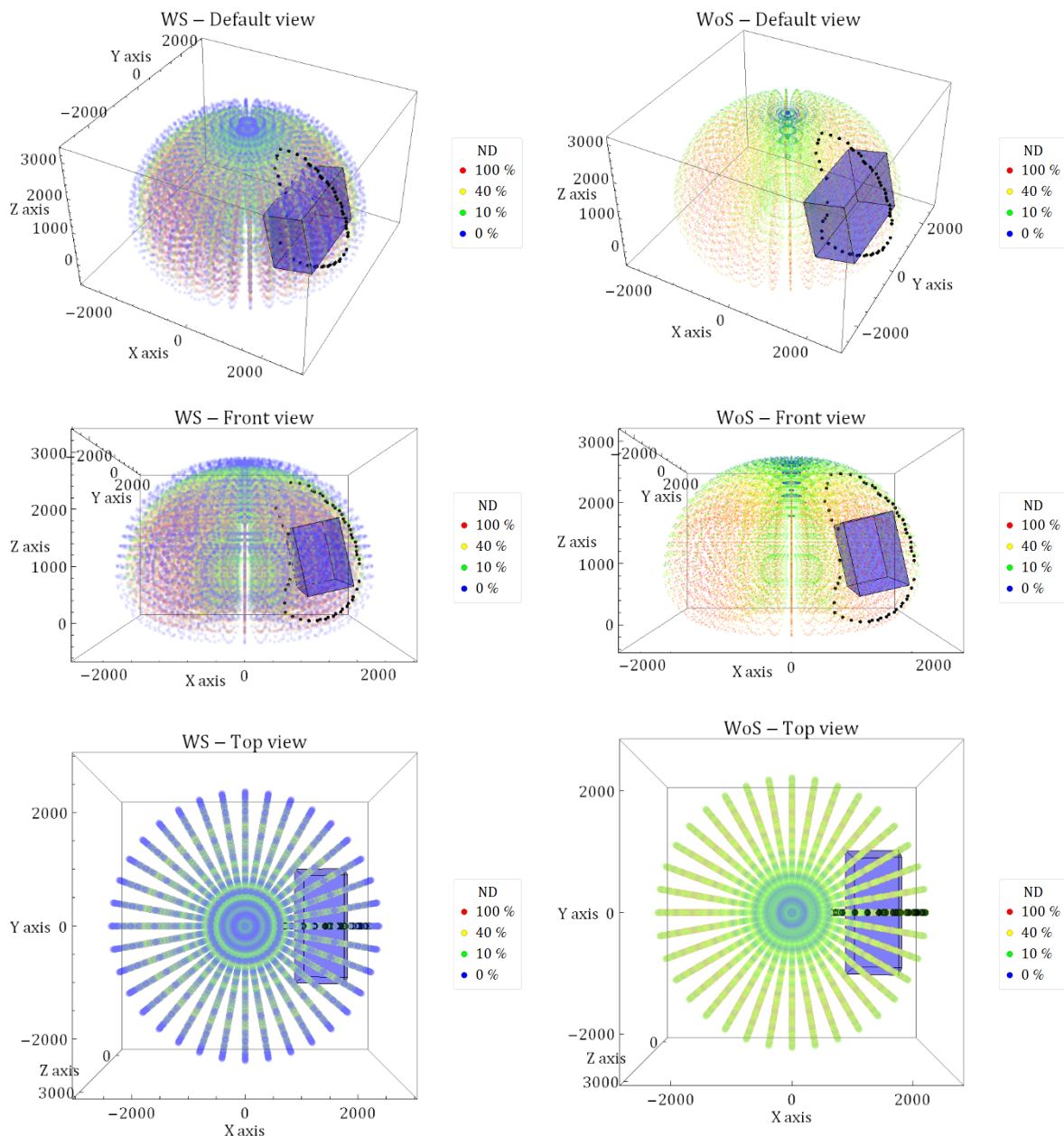
Fig. 7 Optimized workpiece in the robot's workspace with accuracy contour: with singularity points (left), without singularity points (right)

4. Conclusion

The article demonstrates a methodology, based on innovative algorithms, to isolate a refined sub region in a robot's complete workspace that allows milling with suitable accuracy. A parallelepiped shaped, maximum sized workpiece and its position and orientation in a robot's refined workspace were found as an example, but any arbitrary shaped object or a number of objects can be positioned into the refined region instead by using the proposed method.

The presented methodology can also be applied to robotic mechanisms in different manufacturing environments. Instead of defining the accuracy contour by milling measurements, other machining processes to create performance contours can be used. Contours for a robot's posture accuracy for grasping and positioning, measurement accuracy contours for tolerance control of big welded parts and many more can be created to optimally restrict a robot's workspace. Some analytical performance measures that can be used for contour generation were also discussed in this article.

It is to be noted that complex contour shapes may require a different extraction approach than the regional non-dominated sorting algorithm presented in this article. The algorithm could be used because of simplicity of the accuracy contour, containing only two inflection points that were easy to identify and was tailored for quick contour extraction as it was intended to serve as the optimization function for the search algorithm. A general contour extraction algorithm would consider each point from the initial point cloud more often and would be less suitable for workpiece size and placement optimization.

Genetic algorithm was applied for workpiece size, position and orientation search, because of good results reported for similar problems by other authors. Discard penalty was introduced for individuals violating the accuracy contour criteria. Because of the penalty each new iteration took longer to complete as workpieces from earlier generations were closer to contour boundaries and discarded more often. An algorithm with faster convergence is being developed, allowing a more general contour generation algorithm to be used as optimization function.

Acknowledgement

References

[1] Kubela, T., Pochyly, A., Singule, V. (2016). Assessment of industrial robots accuracy in relation to accuracy improvement in machining processes, In: *2016 IEEE International Power Electronics and Motion Control Conference (PEMC),* Varna, Bulgaria, 720-725, doi: 10.1109/EPEPEMC.2016.7752083.

[2] Wu, Y., Klimchik, A., Caro, S., Furet, B., Pashkevich, A. (2015). Geometric calibration of industrial robots using enhanced partial pose measurements and design of experiments, *Robotics and Computer-Integrated Manufacturing*, Vol. 35, 151-168, doi: 10.1016/j.rcim.2015.03.007.

[3] Józwik, J., Ostrowski, D., Jarosz, P., Mika, D. (2016). Industrial robot repeatability testing with high speed camera Phantom V2511, *Advances in Science and Technology Research Journal*, Vol. 10, No. 32, 86-96, doi: 10.12913/22998624/65136.

[4] Marônek, M., Bárta, J., Ertel, J. (2015). Inaccuracies of industrial robot positioning and methods of their correction, *Tehnički vjesnik – Technical Gazette,* Vol. 22, No. 5, 1207-1212, doi: 10.17559/TV-20140416123902.

[5] Jezeršek, M., Kos, M., Kosler, H., Možina, J. (2017). Automatic teaching of a robotic remote laser 3D processing system based on an integrated laser-triangulation profilometry, *Tehnički vjesnik – Technical Gazette,* Vol. 24, No. 1, 89-95, doi: 10.17559/TV-20160504230058.

[6] Papa, G., Torkar, D. (2009), Visual control of an industrial robot manipulator: Accuracy estimation, *Strojniški vestnik – Journal of Mechanical Engineering,* Vol. 55, No. 12, 781-787.

[7] Yoshikawa, T. (1985). Manipulability of robotic mechanisms*, The International Journal of Robotics Research*, Vol. 4, No. 2, 3-9, doi: 10.1177/027836498500400201.

[8] Gotlih, K., Kovac, D., Vuherer, T., Brezovnik, S., Brezocnik, M., Zver, A. (2011). Velocity anisotropy of an industrial robot, *Robotics and Computer-Integrated Manufacturing*, Vol. 27, No. 1, 205-211, doi: 10.1016/j.rcim.2010.07.010.

[9] Merlet, J.P. (2005). Jacobian, manipulability, condition number, and accuracy of parallel robots, *Journal of Mechanical Design*, Vol. 128, No. 1, 199-206, doi: 10.1115/1.2121740.

[10] Mansouri, I., Ouali, M. (2011). The power manipulability – A new homogeneous performance index of robot manipulators, *Robotics and Computer-Integrated Manufacturing*, Vol. 27, No. 2, 434-449, doi: 10.1016/j.rcim.2010.09.004.

[11] Vosniakos, G.-C., Matsas, E. (2010). Improving feasibility of robotic milling through robot placement optimisation, *Robotics and Computer-Integrated Manufacturing*, Vol. 26, No. 5, 517-525, doi: 10.1016/j.rcim.2010.04.001.

[12] Aly, M.F., Abbas, A.T. (2014). Simulation of obstacles' effect on industrial robots' working space using genetic algorithm, *Journal of King Saud University – Engineering Sciences*, Vol. 26, No. 2, 132-143, doi: 10.1016/j.jksues.2012.12.005.

[13] dos Santos, R.R., Steffen, V. Jr., Saramago, S.F.P. (2008). Robot path planning in a constrained workspace by using optimal control techniques, *Multibody System Dynamics*, Vol. 19, No. 1-2, 159-177, doi: 10.1007/s11044-007-9059-1.

[14] Dolgui, A., Pashkevich, A. (2009). Manipulator motion planning for high-speed robotic laser cutting, *International Journal of Production Research*, Vol. 47, No. 20, 5691-5715, doi: 10.1080/00207540802070967.

[15] Lin, Y., Zhao, H., Ding, H. (2017). Posture optimization methodology of 6R industrial robots for machining using performance evaluation indexes, *Robotics and Computer-Integrated Manufacturing*, Vol. 48, 59-72, doi: 10.1016/j.rcim.2017.02.002.

[16] Kuss, A., Drust, M., Verl, A. (2016). Detection of workpiece shape deviations for tool path adaptation in robotic deburring systems, *Procedia CIRP*, Vol. 57, 545-550, doi: 10.1016/j.procir.2016.11.094.

[17] Klancnik, S., Brezovnik, S., Brezocnik, M., Balic, J., Ficko, M., Pahole, I., Buchmeister, B. (2009). Advanced concept of milling with robot, *Scientific Bulletin Series C: Fascicle Mechanics, Tribology, Machine Manufacturing Technology*, Vol. 23, No. 100, 73-78.

[18] Karimi, D., Nategh, M.J. (2016). Contour maps for developing optimal toolpath and workpiece setup in hexapod machine tools by considering the kinematics nonlinearity, *Proceedings of the Institution of Mechanical Engineers, Part B: Journal of Engineering Manufacture,* Vol. 230, No. 9, 1572-1583, doi: 10.1177/0954405415592123.

[19] Wei, H., Tang, X.-S., Liu, H. (2015). A genetic algorithm(GA)-based method for the combinatorial optimization in contour formation, *Applied Intelligence*, Vol. 43, No. 1, 112-131, doi: 10.1007/s10489-014-0633-y.

[20] Truong-Hong, L., Laefer, D.F. (2014). Octree-based, automatic building façade generation from LiDAR data, *Computer-Aided Design*, Vol. 53, 46-61, doi: 10.1016/j.cad.2014.03.001.

[21] Awrangjeb, M., Lu, G. (2015). A triangulation-based technique for building boundary identification from point cloud data, In: *2015 International Conference on Image and Vision Computing New Zealand (IVCNZ)*, Auckland, New Zealand, 1-6, doi: 10.1109/IVCNZ.2015.7761536.

[22] Veber, M. (2016). *Manipulability analysis of milling robot Kuka KR 150-2,* Master thesis, Faculty of Electrical Engineering and Computer Science, University of Maribor, Maribor, Slovenia.

Capabilities of industrial computed tomography in the field of dimensional measurements

Horvatić Novak, A.[a,*], **Runje, B.**[a], **Stepanić, J.**[a]

[a]University of Zagreb, Faculty of Mechanical Engineering and Naval Architecture, Zagreb, Croatia

A B S T R A C T

The paper discusses the capabilities of industrial computed tomography (CT) in the field of dimensional measurement of products with close tolerances. Computed tomography is a method that allows inspection and measurements of both reachable and unreachable characteristics which makes it very desirable and interesting for application in wide range of industries. In order to evaluate the quality of measurement results obtained by industrial CT, two objects with the same geometry, and made from different materials, were measured. Results obtained with CT were compared with the results obtained by coordinate measuring machine, which were considered to be reference values, and deviations between the results have been analysed. Measurements were repeated five times under repeatability conditions. Repeatability is expressed quantitatively in terms of the dispersion characteristics of the results. Statistical analysis showed that in majority of cases, there were no statistically significant differences between measurement results of equal characteristics obtained at different materials. Obtained deviations in the research could be explained by the fact that the measurements were performed at the industrial CT for general applications. Much better results can be achieved by using a metrology CT device.

1. Introduction

Computed tomography (CT) is a method that uses X-ray in order to obtain information about inner and outer geometry and characteristics of inspected objects. It is a well-known method and has been used in medicine and material inspection for over 30 years but its application in dimensional measurements began only about 10 years ago [1]. It is a method with a lot of advantages, which makes it very desirable for industrial purposes of dimensional measurement. Nowadays, requirements on precision and accuracy of production are more rigorous and ever-increasing. There is also a growing need for measurement of objects with more complex geometry and forms [2, 3]. Except standard materials recognized in the industry field, which refers mostly to metals and alloys, great importance is given to application of new materials with better properties and possibilities. The emphasis is on the use of different types of polymers, and in connection to that, different manufacturing methods. In addition to classical metal cutting methods, use of additive manufacturing technology is also increasing. Development and implementation of additive technologies requires development and application of non-destructive inspection and measurement methods. For that reason, the application of industrial CT systems is becoming the basic and foremost requirement [4].

The most significant advantage of computed tomography is the possibility of getting information about internal and external dimensions of inspected object, at the same time, with only one scanning process and without the need to destroy the object [3, 5]. Another advantage, compared to other inspection methods, is the fact that it is suitable for inspection of parts in assembled state without disassembling them [1]. This is of great importance in cases when all parts in disassembled state are manufactured correctly, but do not work properly after being assembled. Furthermore, industrial CT systems enable both dimensional measurements and material analysis to be conducted on the same model. This is especially important when new materials are used in a production process. As such, CT systems are very desirable in many different industries.

However, apart from many advantages, CT dimensional measurement method has also some disadvantages. The main problem for its usage in the field of dimensional measurement is the fact that measurement uncertainty of results is not evaluated, due to the many influential factors in the whole measurement process [6]. This means that metrological traceability is not achieved. In order to assess measurement uncertainty, influence parameters need to be identified and classified. Classification of influence parameters can be done in many different ways. Welkenhuyzen et al [7] proposed dividing influence parameters to: influence of X-ray source, influence of rotation table and workpiece, influence of detector and data processing parameters. Furthermore, Hiller and Reindl [8] divide influence parameters into five groups: CT system, method, test object, environment and human. Another classification of influence parameters can be proposed, according to the step of measurement process. Since the whole CT measurement process can be divided into three sub-processes, where the first sub-process implies scanning of the inspected part, the second one 3D model generation and the third consist of conducting measurements on reconstructed model, influence parameters can be classified into three subclasses: parameters influencing the CT scanning process, parameters influencing reconstruction process and parameters influencing measurement of the model (Fig. 1). CT dimensional measurements are limited by possibilities of CT scanning device, as well as by software tools used for reconstruction and data processing, meaning that operator has great influence on measurement results and measurement uncertainty of obtained results. Operator influence is present throughout the whole CT measurement process, e.g. during selection of CT setups or placing object on rotational table, choosing filters in 3D reconstruction and in data evaluation, and in selecting measurement approach and mathematical algorithm to fit the simple geometry objects. At the moment, use of CT device for industrial measurements largely depends on operator's experience and knowledge. For this reason, estimation of measurement uncertainty is essential and of utmost importance, as well as defining standard procedures for CT measurements.
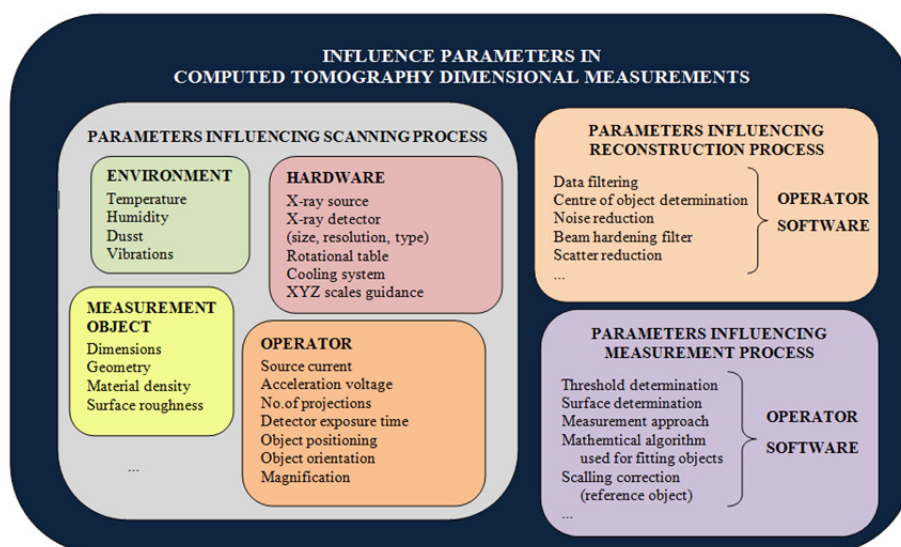
Fig. 1 Influence parameters in CT dimensional measurements

The authors [8-10] evaluate measurement uncertainty in several ways: according to GUM - *Guide to the Expression of Uncertainty in Measurement* [11], where influence of all parameters that affect measurement system has to be determined; with use of computer simulation [12] or by empirical methods that involve use of calibrated workpieces i.e. substitution method [13]. Some authors use the maximum permissible error value (MPE) as an estimator of measurement uncertainty.

This paper researches capabilities of industrial CT device for purposes of dimensional measurements. Two cylinders with the same geometry, made of two different materials were measured and observed. Chosen cylinder geometry represents object where difference in dimensions in different axes is significant, which makes it interesting for the research. Experimental research consisted of dimensional measurements of samples with usage of tactile coordinate measuring machine (CMM) and CT scanning of investigated objects. Results obtained by CT measurements are expressed and observed as deviations from tactile coordinate measurement results which are considered to be reference values. Also, statistical analyses of obtained results were conducted.

2. Materials and method

2.1 Measurement object

Measurements were conducted on two specially shaped cylinders as shown in Fig. 2.

The idea behind the design of such an object was to create an object that will allow for as many different types of measurement and geometrical characteristics as possible. Since the object material has in previous researches [7, 8] been identified as one of the major parameters that influence results, objects for this experiment have been made from two significantly different materials in terms of material density. Cylinder 1 is made of polyamide 6 (PA 6) whose density is 1.4 g/cm^3, and cylinder 2 is made of aluminium with approximately twice as high density, 2.7 g/cm^3. The object was dimensioned based on recommendations for overall penetration depth of an installed CT system [14]. Furthermore, because of the fact that measurement is conducted by fitting simple geometry objects (planes, spheres, cylinders, etc.), the idea was to construct such an object where numerous relationships between different simple objects could be investigated and measured. Except dimensional characteristics, the object is suitable for measurement and investigation of different geometrical characteristics. In this research, only dimensional characteristics were observed. Those were following six different characteristics: outer diameter D, inner diameter d, cylinder length h, distance between two holes l_1, distance between a plane and the centre of a hole l_2 and distance between planes l_3. Fig. 3 shows a drawing of measured objects with the observed measurands.

a) b)
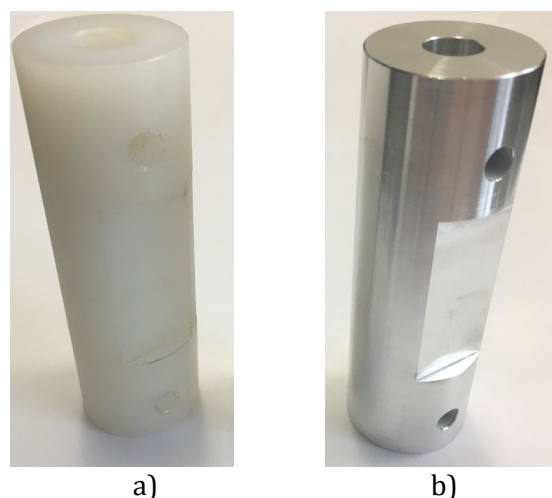
Fig. 2 Measured objects: a) Cylinder 1; b) Cylinder 2
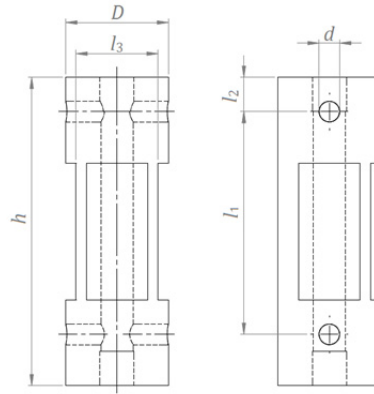
Fig. 3 Measured object with noted measurands

2.2 CMM measurement method

Reference measurements were conducted by coordinate measuring machine Tesa Micro Hite 3D shown in Fig. 4 and measurements were done in software CMM Manager 3.6.

Measurement results with related measurement uncertainties are given in Table 1. Measurement uncertainty in CMM measurements can be conducted in a few ways [15, 16], here CMM measurement uncertainty evaluation is performed in accordance with the Supplement 1 to the 'Guide to the Expression of Uncertainty in Measurement'—Propagation of Distributions Using a Monte Carlo Method which is abbreviated as JCGM 101 [12].

Fig. 4 Coordinate measuring machine Tesa Micro Hite 3D

Table 1 Reference values for cylinder

Measurand	Symbol	Measured results, mm Cylinder 1	Measured results, mm Cylinder 2	Expanded measurement uncertainty U, $k = 2$, $P = 95$ % , μm
Cylinder length	h	89.992	90.066	3
Outer diameter	D	29.986	30.016	3
Inner diameter	d	6.010	6.006	3
Distance between two holes	l_1	65.030	64.972	3
Distance between a plane and the centre of one hole	l_2	9.982	10.019	3
Distance between planes	l_3	24.005	23.932	3

2.3 CT measurement method

Whole CT measuring process consists of CT scanning process, 3D modelling process and data evaluation process. CT scanning was conducted on Nikon X TH 225 (Fig. 5), equipped with 225 kV microfocus X-ray source with a tungsten target and 14 bit Varian 4030 Flat Panel Detector [17].
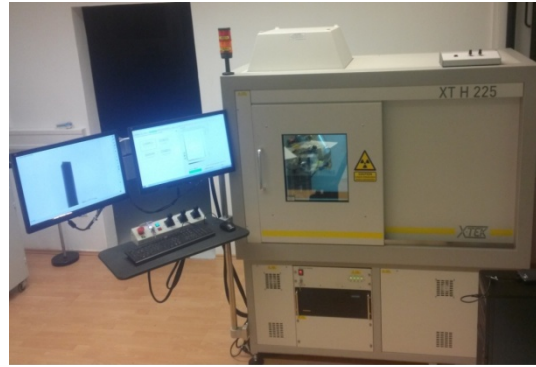
Fig. 5 Industrial CT device – Nikon X TH 225

Measurement process for each cylinder was repeated five times, while all parameters during scanning process were kept constant. Measurements were taken in the same day, by the same operator, on the same measurement device i.e. under repeatability conditions. Quality of results depends on quality of 2D images [18], which means that CT system setup is of great importance. Since there is no prescribed standard for conduction of CT measurements, scanning settings are based on operator's knowledge and experience. Considering the object's geometry, size and material, different setups were determined for each cylinder (given in Table 2). Also, a slightly tilted orientation of object during scanning process was applied.

Object was placed on polystyrene fixture, invisible for the chosen scanning setups. The same fixture was used for both cylinders ensuring the same position of cylinder 1 and cylinder 2 during scanning process. Object orientation during scanning process can be seen in Fig. 6.

Table 2 Scanning parameters

Parameter	Unit	Cylinder 1	Cylinder 2
X-ray source voltage	kV	90	205
X-ray source current	μA	45	130
Copper filter thickness	mm	-	3
Number of projections	-	1000	1000
Source detector distance	mm	984.27	984.27
Source object distance	mm	339.51	339.51
Geometrical magnification	-	2.90	2.90
Detector size	pixel × pixel	3192 × 2296	3192 × 2296
Pixel size	μm	127	127
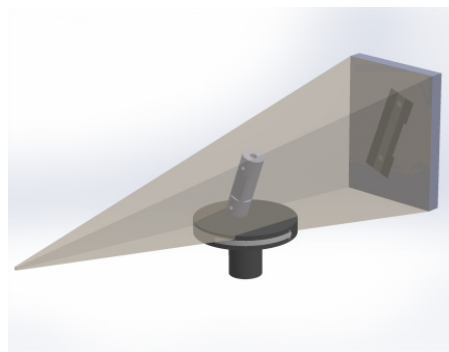
Fig. 6 Display of object during scanning process

3. Results and discussion

In order to estimate capabilities of industrial CT in the field of dimensional measurement of products with close tolerances, measurements of inspected samples were conducted using the same measurement approach in CMM and CT measurements. The approach considered fitting

simple geometry objects such as planes and cylinders using the Gaussian method, where all observations were focused on observing relations between different combinations of those two simple objects. What was observed were dimensions of outer and inner cylinders, distances between two planes, distance between two cylinders and distance between cylinder and plane, as shown in Table 3. All measurements were conducted in good thermal conditions ($t = 20\,°C ± 1\,°C$).

Scaling correction of CT data sets has been performed using the calibrated ball bar where distance between two ruby spheres on carbon rod was used to correct the nominal voxel size.

CT measurements were repeated five times and results are shown in Table 4 as the arithmetic mean of measured results $\bar{x}$ with given standard deviation s for each measurand and for each cylinder.

Table 3 Measurement strategies used for the volumetric data evaluation

Measurand	Symbol	Measurement strategy
Cylinder length	h	Plane-plane
Outer diameter	D	Cylinder
Inner diameter	d	Cylinder
Distance between two holes	l_1	Cylinder-cylinder
Distance between plane and centre of one hole	l_2	Plane-cylinder
Distance between planes	l_3	Plane-plane

Table 4 CT measurement results with given standard deviation for both cylinders

	Cylinder 1		Cylinder 2	
	$\bar{x}$ mm	s, mm	$\bar{x}$, mm	s, mm
D	29.966	0.006	30.011	0.003
d	06.019	0.002	06.013	0.002
h	89.951	0.020	90.043	0.006
l_1	65.032	0.002	64.975	0.001
l_2	09.957	0.024	10.001	0.032
l_3	24.012	0.015	23.942	0.017

Fig. 7 presents deviations between results obtained by CT and CMM measurements. Deviations are expressed as differences between CT measurements and reference CMM measurements and are given in micrometers for each observed measurand and for each cylinder.

Fig. 7 Deviation between CT and CMM results for cylinder 1 and cylinder 2

First, the diameters of simple fitted features were observed: cylinders. The outer diameters of inspected parts were observed as outer cylinders, where the obtained results were in both cases lower than reference values, meaning that all deviations were negative. Slightly better results, in terms of deviation from reference values, were obtained in cases when outer diameter of aluminium cylinder was inspected. Measurements of inner diameter of hole d, showed similar behaviour in cases when measuring inner diameter of the two cylinders. Deviations equalled approximately 10 micrometers, where amounts larger than reference values were obtained with

CT, which resulted in positive deviations. Similar results were also obtained in [8] where outer and inner diameters of observed stainless steel cylinder were inspected. Obtained deviations were in range from (−0.008 to 0.014) micrometers where also different directions of deviations have been noticed when measuring outer and inner diameters. Different directions of deviations when measuring inner, as opposed to outer diameters, can be explained through threshold value. If chosen threshold value was greater than the optimum value, outer diameters would be smaller than actual size i.e. inner diameters would be greater than real value and vice versa [19].

Secondly, distances between different combinations of simple fitted objects were observed. When observing cylinder length h, defined as distance between border objects' planes, the biggest deviation from reference value was obtained. Such result can be explained by the measurement approach where length is observed as distance between two planes parallel to X-ray source. In that case, noise appears on the borders of inspected part and in this case, presence of noise on border planes affects object length. In both cases CT results were lower than reference values which resulted in negative deviations. Furthermore, such negative deviations in cylinder length also can be attributed to the chosen threshold value.

In the case when distance between two holes was observed, characteristic l_1, small or even no deviation between obtained results was expected. No matter what the inner diameter amounts (which depends on determined threshold), the distance between two holes remains the same. This is why measurement of the distance between two holes (or two spheres) is often used for scale error correction [20]. The expectations were proved. In both cases obtained deviations were the lowest in comparison to other results.

When analyzing results of distance between a plane and a hole, marked with l_2, similar behaviour was observed as in case of cylinder length h, despite the fact that different relations between simple objects were observed. Considering the fact that l_2 is defined as the length between border plane and an inner hole (where border noise, which has a great impact on results, appeared) higher deviations in results were expected. Also, when observing length h, higher deviation was obtained in case of cylinder 1. Therefore, higher deviations were again expected in case of cylinder 1. A lot more noise was present in measuring polymer cylinder, which implies impact of object material on measurement results.

One more measurand was observed, distance (l_3) between two planes perpendicular to planes that define object's overall length. Relying on results obtained in the case when h was observed, which was also defined as distance between two planes, the same behaviour of results as in case of cylinder length was expected. Obtained deviations were positive in both cases, which is contrary to expectations. Similar behaviour was also observed in [10] where two lengths (L_T and L_F) on a dose engine made from brass were measured. The measurement approach was held in the same way i.e. lengths were determined as distances between two parallel planes. Observed distances were perpendicular to each other, same as here with distances h and l_3. Explanation can be seen in position of observed characteristic, related to beam radiation.

Furthermore, the standard deviations and arithmetic means of the two samples were compared using the F test and the T test. By applying the F test it was determined that standard deviations do not significantly differ ($p > 0.05$) except in case of measuring the cylinder length. Reason for that can be found in measurement approach. In total, five scans of each cylinder were made. Cylinder length was defined as distance between two planes which were fitted to data by random selection of points. Results are given in Table 5. Furthermore, T test was conducted and results are given in Table 6.

Table 5 Results of F test for all measurands

Measurand	p-value
Outer diameter, D	$p = 0.208$
Inner diameter, d	$p = 1$
Cylinder length, h	$p = 0.039$
Distance between two holes, l_1	$p = 0.208$
Distance between plane and hole, l_2	$p = 0.591$
Distance between two planes, l_3	$p = 0.814$

Table 6 Results of conducted T test

Measurand	p-value
Outer diameter, D	$p = 0.001$
Inner diameter, d	$p = 0.153$
Cylinder length, h	$p = 0.126$
Distance between two holes, l_1	$p = 0.347$
Distance between plane and hole, l_2	$p = 0.706$
Distance between two planes, l_3	$p = 0.775$

Because the p-values are larger than reasonable choice of $\alpha = 0.05$, there is no significant evidence to reject the null-hypothesis stating that arithmetic means are equal. In the case of measuring outer diameters, the p-value is less than alpha risk, meaning there is a significant difference between arithmetical means. In majority of cases, results of equal characteristics obtained from different materials are comparable. It can be concluded that deviations from referent values of all characteristics, except cylinder length obtained at cylinder 1, are approximately ± 25 µm. Pooled experimental standard deviation s_p was estimated and equals 16 µm. Higher deviations could be explained by the fact that the measurements were performed at an industrial CT for general applications. Significantly better results can be achieved using a metrological CT device.

4. Conclusions

Computed tomography has numerous advantages, which makes it very desirable for dimensional measurements and quality control in wide range of industries. However, lack of metrological traceability and accepted procedures still prevents its wider use. In order to define capabilities of industrial computed tomography in the field of dimensional measurements two objects with the same geometry, made from different materials, were measured. Selection of materials and dimensions as well as objects design was conducted in order to cover interesting and frequently used materials in industries.

Measurements of six dimensional characteristics were conducted using the same measurement approach in both the CMM and CT measurements. Results were observed and presented as deviations from CMM results, which were considered to be reference values. Obtained results are in agreement with previous researches. Deviations from referent values of all characteristics, except cylinder length obtained at cylinder 1, are approximately ± 25 µm. Pooled experimental standard deviation s_p was estimated and is equal to 16 µm. Higher deviations could be explained by the fact that the measurements were performed at an industrial CT for general applications. Much better results can be achieved by using a metrological CT device. Also, the standard deviations and arithmetic means of the two samples were compared using the F test and the T test. By applying the F test it was determined that standard deviations do not significantly differ ($p > 0.05$) except when measuring the cylinder length. Reason for that can be found in measurement approach. By applying the T test it was determined that arithmetic means do not significantly differ ($p > 0.05$) except in case when measuring the outer diameters. Explanation of obtained results can be given through selection of threshold value.

Considering the fact that a CT device offers simultaneous examination of more properties (dimensional characteristics, material analysis), growing application and implementation of CT systems in industry is expected. However, the measurement uncertainty of results needs to be assessed, so the further researches should be focused on its evaluation. Considering the requirements, it is advised that further researches are undertaken using Monte Carlo simulations. A bigger emphasis should be placed on identifying and eliminating systematic errors, as well as on developing measurement procedures with the aim of minimizing operator influence.

Acknowledgement

References

[1] Kruth, J.P., Bartscher, M., Carmignato, S., Schmitt, R., De Chiffre, L., Weckenmann, A., (2011). Computed tomography for dimensional metrology, *CIRP Annals – Manufacturing Technology*, Vol. 60, No. 2, 821-842, doi: 10.1016/j.cirp.2011.05.006.

[2] Klobucar, R., Acko, B. (2016). Experimental evaluation of ball bar standard thermal properties by simulating real shop floor conditions, *International Journal of Simulation Modelling*, Vol. 15, No. 3, 511-521, doi: 10.2507/IJSIMM15(3)10.356.

[3] Müller, P., Cantatore, A., Andreasen, J.L., Hiller, J., De Chiffre, L. (2013). Computed tomography as a tool for tolerance verification of industrial parts, *Procedia CIRP*, Vol. 10, 125-132, doi: 10.1016/j.procir.2013.08.022.

[4] De Chiffre, L., Carmignato, S., Kruth, J.-P., Schmitt, R., Weckenmann, A. (2014). Industrial application of computed tomography, *CIRP Annals – Manufacturing Technology*, Vol. 63, No. 2, 655-677, doi: 10.1016/j.cirp.2014.05.011.

[5] Horvatić Novak, A., Runje, B. (2016). Dimensional measurement with usage of computed tomography method, In: *Proceedings of 6th International Conference Mechanical Technologies and Structural Materials*, Faculty of Electrical Engineering, Mechanical Engineering and Naval Architecture, Split, Croatia, 51-54.

[6] Horvatić Novak, A., Runje, B., Butković, D. (2016). Influence of geometrical magnification on computed tomography dimensional measurements, In: *Annals of DAAAM International for 2016, Collection of working papers for 27th DAAAM International Symposium*, DAAAM International Vienna, Vienna, 615-623.

[7] Welkenhuyzen, F., Kiekens, K., Pierlet, M., Dewulf, W., Bleys, P., Kruth, J.-P., Voet, A. (2009). Industrial computer tomography for dimensional metrology: Overview of influence factors and improvement strategies, In: *Proceedings of the 4th International Conference on Optical Measurement Techniques for Structures and Systems: Optimess2009*, Antwerp, Belgium, 401-410.

[8] Hiller, J., Reindl, L.M. (2012). A computer simulation platform for the estimation of measurement uncertainties in dimensional X-ray computed tomography, *Measurement,* Vol. 45, No. 8, 2166-2182, doi: 10.1016/j.measurement.2012.05.030.

[9] Acko, B., Sluban, B., Tasič, B., Brezovnik, S. (2014). Performance metrics for testing statistical calculations in interlaboratory comparison, *Advances in Production Engineering & Management*, Vol. 9, No. 1, 44-52, doi: 10.14743/apem2014.1.175.

[10] Müller, P., Hiller, J., Dai, Y., Andreasen, J.L., Hansen, H.N., De Chiffre, L. (2014). Estimation of measurement uncertainties in X-ray computed tomography metrology using the substitution method, *CIRP Journal of Manufacturing Science and Technology*, Vol. 7, No. 3, 222-232, doi: 10.1016/j.cirpj.2014.04.002.

[11] BIMP (2009). JCGM 100:2008 – Evaluation of measurement data – Guide to the expression of uncertainty in measurement, from *http://www.bipm.org/utils/common/documents/jcgm/JCGM 100 2008 E.pdf*, accessed February 17, 2017.

[12] BIMP (2009). JCGM 101:2008 – Evaluation of measurement data – Supplement 1 to the "Guide to the expression of uncertainty in measurement" – Propagation of distributions using a Monte Carlo method, from *http://www.bipm.org/utils/common/documents/jcgm/JCGM 101 2008 E.pdf*, accessed February 17, 2017.

[13] ISO 15530-3:2011 (2011). *Geometrical product specifications (GPS) -- Coordinate measuring machines (CMM): Technique for determining the uncertainty of measurement*, ISO copyright office, Geneva, from https://www.iso.org/standard/53627.html, accessed March 27, 2017.

[14] NPL: X-ray computed tomography, from *http://www.npl.co.uk/science-technology/dimensional/x-ray-computed-tomography*, accessed May 7, 2017.

[15] Acko, B., McCarthy M., Haertig, F., Buchmeister, B. (2012). Standards for testing freeform measurement capability of optical and tactile coordinate measuring machines, *Measurement Science and Technology*, Vol. 23, No. 9, 1-13, doi: 10.1088/0957-0233/23/9/094013.

[16] Tasic, T., Acko, B. (2011). Integration of a laser interferometer and a CMM into a measurement system for measuring internal dimensions, *Measurement*, Vol. 44, No. 2, 426-433, doi: 10.1016/j.measurement.2010.11.002.

[17] Nikon metrology – CT scanners specifications, from *http://www.nikonmetrology.com/en EU/Products/X-ray-and-CT-Inspection/Computed-Tomography/XT-H-225-Industrial-CT-Scanning/(specifications)*, accessed November 8, 2016.

[18] Schmitt, R., Niggemann, C. (2010). Uncertainty in measurement for x-ray-computed tomography using calibrated work pieces, *Measurement Science and Technology,* Vol. 21, No. 5, doi: 10.1088/0957-0233/21/5/054008.

[19] Cantatore, A., Müller, P. (2011). *Introduction to computed tomography*, DTU Mechanical Engineering, Kgs.Lyngby, Denmark.

[20] Müller, P. (2010). *Use of reference objects for correction of measuring errors in X-ray computed tomography,* DTU Mechanical Engineering, Kgs.Lyngby, Denmark.

Comparison of 3D scanned kidney stone model versus computer-generated models from medical images

Galeta, T.[a], **Pakši, I.**[a], **Šišić, D.**[a], **Knežević, M.**[b]

[a]J.J. Strossmayer University of Osijek, Mechanical Engineering Faculty in Slavonski Brod, Croatia
[b]Department of Urology, General Hospital "Dr. Josip Bencevic", Slavonski Brod, Croatia

A B S T R A C T

With a goal to evaluate accuracy of kidney stone models created from medical images, comparison of computer-generated models against 3D scanned model is performed. Computer-generated models are made using 6 free and one commercial software for medical images obtained by computed tomography (CT) with a slice thickness of 5 mm. Digitized volume of the same kidney stone was obtained after its surgical removal and digitized using a contactless 3D scanner ATOS Compact Scan. Due to the complexity of kidney stone, the scanned reference model is not completely identical to real surgically removed stone from a patient. High maximum deviation is positioned mainly in the areas where the actual kidney stone is not scanned. The average surface deviation is in the range of 0.24354 mm to 0.44719 mm. Results reveals that the accuracy of the computer-generated models depends on quality of algorithms for tissue segmentation implemented in a particular software and on the skill of user. All software enabled us to create a 3D model of the kidney with clearly visible position of a kidney stone inside, accurate enough for planning the operation. It is possible to get a higher model accuracy by reducing the slice thickness during medical imaging; however, it increases the dose of radiation. Therefore, it is necessary to individually determine the optimum balance between the required quality of images and the amount of radiation that the patient is exposed to during recording.

1. Introduction

Medical procedures like radiological diagnosis have become less invasive and more informative. 3D visualization and image processing plays an important role in radiological diagnostics and is of major importance for many clinical disciplines [1]. The ability to create models from medical images can be used in different ways. One of them is monitoring of growth or reduce anomalies and diseases of the body. Using the same parameters when creating the model from different time periods enables easy visualization of changes in the body, allowing doctors quickly and easily diagnose the patient's condition [2]. Generation of three-dimensional models from medical images in combination with rapid prototyping technology enables the production of individual specific model tissue as well as creating custom prosthesis in a simple and fast way [3]. The reconstructed 3D models can provide valuable medical information and powerful diagnostic tool for surgeons to understand the complex internal anatomy of the patient [4]. Shim, Gunay and Shimada in their scientific paper Three-dimensional shape reconstruction of abdominal aortic aneurysm came to conclusion that specific patient's three-dimensional model can be created

with less than 5% deviation with the use of 15 CT images with a section distance of 5 mm. Such level of error is small enough for the purpose of medical diagnosis [5]. These models in combination with 3D scanning also can be use in some non-medical applications like forensic medicine, passenger safety and crash analysis [6]. In this paper we attempt to examine the efficacy and accuracy of different programs for generating 3D models from medical images. Models of kidney stone generated from medical images obtained by computerized tomography (CT) are individually compared to the reference model of kidney stone. The reference kidney stone was removed from a patient by surgery and scanned with high-precision contactless 3D scanner. For testing the accuracy, it is convenient to have a reference model. The kidney stone is very suitable since it can be recorded in high quality with medical devices, later physically removed and then measured. The kidney stone can be even visually compared with 3D printed model generated from medical images. Estimating the accuracy of the generated model provide helpful information in future approach to 3D segmentation of kidney area and in selection of software with enough capabilities.

2. Materials and methods

2.1 Inputs

Data that we used to create reference and computer model came from CT DICOM images of kidney stone but also from physically removed kidney stone from a patient. CT best images dense bones, which makes it very suitable for recording a kidney stone since stone density is mostly very similar to bone. It can be surgically removed and digitized in order to create a referent model for comparison with other models obtained from the computer generation of medical images. The necessary medical images for computer-generated models of kidney stones are made by standard abdominal two channels CT scan device Siemens Somatom Emotion 16 with a slice thickness of 5 mm. In order to create a reference digitalized volume for comparison, kidney stone had to be surgically removed from the patient's body.

Fig. 1 CT section of the abdomen with highlighted region of kidney stone

2.2 3D Scanning

For digitizing volume of kidney stone non-contact 3D scanner, ATOS Compact Scan developed by GOM GmbH was used. It has one sensor head 340 × 130 × 230 mm dimension and supports measuring areas from 35 × 30 mm to 1000 × 750 mm, allowing fast scanning while still delivering high-quality measurement data. ATOS captures an objects full surface geometry and primitives precisely in a dense point cloud or polygon mesh with point spacing from 0.021 mm to 0.615 mm. It is widely utilized in various industries, and can measure different object sizes, surface finishes and shape complexities from 450 mm to 1200 mm working distance [7]. The system is fast, has an accuracy of up to 30 μm, but is very sensitive to glare brightness [8].

The reference model is made by scanning surgically removed kidney stone. The scanning process is composed of several steps such as devices and model preparation, the scanning and processing of the obtained values in the computer program [9]. Before the scanning process, for the more accurate results, the calibration of the system is performed. Calibration is done with

the help of the calibration plate. The most important part of calibration plate are holes with larger diameter, which must be set in the view field of two cameras. The quality of the preparation directly affects the quality of the output. Optical scanner has problems to collect data points in holes of small diameter, what causes errors in the assessment of the position and diameter [10].

Selecting the best position for scanning is very important especially for complex cases such as kidney stones. Supports, which can be standard or individually designed for the needs of each scan, are often used to adjust the position of model. Sometimes it is necessary to scan the subject in many ways, and choose the best result [8].

Fig. 2 Scanning kidney stone with ATOS Compact Scan
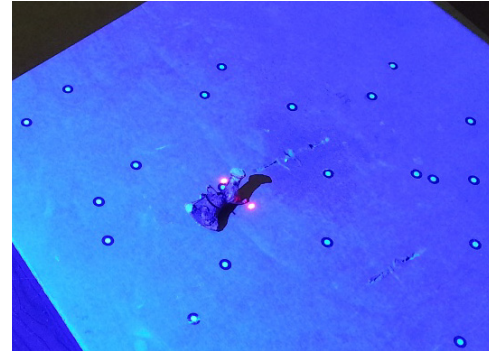
Fig. 3 Scanning a kidney stone

Fig. 2 and Fig. 3 show the way in which we recorded a kidney stone. Prior to scanning it is necessary to set the reference point. The reference points are self-adhesive labels applied to the object, support and work surface. It serves as a link between the individual subjects shooting at different angles. The reference points consist of white dots on a black background which allows them great contrast. The appearance of points, and the way they are set are also visible in presented figures.

After the object and equipment preparation, scanning is performed with the guidance of software support and manually moving the device around of the scanned object. The number of positions necessary for a full scan of the object depends on its complexity, and can amount to several times. According to Barbero B. R. and Ureta E.S. in their paper Comparative study of different digitization techniques and their accuracy, digitization on small pieces and those with sudden changes in their shape remains difficult in those parts that are visible but difficult to access. It is necessary to make several passes for their digitization, which considerably increases noise in the mesh [8]. After kidney stone is scanned, it is necessary to process the collected data. The program creates a mesh made of triangles from collected points which form a cloud. In ATOS Professional created model of kidney stone can be processed, and it is advisable to remove unnecessary parts of mesh model that is recorded by the scanner. Complete processing of object was performed in the program GOM Inspect V7.5 SR2 which is the manufacturer of the used scanner.
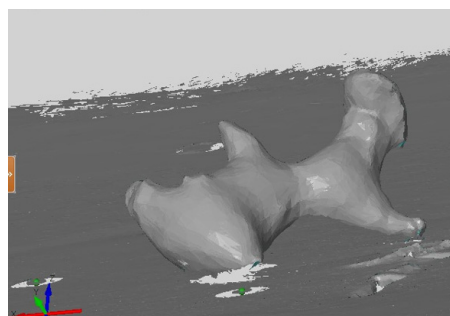
Fig. 4 Created model with all scanned parts

Fig. 4 shows the workspace of the program GOM Inspect and appearance of kidney stone model immediately after scanning. The model can have many unnecessary parts around such as supports and work surface, while at the same time there are places with missing parts.

Small stone size in combination with complex shape causes a lack of some part of information during the scanning so the referent model could not be closed completely. It takes a lot of skill and experience, to obtain usable model from such a scan.

2.3 Generating a model from medical images

To generate models from medical images requires the appropriate computer software designed for this purpose. There are a lot of available software used for medical purposes, some of them are very expensive and professional and some are free open source software. Software that were used in this work were professional 3D Doctor and 6 free open source software's: 3D Slicer, 3DIM Viewer, In Vesalius, ITK-SNAP, Mia Lite, OsiriX.

When generating models from medical images first we created the region of interest. Region of Interest – ROI is created due to the restriction of the area in the picture where we want to perform segmentation and allows us to create models only of the desired part. Segmentation is a function that automatically generates the contour of the desired object at all 2D images in which it is located. The threshold is an interactive tool, in the process of segmentation, which shows the desired object in medical images by setting upper and lower threshold brightness. Thus separates the object from the entire volume. Surface rendering is the creation of three-dimensional polygon mesh for the accurate representation of the model. A polygon mesh can be stored in different 3D formats such as IGES – Initial Graphics Exchange Specification and STL – Stereo lithography.

3D – Doctor 4.0

After loading DICOM files and previously reslice of images by one of the axes, interactive segmentation was performed. Reslice function provides visibility of certain features that would be difficult to see in the original format. Using these functions can override the limitations of the recording device. Results of the reslice function is reducing the cascade model and more accurate models. The models were created without and with the reslice function, with all other factors unchanged, in order to assess the manufacturer's claim that reslice function increases the accuracy of the model. Reslice function creates new cross sections with a much smaller slice thickness using a mathematical algorithm interpolation. Resliced image is stored on in the new image file TIFF – Tagged Image File Format. TIFF is a format for storing high-quality graphics and allows saving multiple images in a single file. After the original CT images were loaded into the program, there where available 106 sections, and only a very small number of sections where the kidney stone was visible. Reslice enabled segmentation to 511 images, and the slice thickness was reduced from 5 mm to 0.734 mm, Fig. 5.

After creating, the model is exported in STL format. The program offers a choice of types of STL files: ASCII STL or Binary STL. Due to its characteristics and advantages, the Binary STL format was selected.
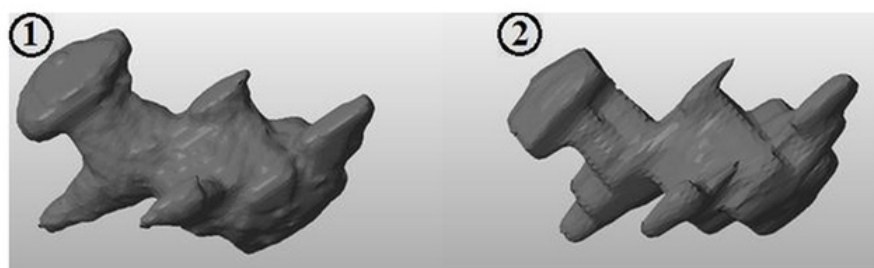
Fig. 5 3D Doctor - Left reslice model and right model without reslice function

3D SLICER 4.3.1

Module Welcome to Slicer is used to load medical images. To get the model from only a specific part of the image, it is necessary to cut the desired part. By inclusion the option ROI visibility, rectangle that indicates the part of the image that we want to maintain is set. The segmentation module requires Editor which has the tool Threshold Effect. The selected section will fill colour (green), and if the fulfilment of the section is poor it is necessary to change the boundaries to achieve the best possible fulfilment. After achieving satisfactory fulfilment, the module for creating the model is used. To create the model, we use the module Model Maker that can be called from the menu of modules or with previously used modules Editor.

Smooth option adjusts the smoothness of the model, but it affects the accuracy of the model and should be used with caution. When saving models, we cannot influence the quality and accuracy of the model, but it is possible to choose between several types of formats of 3D models such as VTK – Visualization Toolkit, PLY – Polygon File Format and STL, Fig. 6.

Fig. 6 Created model in 3D Slicer

3DimViewer 2.0

After loading DICOM files, the program automatically offers the possibility to create region of interest (ROI). The best way for segmentation is setting the threshold. Tool for setting the threshold is available without licensing additional segmentation plugins. The tool has the ability to choose the lowest level of brightness in the picture and the highest new brightness to form the model. Clicking on the edge part of kidney stone lower threshold is set while the upper is possible to leave the maximum value. Segmentation is done with command Perform Thresholding and segmented parts are painted red. Following command Upgrade Model Creation was used to create the model, by which the model is smoothed. The model obtained from current command Create Surface Model gave poor results because the model is very sharp, Fig. 7. To obtain a more accurate model the most important parameters are properly adjusted Smoothing and lower threshold brightness. When saving model, there is no possibility of further quality adjustment and selecting the type of format. The model can be saved only in the STL format. For the purposes of this study we used only publicly available commands. For better results it is advisable to buy additional modules for segmentation.
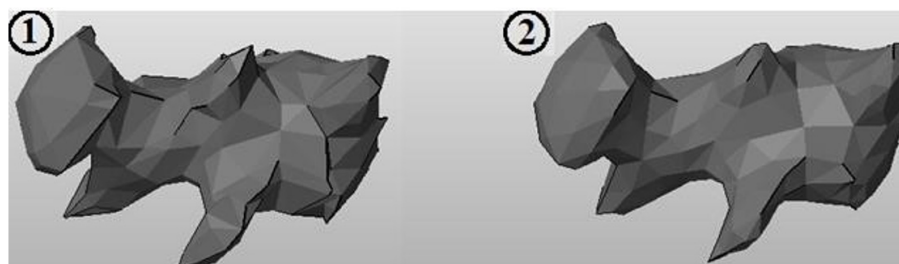
Fig. 7 Left model without upgrading and right model with upgrade Model Creation

In Vesalius 3.0 – Beta

Tools in the program are arranged in the order that is destined to create the model. After downloading the file, user needs to set option Threshold. The program offers already defined boundaries for tissues such as bones, skin and muscles, but also offers manual control. In our example, manual setting of the boundaries was used. After viewing all the sections and detailed adjustment of threshold, model was created using the command Create surface. Because the area of interest is not marked, a created model contains all other bones beside the kidney stone. The program offers several possibilities for separation: selection of largest separated area, manually selection one of the separated area, automatic separation of all separated areas and creation of separate model for each individual part. The manual selection of the separated area is selected because the kidney stone is not the largest separated area. Automatic separation of the all surfaces is not suitable because program will create model for all bones and other parts that has similar density to kidney stone. By running tools Select regions of interest and clicking on the kidney stone in the window, kidney was separated and a new model is automatically created, Fig. 8. Saving to the computer is possible in several types of formats.
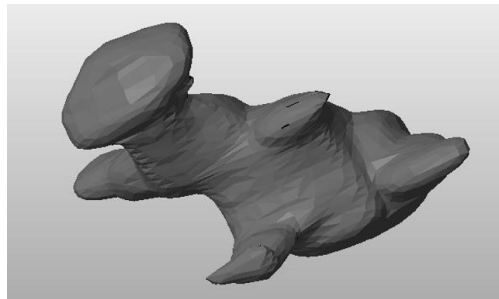
Fig. 8 The created model with In Vesalius software

ITK Snap 2.4.0

After loading the medical images, in the menu Main Toolbox the Snake ROI tool is selected. It uses interactive rectangles and allows accurate adjustment region of interest in all respects. After adjustment of the rectangle, in the Tool Options menu Segment 3D tool is selected. It initiates a process of automatic segmentation of the project that includes only part of the image selected by region of interest.

After determining the region of interest, it is necessary to process the image, in other words to separate the desired element from the rest of the image. The program works on the principle of developing a balloon from the circle to the contour that represents a kidney stone in a particular section. Spreading the contours is defined by a mathematical formula at every point, and parameters of the equation depend on the parameters of the spread curve (Snake parameters). It is necessary to set up as many balloons in the inside area of kidney stone. After setting of all parameters, begins the process of spreading balloons and creating contour that eventually form a model. Large number of iterations is carried out, and a way of creating a model is visible in steps, Fig. 9. Saving model is possible in a number of different formats of 3D models such as VTK, STL and BYU.
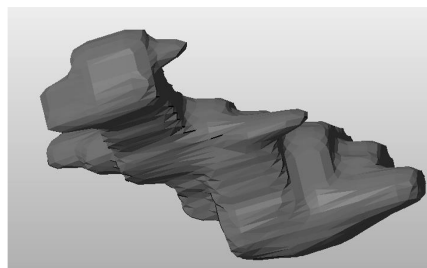
Fig. 9 Created model with ITK Snap software

Mia Lite 2.1

Adjusting smoothness of the model is done with the Smoothing Factor option. With the small changes of Smoothing Factor, a great impact on the accuracy of the model is visible, Fig. 10. On the kidney stone that is created with a very small factor of smoothness of 0.11 a large cascade of model is visible, while stone with the smoothness factor of 0.3 loses one of his parts. Therefore, it is necessary to use this factor very carefully.
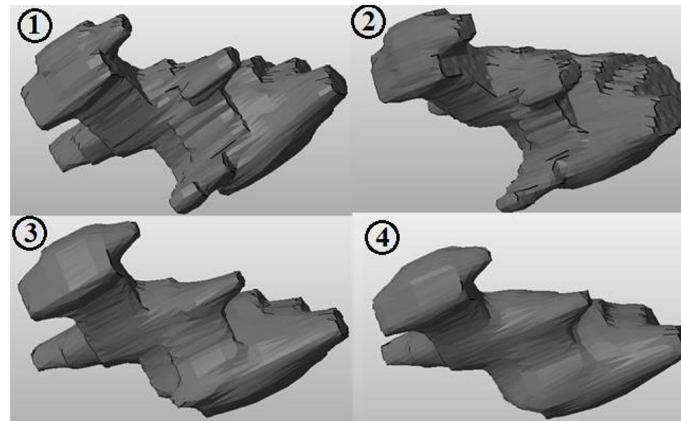
Fig. 10 Influence of smoothness factors on the model: 1) Smoothing factor 0.11; 2) Smoothing factor 0.15; 3) Smoothing factor 0.3; 4) Smoothing factor 0.42

OsiriX 5.6

After loading the medical images, tool for segmentation Grown Regions (2D / 3D Segmentation) is started. The tool includes several algorithms for segmentation. Before making a model, it is necessary to change the pixel values outside the area of interest, so that the created model is limited to the desired area, Fig. 11. This would suggest deletion of all the rest of the image and sections that do not contain created area of interest. Tool used for changing the pixel values was Set Pixel Values to. The brightness level of all pixels outside the ROI is set to the value -1024 which represents the absolute black level of brightness. Software perceived that as a void.
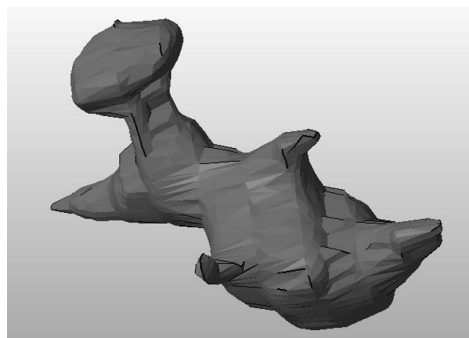
Fig. 11 The model created in OsiriX in the Microsoft Windows environment

2.4 Comparison in GOM inspect

For comparison of the scanned kidney stone and models obtained from medical images, a computer program GOM Inspect V7.5 SR2 is used. GOM Inspect is a free computer program developed by GOM GmbH. In addition to serving as a computer support for 3D scanning process, it has the ability to repair the scanned model, to measure it and to compare scanned model with other models. Due to the complex shape and size of the kidney stone there are some differences between the scan and the actual model that should be taken into account when considering the results of the comparison. In GOM Inspect scanned models have a name Actual Elements and they are shown in grey colour, while all other elements subsequently imported into the program

are marked as Nominal Elements and shown in blue colours, Fig. 12. Before comparing the actual and nominal elements, alignment of elements is executed. Alignment can be divided into pre-alignment and main settlement. Each alignment has several methods that can be selected depending on the needs. Tool for pre-alignment contains several options.

Option CAD allows selection of nominal elements, while option Actual mesh selects the actual elements that will go into the process of alignment. After aligning, the models were analysed with tools for inspection, and reports were created. Search time allows to select a time for calculating the alignment. To place the elements in proper form we can use additional help point function and set a several points on the same positions on real and nominal element. Using the function Compute additional best fit, software calculates the best position for precise alignment. After the pre-alignment, software provides information on the average deviation surface of models. This deviation depends on the performance of tools for alignment, and is not the reference data about the overall accuracy of the model. Because the model has very complex shape the main alignment cannot be ascertained, i.e. it does not contribute to better alignment. The only possible inspection of model is a surface deviation between the real and the nominal model.
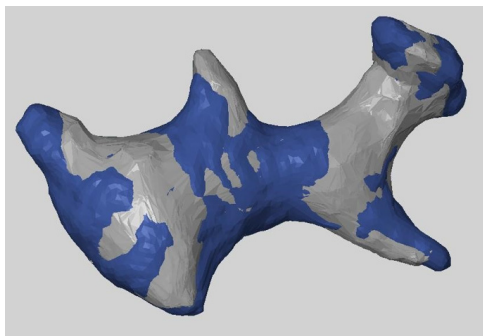
Fig. 12 Aligned Actual element (scanned model) in grey color and Nominal element (imported model) in blue
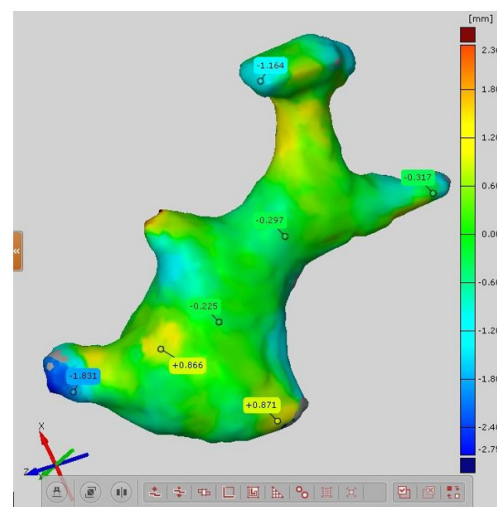
Fig. 13 Deviation with specific values

The function Separate surface comparison per CAD group allows the separation of deviations inspection for each loaded nominal model. The function Actual mesh selects the actual elements that will enter the inspection process. With the parameter Max. distance determines the maximum distance between the points of real and nominal element that will be taken into account. All deviations greater than the maximum distance parameters are not taken into account and they will be indicated in grey on visual display. Tool for inspection of area deviation displays the results in visual form on the silhouette of the nominal model and the range of colours shows the deviation values. Positive values indicate the nominal model surface that is below the surface of the real model, while negative values indicate the nominal model of the surface that is above the actual model. The grey colour indicates the area that is not compared because of the lack of the part of real model.

Tool Deviation Label sets points on a visual representation of where we want to know exactly the amount of deviation. Fig. 13 shows a visual representation of the variation in several points with specific values in millimetres.

3. Results and discussion

To bring the conclusions, we analyse several parameters for each software used in this paper, Table 1. For better visibility, the test results are grouped in Tables 2 and 3. Software showed that despite the simplicity of generating model, setting threshold and smoothness has the major impact on model accuracy. For exact model the parameters of the lower threshold and Smoothness

option factors needs to be set, but each parameter has a very large impact on the model and is difficult to coordinate them. The lower threshold level gives a more accurate model of the shape but increases the model dimensions and make him very cascade. A smoothness factor has a significant impact on model shape. Even small increase of the smoothness factor removes parts of the model. The picture shows the influence of smoothness factor on model shape with a constant parameter of the lower threshold of brightness that is 140.

Model 3 from Fig. 14 was selected as the most accurate because it contains all the parts of the kidney stone although GOM Inspect is not rated it as a model with the lowest average deviation of surface. The results of comparison with OsiriX software show satisfactory deviation on large area of the model, except where the program created a dent deeper than 8 mm. Visual inspection reveals mathematical nature of the dent and that it occurred during the interpolation while model was generated. The largest surface area difference between the used software's is 8.16 cm^2, largest model volume difference is 2.02 cm^3. Maximum deviation above the surface of the real model is 9.983 mm (3D-Doctor) and 8.457 mm under the actual model (MiaLite).
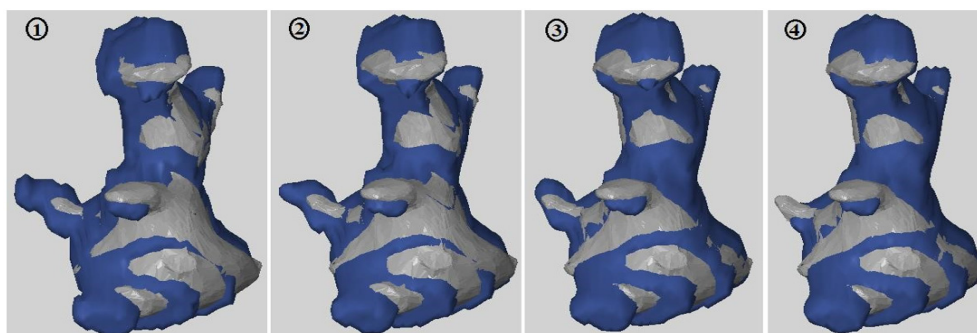
Fig. 14 Models obtained by software Mia Lite – 1) Smoothing factor 0.5; 2) Smoothing factor 0.1; 3) Smoothing factor 0.12; 4) Smoothing factor 0.15

Table 1 Result comparison of the surface area, volume and deviation for all models

Software name	Surface area (cm^2)	Model volume (cm^3)	The average surface deviation (mm)	The maximum deviation of the area under the actual model (mm)	Maximum deviation above the surface of the real model (mm)
3D-Doctor	20.67	4.68	0.24354	7.583	9.983
Slicer	22.20	5.56	0.43737	3.303	9.953
3DimViewer	24.23	5.51	0.29845	3.350	6.751
InVesalius	21.22	4.92	0.24515	2.359	9.480
ITK-SNAP	25.85	6.35	0.44607	8.039	9.612
MiaLite	28.83	6.70	0.44719	8.457	8.086
OsiriX	22.53	5.00	0.25507	2.387	8.483

Table 2 Results for 3D-Doctor, Slicer and 3Dim Viewer software

Software	3D-Doctor		Slicer	3DimViewer
	With Reslice	Without Reslice		
Length / width / model height [mm]	33.69/29.32/40.83	33.33/28.92/39.84	32.52/32.04/29.50	34.37/31.68/44.67
Volume of model (cm^3)	4.68	4.49	5.56	5.51
Surface area (cm^2)	20.67	21.34	22.20	24.23
Number of triangles	11556	4408	3074	1174
Used segmentation	Interactive Segment…	Interactive Segment…	Module: Editor; Option: Threshold Effect	Automatic Segmentation
The lower limit of brightness	1458	1458	300	280
Smoothness factor	no	no	10	4
Additional options / parameters	Used Option: Reslice	no	no	no
The average surface deviation (mm)	0.24354	0.33978	0.43737	0.29845
The maximum deviation of the area under the actual model (mm)	7.583	8.285	3.303	3.350
Maximum deviation above the surface of the real model (mm)	9.983	9.158	9.953	6.751

Table 3 Results for In Vesalius, ITK SNAP, Mia Lite, Osir X

Software	In Vesalius	ITK-SNAP	MiaLite		OsiriX
			MiaLite - selected	MiaLite – minDev	
Length / width / model height (mm)	33.44/29.85/41.26	33.34/30.48/45.00	33.81/30.80/44.98	32.98/29.32/44.89	31.12/43.42/34.05
Volume of model (cm³)	4.92	6.35	6.70	4.49	5.00
Surface area (cm²)	21.22	25.85	28.83	21.65	22.53
Number of triangles	2425	4672	5076	4000	4354
Used segmentation	Manual threshold	Intensity regions	Threshold	Threshold	Threshold (lower/upper bounds)
The lower limit of brightness	400	158	140	1458	100
Smoothness factor	no	3.55	0.12	0.10	30
Additional options / parameters	no	Balloon force: 0.9 Curvature force: 0.26	no	no	Resolution: 100 % Decimate: 0.1
The average surface deviation (mm)	0.24515	0.44607	0.44719	0.30702	0.25507
The maximum deviation of the area under the actual model (mm)	2.359	8.039	8.457	9.262	2.387
Maximum deviation above the surface of the real model (mm)	9.480	9.612	9.998	8.086	8.483
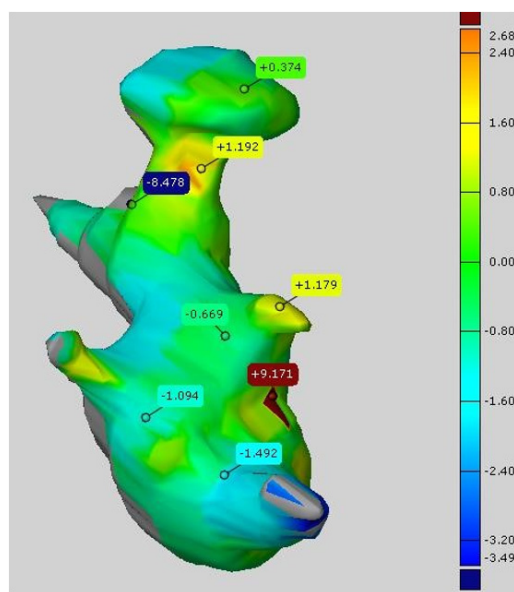
Fig. 15 Surface matching of nominal and real model on the dent area

4. Conclusion

Determination of the most accurate computer generated model is very challenging task, because beside numerical values the deviations of models should be also visually inspected. Shape of the computer-generated models must be compared with the actual kidney stone shape to decide which deviation can be ignored. Additive technologies like 3D printing, which are nowadays more and more in the application, provide the possibility of 3D printing generated models in order to make the visual comparison. 3D printed models obtained from medical imaging are useful when planning surgical procedures. Analysis of the results showed that the software for generating models from medical images could get models with acceptable accuracy for planning medical procedures. High maximum deviation, above and below the surface of the real model, are positioned mainly in the areas where the actual kidney stone is not scanned because of his small size, complex shape and limitations of Atos scanning device. Deviation occurs because of

the inability of computer program GOM Inspect to determine which part of the model surface belongs to not scanned part of the kidney stone. Not all used programs have the ability for processing and precise joining of medical images obtained from multiple directions. 3D-Doctor is the only used software that has the ability to manually merge images from different directions. We can assume that specific algorithm for merging images from multiple directions would increase the accuracy of the resulting model. In a future research we have considered the development of an interpolation algorithm from several sections directions. Created algorithm will be used for an error analysis after automatically merging images from axial, sagittal and coronal plane.

Acknowledgement

References

[1] Rengier, F., Mehndiratta, A., von Tengg-Kobligk, H., Zechmann, C.M., Unterhinninghofen, R., Kauczor, H.-U., Giesel, F.L. (2010). 3D printing based on imaging data: Review of medical applications, *International Journal of Computer Assisted Radiology and Surgery,* Vol. 5, No. 4, 335-341, doi: 10.1007/s11548-010-0476-x.

[2] Ruggeri, M., Tsechpenakis, G., Jiao, S., Jockovich, M.E., Cebulla, C., Hernandez, E., Murray, T.G., Puliafito, C.A. (2009). Retinal tumor imaging and volume quantification in mouse model using spectral-domain optical coherence tomography, *Optics Express,* Vol. 17, No. 5, 4074-4083, doi: 10.1364/OE.17.004074.

[3] Robiony, M., Salvo, I., Costa, F., Zerman, N., Bazzocchi, M., Toso, F., Bandera, C., Filippi, S., Felice, M., Politi, M. (2007). Virtual reality surgical planning for maxillofacial distraction osteogenesis: The role of reverse engineering rapid prototyping and cooperative work, *Journal of Oral and Maxillofacial Surgery,* Vol. 65, No. 6, 1198-1208, doi: 10.1016/j.joms.2005.12.080.

[4] Chougule, V.N., Mulay, A.V., Ahuja, B.B. (2014). Development of patient specific implants for Minimum Invasive Spine Surgeries (MISS) from non-invasive imaging techniques by reverse engineering and additive manufacturing techniques, *Procedia Engineering,* Vol. 97, 212-219, doi: 10.1016/j.proeng.2014.12.244.

[5] Shim, M.-B., Gunay, M., Shimada, K. (2009). Three-dimensional shape reconstruction of abdominal aortic aneurysm, *Computer-Aided Design,* Vol. 41, No. 8, 555-565, doi: 10.1016/j.cad.2007.10.006.

[6] Thali, M.J., Braun, M., Dirnhofer, R. (2003). Optical 3D surface digitizing in forensic medicine: 3D documentation of skin and bone injuries, *Forensic Science International,* Vol. 137, No. 2-3, 203-208, doi: 10.1016/j.forsciint. 2003.07.009.

[7] ATOS Compact Scan – GOM, from *http://www.gom.com/metrology-systems/system-overview/atos-compact-scan.html*, accessed September 8, 2015.

[8] Barbero, B.R., Ureta, E.S. (2011). Comparative study of different digitization techniques and their accuracy, *Computer-Aided Design,* Vol. 43**,** No. 2, 188-206, doi: 10.1016/j.cad.2010.11.005.

[9] Tóth, T., Živčák, J. (2014). A comparison of the outputs of 3D scanners, *Procedia Engineering,* Vol. 69, 393-401, doi: 10.1016/j.proeng.2014.03.004.

[10] Gapinski, B., Wieczorowski, M., Marciniak-Podsadna, L., Dybala, B., Ziolkowski, G. (2014). Comparison of different method of measurement geometry using CMM, optical scanner and computed tomography 3D, *Procedia Engineering,* Vol. 69, 255-262, doi: 10.1016/j.proeng.2014.02.230.

[11] Brisbane, W., Bailey, M.R., Sorensen, M.D. (2016). An overview of kidney stone imaging techniques, *Nature Reviews Urology*, Vol. 13, 654-662, doi: 10.1038/nrurol.2016.154.

[12] Budzik, G., Burek, J., Bazan, A., Turek, P. (2016). Analysis of the accuracy of reconstructed two teeth models manufactured using the 3DP and FDM technologies, *Strojniški vestnik – Journal of Mechanical Engineering*, Vol. 62, No. 1, 11-20, doi: 10.5545/sv-jme.2015.2699.

[13] Mandić, M., Galeta, T., Raos, P., Jugović, V. (2016). Dimensional accuracy of camera casing models 3D printed on Mcor IRIS: A case study, *Advances in Production Engineering & Management*, Vol. 11, No. 4, 324-332, doi: 10.14743/apem2016.4.230.

Simulation of collaborative product development knowledge diffusion using a new cellular automata approach

Kunpeng, Y.[a], Jiafu, S.[b,*], Hui, H.[b]

[a]Managemengt College of Ocean University of China, Qingdao, P.R. China
[b]Chongqing Technology and Business University, Chongqing Key Laboratory of Electronic Commerce & Supply Chain System, Chongqing, P.R. China

A B S T R A C T

In order to quantitatively examine the diffusion process and pattern of collaborative product development (CPD), this paper puts forward a quantitative research model of CPD knowledge diffusion based on improved cellular automata. In light of the idea of SIS epidemic model and the local knowledge interaction characteristic of CPD knowledge diffusion, the influencing factors of knowledge diffusion are abstracted into the parametric variables in the process of knowledge diffusion, and the knowledge-SIS (K-SIS) model is constructed based on improved cellular automata for CPD knowledge diffusion. Finally, the K-SIS model is simulated to study the diffusion process and pattern of CPD knowledge, revealing the influence mechanism of CPD knowledge diffusion influencing factors on the diffusion process. The research results provide valuable reference for improving the efficiency of CPD knowledge diffusion.

1. Introduction

With the increasingly fierce competition in the product market, especially IT industry, pharmaceutical industry and automobile industry, enterprises are attaching more importance to the integration of suppliers, customers and other potential collaborators into product development. Owing to the deep integration of the collaborators' knowledge, collaborative product development (CPD) has become an new product innovation mode practiced by many enterprises, such as Apple, Xiaomi and LEGO, etc. The collaborative product development system (CPDS) consists of such collaborative members as suppliers, customers, other potential collaborators and enterprise professional product developers. Knowledge exchange and diffusion are prevalent in the CPDS owing to the asymmetry of the members in the structure of collaborative production development knowledge (CPD knowledge) and imbalance between them in the level of knowledge stock [1]. CPD knowledge diffusion enables each member to fully access and acquire the knowledge of others, thereby increasing the CPD knowledge stock of individual member and the entire CPDS. Meanwhile, the diffusion of CPD knowledge helps the members complement each other's advantages through the diffusion of CPD knowledge, optimizes the allocation of CPD knowledge resource, enhances the technical content of the CPD and accelerate the process of knowledge innovation and product development [2]. Therefore, the efficient knowledge diffusion opens up an important way to improve the product development capacity of the enterprises, and provides a key support to the successful development of new products.

As an integral part of knowledge management, knowledge diffusion has been followed and studied by many scholars. Probing into the effect of social cohesion and network size on knowledge diffusion, Reagans and McEvily argue that it is easier to diffuse knowledge if the members of society keep closer ties and shorter distances [3]. Kim and Park explore the relationship between the structure of collaborative organization network and knowledge diffusion, suggesting that the small-world network is the most fair and efficient collaborative network structure for knowledge diffusion [4]. Setting out from the motive and impetus to knowledge diffusion, Li et al. point out that the knowledge potential is an important determinant of the speed and breadth of knowledge diffusion [5]. Based on the philosophy of epidemiology, Bass establishes the "epidemic" model innovation diffusion, and expresses the model with mathematical equations [6]. From the perspective of NW small-world network, Sun and Wei build the knowledge diffusion model of high-tech enterprise alliance, and explore the effect of network clustering coefficient, characteristic path length and exchange frequency on the knowledge diffusion of the enterprise alliance [7]. Meng et al. adopt the multi-agent model of disease transmission to simulate the knowledge diffusion process in the network environment [8].

Focusing on the influencing factors, processes and models of knowledge diffusion, the above-mentioned literatures share two common defects: Firstly, most of them concentrate on the social network, the interior and exterior of enterprises, industrial clusters, R & D team, etc. but few pays attention to the diffusion of product development knowledge in the collaboration environment (e.g. CPDS). Secondly, based on mathematical methods, system dynamics and other theoretical methods, the majority of knowledge diffusion models lay too much stress on the macro features like the speed and process of knowledge diffusion, and largely ignore the microscopic basis that the knowledge diffusion is the result of the knowledge activities between the knowledge subjects. Knowledge diffusion is a complex process in which organized and complex knowledge behaviors are formed through the local knowledge interaction between the knowledge subjects [9]. In light of the above, this paper intends to study the CPD knowledge diffusion by the complex system modeling method: cellular automata (CA). Targeted at the complex and inenarrable process of CPD knowledge diffusion, the author draws on the idea of SIS epidemic model, describes the knowledge exchange activities between knowledge subjects and collaborative teams, and illustrates the macro knowledge diffusion phenomenon of the entire CPDS. From the micro-level to the macro-scale, the description and illustration are clear and intuitive. On this basis, the traditional CA model is improved, and the quantitative model of CPD knowledge diffusion is constructed based on improved CA. The model is used to examine the process and pattern of CPD knowledge diffusion, revealing the influence mechanism of CPD knowledge diffusion influencing factors on the diffusion process.

This paper aims to quantitatively analyze the CPD knowledge diffusion process and effectively predict the diffusion trend, which can help managers to better improve the management performance of CPD.

2. Analysis of CPD knowledge diffusion process

Knowledge diffusion refers to the transfer and sharing of knowledge between different subjects across time and space. In the context of CPD, the knowledge in the CPDS is diffused between different knowledge subjects, between knowledge subjects and collaborative teams, and between different collaborative teams [10, 11]. During the diffusion of CPD knowledge, the knowledge exchange happens between knowledge receivers and knowledge transmitters. Based on their own demand of knowledge and understanding of transmitters' knowledge resources, receivers seek for in-depth exchange with transmitters to acquire valuable knowledge. Then, the receivers digest and absorb the acquired knowledge, internalize it into their own knowledge, and transform themselves into knowledge transmitters, aiming to spread the acquired knowledge to other subjects. In addition, the CPDS is a virtual organization involving multiple units and subjects. In the system, the CPD activities are mainly implemented by virtual teams that closely collaborate with each other. Thus, knowledge exchanges also take place between subjects belonging to different teams, that is, knowledge diffusion occurs both inside the teams

and between the teams. In this way, the CPD knowledge is eventually diffused and shared across the CPDS.

3. Construction of CPD knowledge diffusion model

3.1 Cellular automata

Cellular automata (CA) is a network dynamics model discrete in time and space. The CA is composed of a finite number of locally interacting cells. At a certain moment, the state of a cell only depends on its own state and the state of neighborhood cells. As time goes, the simple local rule between the cells can evolve into the complex global behavior of the macro system [12-14]. The "evolution from simple local rule to complex global behavior" is one of the unique strengths of the CA model. Once it is applied to knowledge diffusion, the model will be able to depict the phenomenon of knowledge diffusion in real system from the microscopic angle: simulate the local knowledge exchanges between subjects with simple rules and evolve into the macroscopic results of global knowledge diffusion. Through the control of the initial parameters, the model can simulate the diffusion process of different types and forms of knowledge, and explain the influence of factors like organizational characteristics and knowledge subject features in the knowledge diffusion process. Therefore, the CA model is an ideal choice for CPD knowledge diffusion simulation.

The CA can be expressed by a four-tuple:

$$CA = (C, Q, V, F) \tag{1}$$

where C is the cell space; Q is the cell state set; V is the cell neighborhood; F is the cell state transition rule.

3.2 CA model of CPD knowledge diffusion

A large number of studies have shown that the spread of social phenomena is an infection process [15]. In this research, the CPDS members in possession of a specific piece of CPD knowledge are regarded as "infectors" and those who do not possess such knowledge are viewed as "susceptibles". For a specific type of CPD knowledge, the "infectors" can "infect" the "susceptibles" with the knowledge so that the latter acquire the knowledge and the ability to "infect" others with the knowledge. In the meantime, the knowledge "infectors" may give up the knowledge because of their memory ability. There is a certain probability for the "infectors" to transform into "susceptibles" by forgetting the knowledge. The transformation is the "restoration of health". Here, an "infector" is denoted as I and a "susceptible" is denoted as S. In light of the CA model proposed above, this paper names the CPDS knowledge diffusion model as the Knowledge-SIS (K-SIS) model.

According to the four elements of the four-tuple expression of the CA, the K-SIS model is constructed as below.

Cell space C: Let C be a 2D cell space containing *nxn* cells, representing the entire CPDS. The cells in C are expressed as $c(i,j)$, representing the CPD teams in the CPDS. Hence, can be expressed as:

$$C = \{c(i,j)|1 \leq i \leq n, 1 \leq j \leq n\} \tag{2}$$

As discussed before, the CPD are mainly implemented by virtual teams in the CPDS, and knowledge exchanges occur between different subjects and collaborative teams. Hence, this paper sets each cell as a CPDS collaborative team, each containing a certain number of knowledge subjects.

Cell state set Q: By the above definition, each cell represents a collaborative team containing a certain number of knowledge subjects. Thus, the cell state can be expressed by the proportions of knowledge infectors and susceptibles in the cell. Let $S_{c(i,j)}^t$ be the proportion of knowledge

susceptibles in cell $c(i,j)$ at the moment t, and $I_{c(i,j)}^t$ be the proportion of knowledge infectors, and we have:

$$S_{c(i,j)}^t + I_{c(i,j)}^t = 1 \tag{3}$$

The state of a cell can be expressed by a two-tuple $q_{c(i,j)}^t = (S_{c(i,j)}^t, I_{c(i,j)}^t) \in Q$.

Neighborhood V: This paper uses a Moore neighborhood with a 2-radius. As shown in Fig. 1, the neighborhood of the central black cell is expressed by the area of the grids marked by dotted lines.
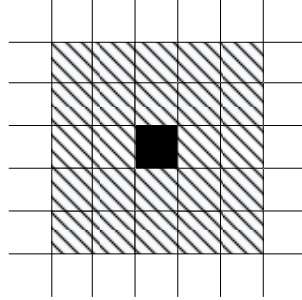
Fig. 1 Moore neighborhood with a 2-radius

Transition rule F: During the diffusion of CPD knowledge, there is a certain probability for the susceptibles to transform into infectors after acquiring knowledge from the infectors via knowledge exchanges. The infectors are either internal or external to the team. Meanwhile, the infectors are transforming into susceptibles at a certain probability by forgetting the knowledge. For the K-SIS model, the core objective is to obtain the proportions of susceptibles and infectors $S_{c(i,j)}^t$ and $I_{c(i,j)}^t$ in the cell $c(i,j)$. Through the analysis of CPD knowledge diffusion process and its influencing factors, the cell state transition rule F can be expressed as:

$$S_{c(i,j)}^t = S_{c(i,j)}^{t-1} + \delta \cdot I_{c(i,j)}^t - \frac{\rho_{In} \cdot S_{c(i,j)}^{t-1} \cdot I_{c(i,j)}^{t-1}}{v_{In}} - S_{c(i,j)}^{t-1} \cdot \sum_{c(i+\alpha,j+\beta)\in V} \frac{N_{c(i+\alpha,j+\beta)}}{N_{c(i,j)}} \cdot \mu_{\alpha\beta}^{ij} \cdot I_{c(i+\alpha,j+\beta)}^{t-1} \tag{4}$$

$$S_{c(i,j)}^t = (1-\delta) \cdot I_{c(i,j)}^t + \frac{\rho_{In} \cdot S_{c(i,j)}^{t-1} \cdot I_{c(i,j)}^{t-1}}{v_{In}} + S_{c(i,j)}^{t-1} \cdot \sum_{c(i+\alpha,j+\beta)\in V} \frac{N_{c(i+\alpha,j+\beta)}}{N_{c(i,j)}} \cdot \mu_{\alpha\beta}^{ij} \cdot I_{c(i+\alpha,j+\beta)}^{t-1} \tag{5}$$

Where $c(i+\alpha, j+\beta)$ is a neighborhood cell of the cell $c(i,j)$; $N_{c(i+\alpha,j+\beta)}$ and $N_{c(i,j)}$ are the number of knowledge subjects in the cell; $\delta \in [0,1]$ is the knowledge forgetting rate of CPD knowledge subjects; $\rho_{In} \in [0,1]$ is the intra-team trust level between the CPD knowledge subjects; $v_{In} \in [1,9]$ is the intra-team knowledge stickiness; $\mu_{\alpha\beta}^{ij} = \frac{\rho_{out} \cdot c_{\alpha\beta}^{ij}}{v_{out}}$ is the inter-team transfer and diffusion ability of CPD knowledge; $\rho_{out} \in [0,1]$ is the inter-team trust level between the CPD knowledge; $v_{out} \in [1,9]$ is the inter-team knowledge stickiness; $c_{\alpha\beta}^{ij} \in [0,1]$ is the inter-team collaboration strength.

For the inter-team collaboration strength, this paper adopts the second-order extended Moore neighborhood, which naturally leads to the concept of "cell distance". If the coordinates of a cell in the cell space is represented with a pair of integers (x,y), the distance between the cells $c(i,j)$ and $c(i+\alpha, j+\beta)$ can be expressed in Euclidean distance.

$$d_{\alpha\beta}^{ij} = \sqrt{(x_i - x_{i+\alpha})^2 + (y_i - y_{i+\alpha})^2} \tag{6}$$

The inter-team collaboration strength in the CPDS is defined as the reciprocal of the cell distance.

$$c_{\alpha\beta}^{ij} = 1 \Big/ d_{\alpha\beta}^{ij} \tag{7}$$

4. Simulation analysis of CPD knowledge diffusion

4.1 Settings of simulation parameters

Based on the K-SIS knowledge diffusion model, the trust level between CPD knowledge subjects, knowledge stickiness, the inter-team collaboration strength and knowledge forgetting rate are set as the parameters of the simulation experiment. The simulation aims to explore the effect of these influencing factors on CPD knowledge diffusion. Suppose the CPD knowledge diffusion cell space contains 20*20 cells, each of which has 30 knowledge subjects; at the initial moment of knowledge diffusion, the cell $c(10,10)$ is the only one that contains knowledge infectors, i.e. only one team in the CPDS possesses a specific type of CPD knowledge. The initial state is denoted as $q_{c(10,10)}^0 = (0.3, 0.7)$, and the simulation time $T = 100$.

In this research, the level of CPD knowledge diffusion performance is measured by the average knowledge level and the standard deviation of knowledge level [16, 17]. At the moment, the average knowledge level of the collaborative teams within CPDS is:

$$KP_t = \frac{1}{N} \sum_{i \in V} l_{i,t} \tag{8}$$

where N is the number of collaborative teams; $l_{i,t}$ is the amount of knowledge possessed by team i at the moment t.

At the moment t, the standard deviation of knowledge level $K\sigma_t$ reflecting the uniformity of knowledge possessed by each team is:

$$K\sigma_t = \sqrt{\frac{1}{N} \sum_{i \in V} l_{i,t}^2 - KP_t^2} \tag{9}$$

4.2 Analysis of the effect pattern of influencing factors on CPD knowledge diffusion

(1) The effect of trust level between knowledge subjects on CPD knowledge diffusion

In the CPDS, assume that the intra-team trust level between knowledge subjects $\rho_{In} = 0.9$, the inter-team trust level between knowledge subjects $\rho_{out} = 0.6$, the intra-team knowledge stickiness $v_{In} = 3.5$, the inter-team knowledge stickiness $v_{out} = 4.5$, and the knowledge forgetting rate $\delta = 0.1$. To disclose the effect of trust level on CPD knowledge diffusion, the simulation is conducted with the initial trust levels of 60 %, 70 %, 80 %, 90 % and 100 %. The simulation results are exhibited in Fig. 2.

Fig. 2 The effect of trust level between knowledge subjects on CPD knowledge diffusion

Under the same knowledge stickiness, inter-team collaboration strength and knowledge forgetting rate, the average knowledge level of CPDS teams grows higher and faster with the increase of the trust level between the knowledge subjects in the CPDS, Fig. 2(a). This means the trust level between knowledge subjects is positively correlated with the average knowledge lev-

el and the diffusion rate of CPD knowledge diffusion. At the beginning of knowledge diffusion, higher trust level between knowledge subjects is accompanied by wider gaps on the knowledge levels of different teams; when it comes to the later stage of knowledge diffusion, however, the gaps narrow down rapidly and eventually converge to a low variance, Fig. 2(b). Thus, higher trust level helps CPDS teams to achieve more uniform distribution of knowledge. Summing up the simulation results of Figs. 2(a) and 2(b), it is concluded that: when the knowledge subjects maintain a high level of trust, the CPDS teams would possess high knowledge levels and achieve uniform knowledge distribution at a fast knowledge diffusion rate. In practice, the CPDS managers should recognize that the effective knowledge diffusion is the outcome of trust, cooperation and other emotional factors, and the foundation of knowledge interaction and collaboration is trust. Therefore, the managers should take organizational, institutional and cultural measures to boost the trust between CPDS members on knowledge sharing, encourage them more willing to cooperate and share knowledge with each other, promote them to build closer ties, and arouse their interests in actively exchanging and sharing knowledge.

(2) The effect of knowledge stickiness on CPD knowledge diffusion

In the CPDS, assume that the intra-team knowledge stickiness $v_{In} = 3.5$, the inter-team knowledge stickiness $v_{out} = 4.5$, the intra-team trust level between knowledge subjects $\rho_{In} = 0.9$, the inter-team trust level between knowledge subjects $\rho_{out} = 0.6$, and the knowledge forgetting rate $\delta = 0.1$. To study the effect of knowledge stickiness on CPD knowledge diffusion, the simulation is conducted with the initial trust levels of 60 %, 70 %, 80 %, 90 % and 100 %. The simulation results are displayed in Fig. 3.
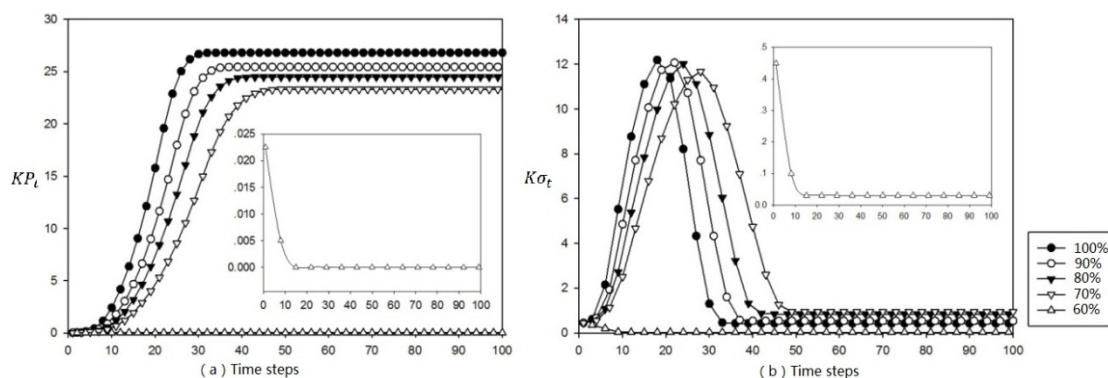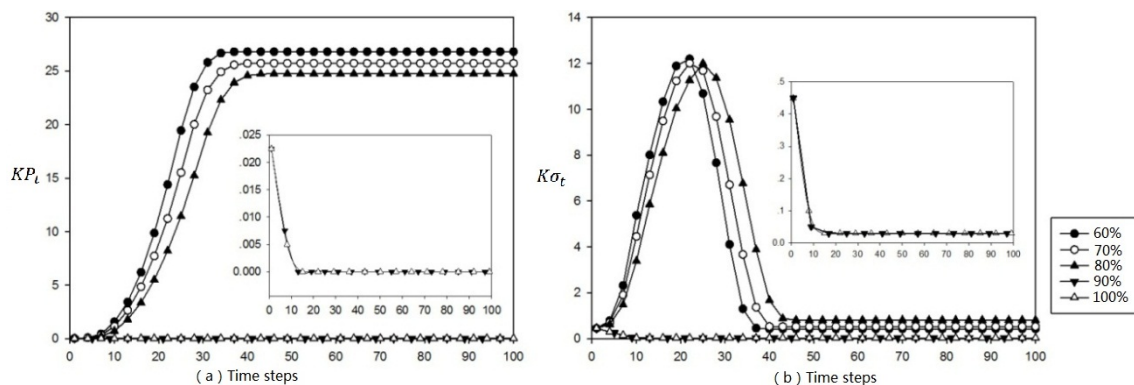

Fig. 3 The effect of knowledge stickiness on CPD knowledge diffusion

Under the same trust level, inter-team collaboration strength and knowledge forgetting rate, the average knowledge level of CPDS teams grows lower and slower with the increase of CPD knowledge stickiness, Fig. 3(a), that is, the knowledge stickiness is negatively correlated with the average knowledge level and the diffusion rate of CPD knowledge diffusion. The result bears testimony to the fact that: due to the objective nature of knowledge, stickier CPD knowledge is often more complex, ambiguous and specific, and harder to encode or transfer. The fact adds to the difficulty in knowledge diffusion. In the most severe cases, it would be too difficult for the CPD knowledge diffusion to occur. As shown in Fig. 3(b), the beginning of knowledge diffusion is marked by low CPD knowledge stickiness and wide gaps on the knowledge level between different teams. Upon entering the later stage of knowledge diffusion, however, the gaps shrink rapidly and finally converge to a low variance, Fig. 3(b), indicating that lower knowledge stickiness contributes to more uniform distribution of knowledge across the CPDS teams. Overall, the simulation results in Fig. 3 show the following practical implications: the less sticky the CPD knowledge is, the higher knowledge levels and faster knowledge diffusion rate the CPDS teams would possess. In real industrial environment, especially these high-technology industries with complex and quick-update knowledge, the CPDS should confer the members more permissions to contact and use its knowledge resource. On the other hand, the complex CPD knowledge should be described in simple, clear and versatile languages which can be effectively learned and

used by members, and the CPDS members should be trained to improve their comprehension and acceptance ability for the CPD knowledge. In these ways, managers can reduce the stickiness of knowledge transfer and raise the efficiency of knowledge diffusion.

(3) The effect of inter-team collaboration strength on CPD knowledge diffusion

According to the formula (6)-(7), the inter-team collaboration strength is a fixed value determined by "cell distance". To reveal the effect of inter-team collaboration strength on CPD knowledge diffusion, the simulation is conducted with the initial trust levels of 60 %, 70 %, 80 %, 90 % and 100 %. Assume that the trust level between the CPD knowledge subjects within a CPDS team $\rho_{In} = 0.9$, the inter-team trust level between knowledge subjects $\rho_{out} = 0.6$, the intra-team knowledge stickiness $v_{In} = 3.5$, the inter-team knowledge stickiness $v_{out} = 4.5$, and the knowledge forgetting rate $\delta = 0.1$. The simulation results are presented in Fig. 4.
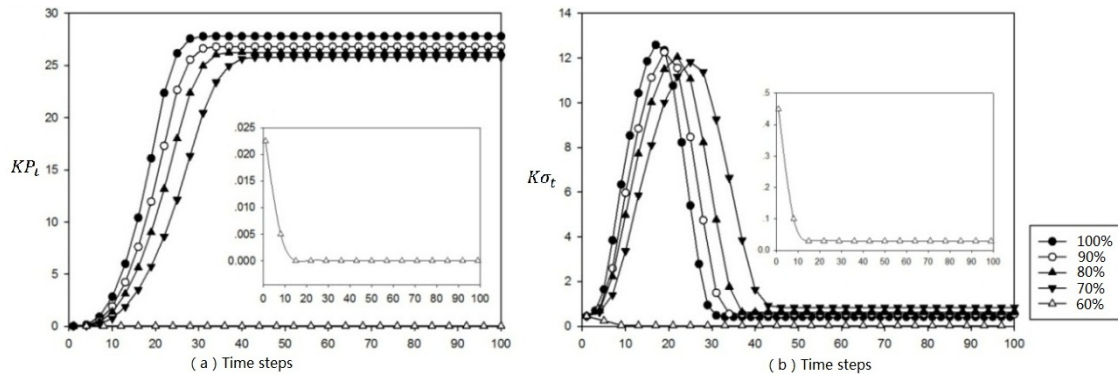
Fig. 4 The effect of inter-team collaboration strength on CPD knowledge diffusion

Fig. 4 shows the positive correlation between inter-team collaboration strength and the average knowledge level, diffusion rate and knowledge distribution uniformity in CPD knowledge diffusion. The revelation provides an evidence to the fact that: good inter-team collaboration strength improves the stability of CPDS so that the teams cooperate more closely in CPD tasks, and also offers knowledge subjects belonging to different teams more channels and chances for exchanges, thereby advancing the diffusion of knowledge across the collaborative teams. Thus, the practical conclusion can be drawn that: stronger inter-team collaboration enables CPDS teams to possess higher knowledge levels and achieve more uniform knowledge distribution at a faster knowledge diffusion rate. In the management of CPDS, the inter-team collaboration strength determines the effectiveness of knowledge exchange and collaboration among different teams, and then further affects the knowledge diffusion efficiency of the whole CPDS. Therefore, mangers should strengthen the flat management of CPDS to make each team the centrality of CPDS by reducing organizational hierarchy and intensifying intra-team exchanges and collaborations in a relatively stable operating environment. On ther other hand, managers also need to encourage members of different teams to actively communicate and collaborate with each other in formal and informal ways, in order to enhance the efficiency of knowledge exchange among members. Finnaly, the knowledge of individual team members can be elevated into team resources, and then turned into the knowledge resources of the entire organization, making knowledge a competitive edge over other groups and organizations.

(4) The effect of knowledge forgetting on CPD knowledge diffusion

Assume that the trust level between the CPD knowledge subjects within a CPDS team $\rho_{In} = 0.9$, the inter-team trust level between knowledge subjects $\rho_{out} = 0.6$, the intra-team knowledge stickiness $v_{In} = 3.5$, and the inter-team knowledge stickiness $v_{out} = 4.5$. To discover the effect of knowledge forgetting on knowledge diffusion, the simulation is conducted with $\delta = 0, 0.1, 0.2, 0.3, 0.4$. The simulation results are listed in Fig. 5.
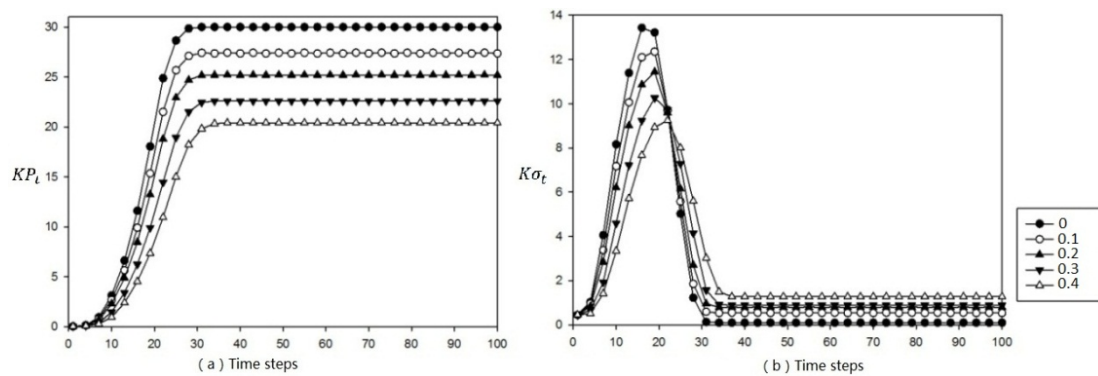
Fig. 5 The effect of knowledge forgetting on CPD knowledge diffusion

Fig. 5 showcases the negative correlation between knowledge forgetting and the average knowledge level, diffusion rate and knowledge distribution uniformity in CPD knowledge diffusion. It reflects the reality that: knowledge forgetting reduces the attachment of CPD knowledge to knowledge subjects, making it more difficult for knowledge subjects to possess CPD knowledge in the long run; meanwhile, the knowledge diffusion process is discretized due to the knowledge loss during knowledge forgetting, which hinders the effective diffusion of CPDS knowledge. Based on the simulation results in Fig. 5, it can be drawn the following practical conclusions: low knowledge forgetting rate allows CPDS teams to possess high knowledge levels and achieve uniform knowledge distribution at a fast knowledge diffusion rate. Therefore, managers should strictly control the introduced knowledge to avoid the waste or wrong guidance by introducing valueless knowledge. For the introduced knowledge with great value, managers should effectively promote the knowledge to improve members' cognition and recognition of the knowledge value. Secondly, the repeated backups and personnel trainings are necessary in the diffusion of the valuable CPD knowledge to lower the level of difficulty of the introduced knowledge. Moreover, the members possessing the knowledge should be regarded as key management object to give them more change to strengthen the memory through constant application of the knowledge, so as to minimize the loss of CPD knowledge caused by knowledge forgetting.

5. Conclusion

Effective knowledge diffusion plays a vital role for the success of CPD. Aiming to deeply understand the diffusion process and rule of CPD knowledge, this paper constructs the K-SIS model of CPD knowledge diffusion in the context of CPD based on an improved CA method. Compared with the traditional knowledge diffusion model, the proposed model makes the following innovations: (1) Based on the improved CA method, the model involves a "bottom up" examination of the CPD knowledge diffusion process from the perspective of microscopic knowledge exchange activities; (2) The traditional CA model is improved to fit the quantitative research of CPD knowledge diffusion in the context of collaboration. During the simulation on the proposed model, the author digs deep into the diffusion process and pattern of CPD knowledge, using the average knowledge level of the teams and the standard deviation of the knowledge level to measure the performance of CPD knowledge diffusion; Moreover, the author performs multiple simulations with the influencing factors of CPD knowledge diffusion given different parameter values, and compares and analyzes the simulation results to obtain the effect pattern of each influencing factor on CPD knowledge diffusion. The simulation results show that the proposed model can realize the quantitative analysis of the CPD knowledge diffusion process and the prediction of the diffusion trend. Moreover, different factors play different roles for the knowledge diffusion efficiency. Monitoring and managing these factors will help enterprises to optimizing the CPD knowledge diffusion process. This study offers theoretical guidance and methodological support for improving CPD knowledge diffusion efficiency.

This research provides a feasible reference for the study of the CPD knowledge diffusion mechanism and the prediction of CPD knowledge diffusion trend from the knowledge activities of micro individuals. Nevertheless, this paper fails to discuss the individual differences and the diffusion time delay in CPD knowledge diffusion. These issues are meaningful objects in further research.

Acknowledgement

References

[1] Yang, Y., Guo, B., Yin, S., Wang, W.-L., Zhang, X.-D. (2008). Connotation, theory framework and application of customer collaborative innovation, *Computer Integrated Manufacturing Systems*, Vol. 14, No 5, 944-950.

[2] Xing, L. (2016). Research on cooperation innovation among enterprises in the strategic emerging industrial cluster, *Advances in Modelling and Analysis A*, Vol. 53, No. 2, 93-107.

[3] Reagans, R., McEvily, B. (2003). Network structure and knowledge transfer: The effects of cohesion and range, *Administrative Science Quarterly*, Vol. 48, No. 2, 240-267, doi: 10.2307/3556658.

[4] Kim, H., Park, Y. (2009). Structural effects of R&D collaboration network on knowledge diffusion performance, *Expert Systems with Applications*, Vol. 36, No. 5, 8986-8992, doi: 10.1016/j.eswa.2008.11.039.

[5] Li, L., Dang, X., Zhang, S. (2007). Study on the knowledge diffusion of enterprises technology innovation cooperation based on the perspective of knowledge potential, *Science of Science and Management of S. & T.*, Vol. 28, No. 4, 107-112, doi: 10.3969/j.issn.1002-0241.2007.04.024.

[5] Bass, F.M. (1969). A new product growth for model consumer durables, *Management Science*, Vol. 15, No. 5, 215-227, doi: 10.1287/mnsc.15.5.215.

[7] Sun, Y.-W., Wei, Y.-P. (2011). EmpiricaI study on AIDA model of knowledge diffusion in high-tech enterprise alliance based on complex network, *China Soft Science*, Vol. 6, 17-26, doi: 10.3969/j.issn.1002-9753.2011.06.013.

[8] Meng, X., Liu, H., Wu, H. (2004). Multi-agent model of knowledge diffusion in network space, *Studies in Science of Science*, Vol. 21, No. 6, 636-641, https://doi.org/10.3969/j.issn.1003-2053.2003.06.014.

[9] Xue, J., Ding, C.Q., Yuan, J. (2016). The evolutionary game study of knowledge transfer behavior in cooperative crowdsourcing community of innovation, *Advances in Modelling and Analysis A*, Vol. 53, No. 2, 164-175.

[10] Yang, Y., Wang, X.-L., Zeng, Q., Yang, J., Xing, Q.-S. (2011). Multi-agent conflict coordination in collaborative product creative design optimization, *Computer Integrated Manufacturing Systems*, Vol. 17, No. 1, 1-9, doi: 10.13196/j.cims.2011.01.3.yangy.001.

[11] Yang, J., Yang, Y., Zhao, C., Liang, X.-D., Zeng, Q. (2009). Integrated technology of customer knowledge in collaborative product innovation based on ontology theory, *Computer Integrated Manufacturing Systems*, Vol. 15, No. 12, 1-9, doi: 10.13196/j.cims.2009.12.17.yangj.024.

[12] Wolfram, S. (1984). Cellular automata as models of complexity, *Nature*, Vol. 311, No. 5985, 419-424, doi: 10.1038/311419a0.

[13] Guseo, R., Guidolin, M. (2009). Modelling a dynamic market potential: A class of automata networks for diffusion of innovations, *Technological Forecasting and Social Change*, Vol. 76, No. 6, 806-820, doi: 10.1016/j.techfore.2008.10.005.

[14] Gu, B., Sheng, V.S. (2016). A robust regularization path algorithm for ν-support vector classification, *IEEE Transactions on Neural Networks and Learning Systems*, Vol. 28, No. 5, 1241-1248, doi: 10.1109/TNNLS.2016.2527796.

[15] Sun, G.-Q., Jin, Z., Li, L. (2011). Emergent turing pattern in epidemic spreading using cellular automaton, *International Journal of Modern Physics B*, Vol. 25, No. 32, 4605-4613, doi: 10.1142/S0217979211059401.

[16] Xia, Z., Wang, X., Sun, X., Liu, Q., Xiong, N. (2016). Steganalysis of LSB matching using differences between nonadjacent pixels, *Multimedia Tools and Applications*, Vol. 75, No. 4, 1947-1962, doi: 10.1007/s11042-014-2381-8.

[17] Luo, S., Du, Y., Liu, P., Xuan, Z., Wang, Y. (2015). A study on coevolutionary dynamics of knowledge diffusion and social network structure, *Expert Systems with Applications*, Vol. 42, No. 7, 3619-3633 doi: 10.1016/j.eswa.2014.12.038.

An integer programming approach for process planning for mixed-model parts manufacturing on a CNC machining center

Kongchuenjai, J.[a,*], Prombanpong, S.[a]

[a]King Mongkut's University of Technology Thonburi (KMUTT), Department of Production Engineering, Bangkok, Thailand

ABSTRACT

The manufacturing system for mixed parts is prevalent in many industries due to the continuous demand of product variety. Thus, the mathematical model to design a process plan is developed by using the integer linear programming. The main aim is to minimize the total production time. The main time factors included in the model composed of the machining time, tool traveling time, tool changing time and tombstone face changing time. The significant design constraints are the precedence operations, fixture design, and available cutting tool constraints. Furthermore, a variation of part styles is also accounted for in this study as the different types of a part can be concurrently mounted and processed which makes this problem unique. Therefore, this problem is much more complex than the normal single model process planning. The model developed using integer programming will determine a sequence of required machining operations. It also decides the face of the machining part to be fastened on the tombstone face. In addition, a suitable cutting tool will be selected based on minimum total production time. The result of this paper, aids in solving process planning difficulty in the dedicated flexible manufacturing system in the era of Industry 4.0.

1. Introduction

A process planning task is a part of product design and manufacturing system that considers the manufacturing resources and engineering techniques for decision making. The efficient process plan can be achieved by well-designed logical operation steps, technological machine capabilities, cutting tool selection, and machining conditions. Moreover, the process planner has to be concerned about the jigs and fixtures, route sheet generation, setup requirements and other important processes during the design process [1]. Thus, the process plan reveals the sequence of machining operations together with a selection of cutting tools as well as cutting conditions. A process planner must make a crucial decision whether to continue using the same cutting tool or to change to a new cutting tool for the next set of remaining operations. This decision dictates a resulted machining sequence and challenges researchers to innovate efficient algorithms for a generation of the process plan which is considered a classic problem for automated process planning. It is obvious that the process planning task is crucial and affects both of the economic and efficiency of the manufacturing system.

In decades, the CNC machine technology has been developed by equipment to make it more highly automated, flexible, and run at a faster speed. The ATC (Automated Tool Change) and APC

(Automated Pallet Change), as an example, augment the capability of the CNC machine and become the standard equipment equipped in the available CNC machine in the market. The ATC enables the quick tool change from the tool magazine at no time loss. In addition, a modular fixture for the CNC machine is developed to allow a quick change of fixture design to cope with the FMS (Flexible Manufacturing System) demand. The automated process plan plays a crucial role in FMS since it enables the system to work effectively and meet its design objective. The FMS is deemed an intelligent manufacturing system (IMS) which corporates computer networks, automation, and intelligent software.

It can be enhanced by wireless communication, sensor networks, Internet of thing (IoT) to become a high-resolution manufacturing system (HMS) [2, 3]. Therefore, the FMS is considered an indispensable physical platform for the fourth industrial revolution so-called Industry 4.0. The FMS enables mixed-model parts to be concurrently performed in the same environment. It requires an automated part identification and a quick change of both physical setup and part programming. To enable the quick physical setup change, a tower-type pallet or tombstone type, together with APC, are invented to fasten multi-part faces and to change a pallet in the CNC machine. The tower-type tombstone allows up to four different part faces to be machined in only one set up. The part fastened at each tombstone face is not necessarily from the same work-part. It can be from a totally different part which shares the same set of available tools. Logically, a process planner must make an additional effort to decide which part face has to be fastened at each tombstone face which certainly relates and affects a tool selection. Meaning that a planner can continue using the same tool for more operations at different locations by rotating the tombstone. The APC allows a quick pallet change in order to eliminate setup time due to part change. In other words, the machining operations can be performed at one tombstone whereas unloading/loading of the new parts can be concurrently operated at the other tombstone. As a result, the process planner's task is enlarged by the advent of the new mentioned technology. This will increase the complexity of a process planning task since more machining operations and more part-faces options needed to be considered. Therefore, a sound process plan must integrate the machining operation requirement and the work-part sequence and orientation in the consideration, in order to suit today's CNC machine. In other words, the intelligent process plan must be able to reveal a sequence of machining operations with the selected tool and also identify the tombstone rotation sequence with the selected work-part face. Obviously, this task cannot be accomplished without the intelligent mathematical program.

This research deals with the flexible job shop scheduling problem, which is the NP-hard problem [2]. This research also proposed the Computer-Aided Process Planning (CAPP) by using the integer linear programming.

2. Literature review

The research in an automated process planning area has been long conducted for a few decades. A myriad of algorithms and methods both in heuristic and optimization are presented in the previous literature [2]. However, one of the most classic methods used to generate a process plan is an optimization through a linear programming method and can be applied in all manufacturing systems. As referred to in [3], the mixed integer programming was a mathematical technique used for flow shop scheduling and job shop scheduling, including the fuzzy processing time under precedence constraints and machine capacity constraints. The mixed integer programming also was used in the work of [4] for integrated process planning and scheduling model (IPPS) in a job shop problem. The linear integer programming model was also used to select the machine-tool and allocate the operations for FMS, in which each part and tool can be dynamically changed during the production [5]. In addition, the mixed integer linear programming is used to minimize the maximum completion time of final products in a flexible assembly job shop with the sequential-dependent setup time [6]. The mathematical programming in the form of linear programming is always utilized in the flexible job shop scheduling problem (FJSP) [7-9], flow shop scheduling [3, 10, 11] and integrated process planning and scheduling (IPPS) systems [4, 6, 12]. Different conditions were found in the objective functions, for example, minimization

of the make-span (the completion time of the last operation of all jobs), the lateness, the total processing cost and the weighted number of tardy jobs. In [8-9], the integrated constraint programming (CP) model was developed for the FMS, including the tool allocation, machine loading, part routing, and scheduling. The branch-and-bound algorithm was also used by the process planning and sequencing of CAPP and flexible manufacturing systems was studied [8]. Tool changing and tool sequencing problems were proposed where the branch-and-bound was used to minimize the total flow time [11, 12]. Various algorithms are proposed in the previous literature. For example, the game theory was proposed in [12] to investigate the interaction among decision alternatives, whereas the game theory was developed for multi-objective integrated process planning and scheduling (IPPS). The objective function is composed of the minimization of the maximal completion time of machines (makespan), the minimization of the maximal machine workload and the minimization of the total workload of machines. The integrated process planning was presented as a mathematical model to consider optimum machining conditions, operations sequences and tool magazine arrangement by CNC machine. The objective was to minimize the production costs [13]. As mentioned in [14], the process-planning model for machining cylinder heads in a typical engine manufacturing plant was designed in the form of mathematical programming models that include the part grouping, machine loading, tool allocation, and scheduling. The minimization of the non-productive tool movement and orientation time was proposed in this research. The problem of production scheduling, which consists of loading and sequencing, is tackled in this study [15]. The major findings of this study are generating an optimum sequencing order to complete all the jobs which have been assigned already on each machine. The solution for maximizing machine and labor utilization and reducing the idle time on each machine is presented. The two NP-hard optimization problems are composed of the flexible job-shop scheduling problems (FJSPs) and the FJSPs with process plan flexibility (FJSP-PPFs) [16]. The FJSPs is the routing and sequencing sub-problems, whereas the FJSP-PPFs, with process plan flexibility, is the process plan selection sub-problem. There are two steps for probe solving, which are 1) a mixed-integer linear programming model (MILP-1) is developed for FJSPs and 2) a modification of mixed-integer linear programming model is developed for FJSP-PPFs.

Table 1 Summary of mathematical modeling approach for process planning of machining processes

Author/ Year	Mathematical modeling approach				Process planning activities							
	Mathematical programming	Constraint programming	Branch and bound	Game theory	Sequencing of operation	Set-up planning	Selection of plan	Selection of part	Selection of machine	Selection of cutting tool	Cal. process times	Cal. cost
Li et al., 2010	*											*
Demir and Isleyen, 2013	*					*			*		*	
Milos Seda, 2007	*					*			*			
Li et al., 2010	*								*		*	*
Jahromi and Moghaddam, 2012	*							*	*	*		*
Nourali et al., 2012	*					*	*		*		*	
Deja et al., 2013			*		*				*			
Zeballos et al., 2010		*						*	*	*		
Zeballos, 2010		*						*	*	*	*	
Karakayal and Azizoglu, 2006			*			*					*	
Chandra et al., 2007			*			*					*	
Li et al., 2012				*			*				*	
Akturk and Avci, 1996	*					*				*		*
Das et al., 2009	*									*	*	
Ahmad et al., 2009	*					*			*		*	
Ozguven et al., 2010	*					*			*			*
Prombanpong et al.,1992	*					*						
Kongchuenjai and Prombanpong, 2015	*				*						*	

Computerized fixture design system and process plans generated by computer, aided the process planning. A mathematical model was developed to integrate process planning in fixture design considerations [17]. The objective of this research was to review the CAPP research works including the critical analysis of journals that publish CAPP research works and also the direction of future research work. The general information, past reviews, and discussions are summarized in various aspects [18]. In the work of [19], the mathematical programming to determine the optimal sequence of single part manufactured on a CNC machining center equipped with the single tombstone-type fixture was developed. The minimization of total production time required for machining, tool traveling and tool changing were then determined under relevant constraints. Table 1 concludes the relevant literature reviews of process planning of machining processes.

3. Materials and methods

3.1 Pallet configuration

As mentioned earlier the mixed-model parts in the FMS are to be machined by the CNC machining center with an automatic tool change and a pallet change function. With a pallet change function, it allows up to two pallets to be installed. Thus, the two tower-type pallets are equipped in the machine which allows up to eight-part faces to be ready for machining as shown in Fig. 1.
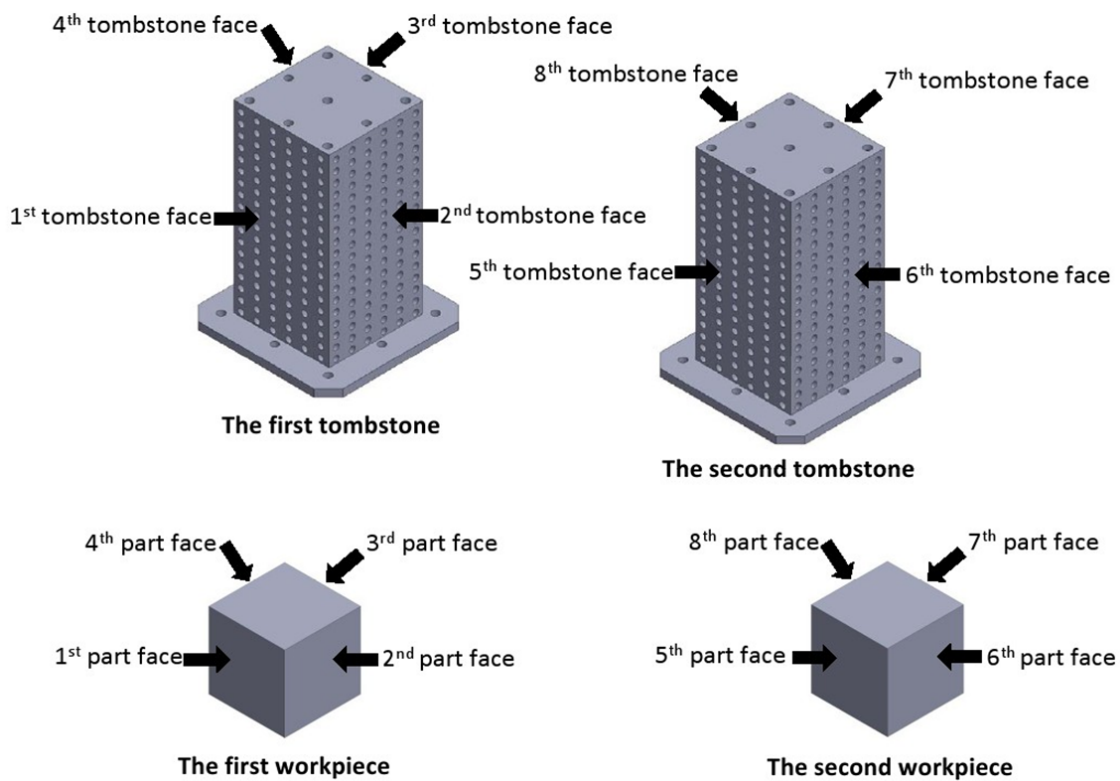
Fig. 1 Tower-type pallet

3.2 Mathematical model

Let assume two different parts are in the system and each four-part faces are required for machining operations. Thus, a process planner must integrate two planning tasks i.e. which part face to be fastened on each tombstone face and what is the sequence of machining operations and tools to complete all the required machining operations. It is obvious that the planning task is much more complicated; consequently, the mathematical model using 0-1 Integer programming is proposed and can be described as follows:

Indices:

i	Index of tombstone fixture faces ($i = 1, 2, \cdots, I$).
j	Index of part faces of the work-piece ($j = 1, 2, \cdots, J$).
k, k'	Index of machining operation number ($k, k' = 1, 2, \cdots, N$).
l	Index of tool number ($l = 1, 2, \cdots, L$).
m, m'	Index of consecutive number ($m, m' = 1, 2, \cdots, N$).

Parameters:

I	Maximum number of tombstone fixture faces
J	Maximum number of part faces
k_j	Set of operations to be performed in part face j
L	Number of all available cutting tools in the tool magazine
l_k	Set of tools that can perform for operation k; $l_k \subseteq L$
N	Maximum number of operations or sequences
MT_{kl}	Machining time for operation k by tool l
TC	Tool change time which is constant
$TT_{kk'}$	Tool travel time from operations k to operation k'
$TFC_{ii'}$	Tombstone face change time from the tombstone face i to i'

Decision variables:

E	Total production time.
X_{ijklm}	= 1 when operation k is performed on the tombstone face i on part face j with tool l in sequence m.
	= 0, otherwise
$Z_{kk'lm}$	= 1 when tool l is used for both operation k and k' in sequence m and $m + 1$, respectively.
	= 0, otherwise
$W_{ii'm}$	= 1 when sequence m is setup the tombstone face i and sequence $m + 1$ is setup the tombstone face i'.
	= 0, otherwise
Y_{ij}	= 1 when part face j occupies in tombstone face i.
	= 0, otherwise

The objective function is to minimize the total production time of the parts which is a summation of the four entities, namely, the machining times, the tool change time, the tool travel times and the pallet change time and can be formulated as (1).

$$\text{Minimize } E = \sum_{i=1}^{I}\sum_{j=1}^{J}\sum_{k=1}^{N}\sum_{l\in l_k}\sum_{m=1}^{N}\left(MT_{kl} \times X_{ijklm}\right) + TC \times \sum_{m=1}^{N-1}\left(1 - \sum_{k=1}^{N}\sum_{\substack{k'=1\\k'\neq k}}^{N}\sum_{l\in l_k} Z_{kk'lm}\right)$$

$$+ \sum_{m=1}^{N-1}\sum_{l\in l_k}\left[\sum_{k=1}^{N}\sum_{\substack{k'=1\\k'\neq k}}^{N}\left(TT_{kk'} \times Z_{kk'lm}\right)\right] + \sum_{m=1}^{N-1}\left[\sum_{i=1}^{I}\sum_{i'=1}^{I}\left(TFC_{ii'} \times W_{ii,m}\right)\right] \tag{1}$$

The set of constrains in the mathematical model are as follows:

1. Operation completion constraint. All the required machining operations must be completed and must not be repeated. Each required machining operation must be completed at the only one tombstone pallet face and the only one cutting tool.

$$\sum_{i=1}^{I}\sum_{j=1}^{J}\sum_{l\in l_k}\sum_{m=1}^{N} X_{ijklm} = 1 \text{ for all } k \tag{2}$$

2. Singular sequence constraint. This constrain ensures that the operation k must be scheduled in only one sequence slot in the sequence plan.

$$\sum_{i=1}^{I}\sum_{j=1}^{J}\sum_{k\in k_j}\sum_{l\in l_k} X_{ijklm} = 1 \text{ for all } m \tag{3}$$

3. Precedence constraint. An operation k' must follow the precedence constraint rule. That is if k must precede k' and k is sequenced in sequence m, k' must be able to perform only at sequence m' where $m' \geq m+1$. This constrain ensures a logical process sequence.

$$\sum_{i=1}^{I}\sum_{j=1}^{J}\sum_{k\in k_j}\sum_{l\in l_k} X_{ijklm} - \sum_{i=1}^{I}\sum_{j=1}^{J}\sum_{k'\in k'_j}\sum_{l\in l_k}\sum_{m'\geq m+1} X_{ijk'lm'} \geq 0 \tag{4}$$

4. Tool change time equation. Whenever there is any tool change, it will increase a production time in the objective function. Thus, this tool change time must be taken into account. In the mathematical expression, $Z_{kk'lm}$ is used as a flag to indicate the tool change action. That is, if the same tool l is used in operation k and k', then $Z_{kk'lm}$ is set to be 1; otherwise 0.

$$\sum_{i=1}^{I}\sum_{j=1}^{J}\sum_{\substack{k\in k_j \\ k\neq k'}} X_{ijklm} + \sum_{i=1}^{I}\sum_{j=1}^{J}\sum_{\substack{k\in k'_j \\ k\neq k'}} X_{ijk'lm'} - 2Z_{kk'lm} \geq 0 \tag{5}$$

$$\text{for all } k \neq k', m' = m+1, \ldots, N-1, l \in l_k$$

$$\sum_{i=1}^{I}\sum_{j=1}^{J}\sum_{\substack{k\in k_j \\ k\neq k'}} X_{ijklm} + \sum_{i=1}^{I}\sum_{j=1}^{J}\sum_{\substack{k\in k'_j \\ k\neq k'}} X_{ijk'lm'} - 2Z_{kk'lm} \leq 1 \tag{6}$$

$$\text{for all } k \neq k', m' = m+1, \ldots, N-1, l \in l_k$$

5. Tombstone face occupied constraint. This constrain ensures that there will be only one part face j occupied at each tombstone face i.

$$\sum_{j=1}^{J} Y_{ij} = 1 \text{ for all } i \tag{7}$$

$$\sum_{i=1}^{I} Y_{ij} = 1 \text{ for all } j \tag{8}$$

$$Y_{ij} - \sum_{k=1}^{N}\sum_{l\in l_k}\sum_{m=1}^{N} x_{ijklm} \geq 0 \text{ for all } i, j \tag{9}$$

6. Tombstone rotation or change time equation. Whenever there is any tombstone rotation by itself or completely alternate to the other one, the lost time of this action must be taken into account in the production time. The logic is that time required to rotate itself is less than that of changing to the other where *TFCii'*, tombstone face change time is varied depending upon face index relationship. In the mathematical expression, $W_{ii'm}$ is utilized as a flag to indicate the change action. If tombstone face i is scheduled in sequence m and the tombstone face i' is scheduled in sequence $m + 1$, then $W_{ii'm}$ is assigned to 1; otherwise 0.

$$\sum_{\substack{j=1 \\ }}^{J} \sum_{\substack{k \in k_j \\ k \neq k'}} X_{ijklm} + \sum_{\substack{j=1 \\ }}^{J} \sum_{\substack{k \in k_j \\ k \neq k'}} X_{i'jk'lm'} - 2W_{ii'm} \geq 0 \tag{10}$$

for all $i, i' = 1,2, \dots, I, k \neq k', m' = m + 1, \dots, N - 1, l \in l_k$

$$\sum_{\substack{j=1 \\ }}^{J} \sum_{\substack{k \in k_j \\ k \neq k'}} X_{ijklm} + \sum_{\substack{j=1 \\ }}^{J} \sum_{\substack{k \in k_j \\ k \neq k'}} X_{i'jk'lm'} - 2W_{ii'm} \leq 1 \tag{11}$$

for all $i, i' = 1,2, \dots, I, k \neq k', m' = m + 1, \dots, N - 1, l \in l_k$

7. Integrality constraints. All decision variables are 0-1 integer.

$$X_{ijklm} = 0,1 \text{ for all } i, j, k, l \text{ and } m \tag{12}$$

$$Z_{kk'lm} = 0,1 \text{ for all } k, k' \text{ and } m \tag{13}$$

$$W_{ii'm} = 0,1 \text{ for all } i, i' \text{ and } m \tag{14}$$

4. Case study: Automotive parts (manifold inlet and console)

4.1 Preparatory steps

The developed mathematical model is applied with actual two automotive parts. It is assumed that the manifold inlet and the console as shown in Fig. 2 and Fig. 3, respectively, are the mixed-model parts needed to be machined by the machining center in FMS.

A list of available cutting tools, equipped with the tool magazine, is shown in Table 2. It is assumed that there are nine cutting tools available that can be used to complete all the operations in the above-mentioned parts.

The manifold requires ten machining operations which are arbitrarily assigned as operation numbers 1 to 10. The console, on the other hand, also needs ten machining operations which will be assigned as operation numbers 11 to 20. Thus, there will be a total of twenty required machining operations to complete as shown in Table 3. Note that machining operation numbers 1 to 10 belong to the manifold and the remaining operations belong to the console. For each operation, there is a list of feasible tools as well as the calculated machining time associated with the tool. In addition, the required precedence operation is also provided. As an example, operation number 6 must precede operation number 7. Likewise, operation number 13 must precede operation number 14 and so on.
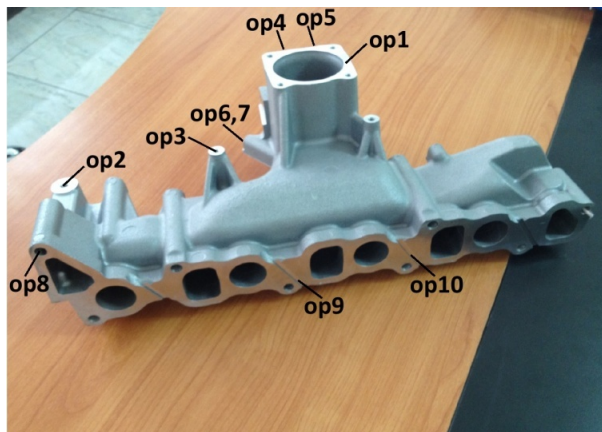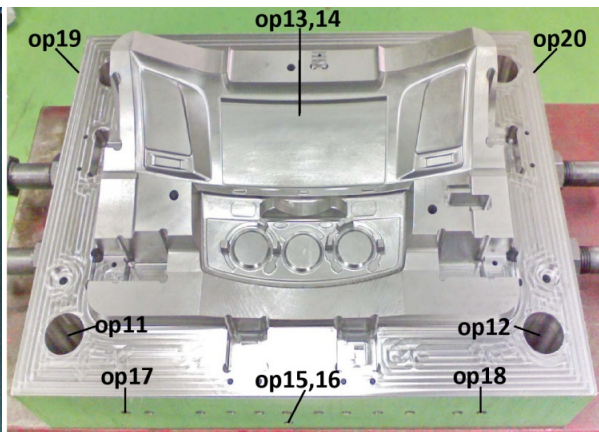

Fig. 2 Manifold

Fig. 3 Console

Table 2 A list of cutting tools in the tool magazine

Tool number	Tool specification
1	Face cutter 32 mm
2	Face cutter 80 mm
3	Spot face cutter 24 mm
4	Drill bit 6.8 mm
5	Drill bit 10 mm
6	Tap cutter M8
7	End mill 3 mm
8	End mill 6 mm
9	Ball mill 6 mm

Table 3 The necessary machining data

Work-piece	Part face	Operation no. and description		Tool no. and description		Machining time (min)	Precedence
Manifold	1	1	Face milling	1	Face cutter 32 mm	30	-
				2	Face cutter 80 mm	15	-
		2	Drilling	5	Drill bit 10 mm	1	-
		3	Drilling	5	Drill bit 10 mm	1	-
	2	4	Face milling	3	Spot Face cutter 24 mm	3	-
		5	Face milling	3	Spot Face cutter 24 mm	3	-
	3	6	Drilling	4	Drill bit 6.8 mm	1	-
		7	Tapping M8	6	Tap M8	2	Operation 6
	4	8	Boring	7	End mill 3 mm	8	-
				8	End mill 6 mm	5	
		9	Grooving	7	End mill 3 mm	7	-
		10	Grooving	7	End mill 3 mm	7	-
Console	5	11	Boring	7	End mill 3 mm	6	-
				8	End mill 6 mm	3	
		12	Boring	7	End mill 3 mm	6	-
				8	End mill 6 mm	3	
		13	Rough milling	7	End mill 3 mm	360	-
				8	End mill 6 mm	240	
		14	Finish milling	9	Ball mill 6 mm	480	Operation 13
		15	Drilling	4	Drill bit 6.8 mm	1	-
	6	16	Tapping M8	6	Tap M8	2	Operation 15
		17	Drilling	5	Drill bit 10 mm	1	-
		18	Drilling	5	Drill bit 10 mm	1	
	7	19	Grooving	7	End mill 3 mm	7	-
	8	20	Grooving	7	End mill 3 mm	7	-

Other essential data are tool travel time between the operations and tombstone rotation time and tombstone change time as shown in Table 4 and 5, respectively. The tool travel time is the time required to reach the next operation. It assumed that the travel speed used to calculate tool travel time is the maximum speed that the machine can provide. In this study case, the maximum travel speed (rapid traverse) is 1.969 in/min. In Table 5 the tombstone rotation time is around 0.6 min whereas the tombstone change time is around 2 min since this operation requires the other tombstone to swing into place.

All the necessary data in Tables 2 to 5 will be used to formulate the mathematical model and executed using Gurobi Optimizer. A total number of generated variables is 2.252 and the computer processing time is 31.18 min. The obtained result is summarized in Table 6 and 7.

Table 4 Tool travel time between the operations

From operation k	Tool travel time (min) To operation number k'																			
	1	2	3	4	5	6	7	8	9	10	11	12	13	14	15	16	17	18	19	20
1	-	0.2	0.1	0.1	0.1	0.1	0.1	0.2	0.1	0.1	0.4	0.4	0.2	0.2	0.2	0.1	0.3	0.3	0.4	0.4
2	0.2	-	0.2	0.2	0.3	0.2	0.2	0.1	0.3	0.3	0.3	0.4	0.3	0.3	0.3	0.3	0.1	0.4	0.2	0.4
3	0.1	0.2	-	0.1	0.2	0.1	0.1	0.2	0.2	0.2	0.3	0.4	0.2	0.2	0.2	0.2	0.4	0.4	0.4	0.4
4	0.1	0.2	0.1	-	0.2	0.1	0.1	0.2	0.2	0.2	0.3	0.4	0.2	0.2	0.2	0.2	0.2	0.3	0.3	0.3
5	0.1	0.3	0.2	0.2	-	0.2	0.2	0.3	0.1	0.1	0.3	0.3	0.2	0.2	0.2	0.2	0.3	0.3	0.3	0.3
6	0.1	0.2	0.1	0.1	0.2	-	-	0.1	0.1	0.1	0.3	0.3	0.2	0.2	0.1	0.2	0.3	0.3	0.3	0.3
7	0.1	0.2	0.1	0.1	0.2	-	-	0.1	0.1	0.1	0.3	0.3	0.2	0.2	0.2	0.1	0.3	0.3	0.3	0.3
8	0.2	0.1	0.2	0.2	0.3	0.1	0.1	-	0.3	0.3	0.2	0.4	0.1	0.1	0.2	0.2	0.1	0.3	0.1	0.4
9	0.1	0.3	0.2	0.2	0.1	0.1	0.1	0.3	-	0.2	0.2	0.2	0.2	0.2	0.2	0.2	0.2	0.2	0.2	0.4
10	0.1	0.3	0.2	0.2	0.1	0.1	0.1	0.3	0.2	-	0.4	0.1	0.2	0.2	0.2	0.2	0.3	0.1	0.4	0.2
11	0.4	0.3	0.3	0.3	0.3	0.3	0.3	0.2	0.2	0.4	-	0.4	0.4	0.4	0.2	0.2	0.2	0.3	0.1	0.4
12	0.4	0.4	0.4	0.4	0.3	0.3	0.3	0.4	0.2	0.1	0.4	-	0.4	0.4	0.2	0.2	0.4	0.1	0.4	0.1
13	0.2	0.3	0.2	0.2	0.2	0.2	0.2	0.1	0.2	0.2	0.4	0.4	-	-	0.1	0.1	0.2	0.2	0.2	0.3
14	0.2	0.3	0.2	0.2	0.2	0.2	0.2	0.1	0.2	0.2	0.4	0.4	-	-	0.1	0.1	0.2	0.2	0.2	0.3
15	0.2	0.3	0.2	0.2	0.2	0.1	0.2	0.2	0.2	0.2	0.2	0.2	0.1	0.1	-	-	0.2	0.2	0.2	0.2
16	0.1	0.3	0.2	0.2	0.2	0.2	0.1	0.2	0.2	0.2	0.2	0.2	0.1	0.1	-	-	0.2	0.2	0.2	0.2
17	0.3	0.1	0.4	0.2	0.3	0.3	0.3	0.1	0.2	0.3	0.2	0.4	0.2	0.2	0.2	0.2	-	0.4	0.1	0.4
18	0.3	0.4	0.4	0.3	0.3	0.3	0.3	0.3	0.2	0.1	0.3	0.1	0.2	0.2	0.2	0.2	0.4	-	0.4	0.1
19	0.4	0.2	0.4	0.3	0.3	0.3	0.3	0.1	0.2	0.4	0.1	0.4	0.2	0.2	0.2	0.2	0.1	0.4	-	0.4
20	0.4	0.4	0.4	0.3	0.3	0.3	0.3	0.4	0.4	0.2	0.4	0.1	0.3	0.3	0.2	0.2	0.4	0.1	0.4	-

Table 5 Tombstone rotation and change time

From tombstone face i	Tombstone rotation and change time (min) To tombstone face i'							
	1	2	3	4	5	6	7	8
1	-	0.6	0.6	0.6	2	2	2	2
2	0.6	-	0.6	0.6	2	2	2	2
3	0.6	0.6	-	0.6	2	2	2	2
4	0.6	0.6	0.6	-	2	2	2	2
5	2	2	2	2	-	0.6	0.6	0.6
6	2	2	2	2	0.6	-	0.6	0.6
7	2	2	2	2	0.6	0.6	-	0.6
8	2	2	2	2	0.6	0.6	0.6	-

4.2 Results and discussion

The obtained results in Tables 6 and 7 can be delineated as the following. In Table 6 the optimal solution indicates which part face of the part will be fastened on which face of the tombstone. It can be seen from Table 4 that both Manifold and Console are fastened at both tombstone towers and each part face is assigned to a particular tombstone face. Without the solution obtained from this paper, one will randomly fasten the part on a tombstone tower. In addition, it is common that one tombstone fixture for one type of work-part, not a mixed part. Therefore, the process planner will not be able to generate the most optimal process plan due to the work-part setup constraint. This paper therefore advances a further step in the process planning research field. According to Table 7, the resulted optimal operation sequence is 18-17-15-6-7-16-.....-5. In other words, the first operation sequence in the plan is operation number 18 which is a drilling operation with tool number 5. Note that the operation 18 is located at part face 6 which is fastened at tombstone face 2. The second sequence is operation number 17 using the same cutting tool as the first sequence. Next, the cutting tool will be changed to tool number 4 in order to perform another drilling operation number 15.

Table 6 Summary of part and fixture setup result

Tombstone face i	Part	Part face j	Operation k
1	Manifold	4	10,9,8
2	Console	6	18,17,15,16
3	Console	8	20
4	Manifold	3	6,7
5	Console	5	12,11,13,14
6	Manifold	1	3,2,1
7	Console	7	19
8	Manifold	2	4,5

Table 7 Summary of the result obtained from the optimization model

Sequence m	Operation no. k	Tool no. l	Machining time (min) MT_{kl}	Tool change time (min) TC	Tool travel time (min) $TT_{kk'}$	Tombstone face change time (min) $TFC_{ii'}$
1	18	5	1	-	0.4	-
2	17	5	1	0.5	-	-
3	15	4	1	-	0.1	0.6
4	6	4	1	0.5	-	-
5	7	6	2	-	0.1	0.6
6	16	6	2	0.5	-	0.6
7	20	7	7	-	0.2	0.6
8	10	7	7	-	0.2	-
9	9	7	7	0.5	-	-
10	8	8	5	-	0.4	2
11	12	8	3	-	0.4	-
12	11	8	3	-	0.4	-
13	13	8	240	0.5	-	-
14	14	9	480	0.5	-	0.6
15	19	7	7	0.5	-	0.6
16	3	5	1	-	0.2	-
17	2	5	1	0.5	-	-
18	1	2	15	0.5	-	0.6
19	4	3	3	-	0.2	-
20	5	3	3	-	-	-
			790	4.5	2.6	6.2
Total production time(min)			803.3			

After that, the tombstone pallet rotates from face 2 to face 4 in order to perform the drilling operation number 6. Next, the cutting tool will be changed to tool number 6 in order to operate operation number 7 and then rotate the tombstone to face 2 in order to perform machining operation number 16 by the same cutting tool number 6.

In brief, all the twenty required machining operations are completed using a total of nine cutting tools. There is a total of nine tool changes with six times tombstone rotation and one tombstone change. The total machining time is 790 min, and tool change time, tool travel time, and tombstone face change time are 4.5 min, 2.6 min, and 6.2 min, respectively. Therefore, the total production time is 803.3 min.

5. Conclusion

This research proposed the integer linear programming approach to determine an optimal solution for a mixed-model part manufactured on a machining center. This model is a part of automated process planning, which has long been researched by using various mathematical techniques. The mathematical model has been verified by a number of examples in real cases and it is deemed practical. The uniqueness of this developed model is the ability to generate the optimal process plan with a simultaneous consideration of tombstone change for mixed-model parts. The mathematical model that was developed in this research deals with the complexity and decision making of Tombstone problems. In addition, based on the literature review, none of the research has addressed this type of problem. This research applied the real-world case study

of automotive parts manufacturing. The two types of automotive parts under consideration compose of manifold inlet and console. The results provide the process planning that can simultaneously produce two different part types. This mixed-model part consumes the minimum total production time. For future research work, the heuristics methods and mathematical relaxation techniques should be applied to solve the problems that have more machining operations, whereas the large-scale problems cannot easily be solved by general linear programming.

Acknowledgement

References

[1] Li, X., Zhang, C., Gao, L., Li, W., Shao, X. (2010). An agent-based approach for integrated process planning and scheduling, *Expert Systems with Applications,* Vol. 37, No. 2, 1256-1264, doi: 10.1016/j.eswa.2009.06.014.

[2] Demir, Y., İşleyen, S.K. (2013). Evaluation of mathematical models for flexible job-shop scheduling problems, *Applied Mathematical Modeling*, Vol. 37, No. 3, 977-988, doi: 10.1016/j.apm.2012.03.020.

[3] Šeda, M. (2007). Mathematical models of flow shop and job shop scheduling problems, *International Journal of Applied Mathematics and Computer Science,* Vol. 4, No. 4, 241-246.

[4] Li, X., Gao, L., Shao, X., Zhang, C., Wang, C. (2010). Mathematical modeling and evolutionary algorithm-based approach for integrated process planning and scheduling, *Computers & Operations Research*, Vol. 37, No. 4, 656-667, doi: 10.1016/j.cor.2009.06.008.

[5] Jahromi, M.H.M.A., Tavakkoli-Moghaddam, R. (2012). A novel 0-1 linear integer programming model for dynamic machine-tool selection and operation allocation in a flexible manufacturing system, *Journal of Manufacturing Systems*, Vol. 31, No. 2, 224-231, doi: 10.1016/j.jmsy.2011.07.008.

[6] Nourali, S., Imanipour, N., Shahriari, M.R. (2012). A mathematical model for integrated process planning and scheduling in the flexible assembly job shop environment with sequence dependent setup times, *International Journal of Mathematical Analysis*, Vol. 6, No. 43, 2117-2132.

[7] Deja, M., Siemiatkowski, M.S. (2013). Feature-based generation of machining process plans for optimized parts manufacture, *Journal of Intelligent Manufacturing*, Vol. 24, No. 4, 831-846, doi: 10.1007/s10845-012-0633-x.

[8] Zeballos, L.J., Quiroga, O.D., Henning, G.P. (2010). A constraint programming model for the scheduling of flexible manufacturing systems with machine and tool limitations, *Engineering Applications of Artificial Intelligence*, Vol. 23, No. 2, 229-248, doi: 10.1016/j.engappai.2009.07.002.

[9] Zeballos, L.J. (2010). A constraint programming approach to tool allocation and production scheduling in flexible manufacturing systems, *Robotics and Computer-Integrated Manufacturing,* Vol. 26, No. 6, 725-743, doi: 10.1016/j.rcim. 2010.04.005.

[10] Karakayalı, I., Azizoğlu, M. (2006). Minimizing total flow time on a single flexible machine, *International Journal of Flexible Manufacturing Systems*, Vol. 18, No. 1, 55-73, doi: 10.1007/s10696-006-9000-6.

[11] Chandra, P., Li, S., Stan, M. (1993). Jobs and tool sequencing in an automated manufacturing environment, *International Journal of Production Research,* Vol. 31, No. 12, 2911-2925, doi: 10.1080/00207549308956907.

[12] Li, X., Gao, L., Li, W. (2012). Application of game theory based hybrid algorithm for multi-objective integrated process planning and scheduling, *Expert Systems with Applications*, Vol. 39, No. 1, 288-297, doi: 10.1016/j.eswa. 2011. 07.019.

[13] Akturk, M.S., Avci, S. (1996). An integrated process planning approach for CNC machine tools, *The International Journal of Advanced Manufacturing Technology,* Vol. 12, No. 3, 221-229, doi: 10.1007/bf01351201.

[14] Das, K., Baki, M.F., Li, X. (2009). Optimization of operation and changeover time for production planning and scheduling in a flexible manufacturing system, *Computers & Industrial Engineering*, Vol. 56, No. 1, 283-293, doi: 10.1016/j.cie.2008. 06.001.

[15] Ahmad, I., Al-aney, K.I.M. (2009). Loading and sequencing jobs with the fastest machine among others, *Advances in Production Engineering & Management*, Vol. 4, No. 3, 127-138.

[16] Özgüven, C., Özbakır, L., Yavuz, Y. (2010). Mathematical models for job-shop scheduling problems with routing and process plan flexibility, *Applied Mathematical Modelling*, Vol. 34, No. 6, 1539-1548, doi: 10.1016/j.apm.2009. 09.002.

[17] Prombanpong, S., Lewis, R.L., Bishop, A.B. (1992). A computer-aided fixture design system with process planning integration for prismatic parts manufactured on CNC machining centers, In: *Computers in Engineering, Proceedings of the International Computers in Engineering Conference and Exhibition, San Francisco*, USA, 369-380.

[18] Xu, X., Wang, L., Newman, S.T. (2011). Computer-aided process planning – A critical review of recent developments and future trends, *International Journal of Computer Integrated Manufacturing*, Vol. 24, No. 1, 1-31, doi: 10.1080/0951192X.2010.518632.

[19] Kongchuenjai, J., Prombanpong, S. (2015). An optimization of automated process planning for manufacturing prismatic parts on a machining center, *Applied Mechanics and Materials,* Vol. 789-790, 1270-1274, doi: 10.4028/www.scientific.net/AMM.789-790.1270.

Predicting the availability of production lines by combining simulation and surrogate model

Jia, Y.[a], **Tian, H.**[a], **Chen, C.**[a,*], **Wang, L.**[a,b]

[a]School of Mechanical Science and Engineering, Jilin University, Changchun, P.R. China
[b]School of Mechanical Engineering, Dalian University of Technology, Dalian, P.R. China

A B S T R A C T

The availability analysis plays a significant part in both the design and operations management of production lines. In this paper, a method combining discrete event simulation (DES) and surrogate model is presented to predict the availability of production lines with unreliable workstations and finite intermediate buffers. The DES can conduct computer experiments for production lines with the help of design of experiments (DOE) under the Matlab environment. The surrogate model is constructed by using Kriging model integrated with Latin hypercube sampling (LHS), which can predict the responses based on a limited set of simulation results. The major advantages of the proposed approach are its flexibility and convenience. Also, it is the first time to investigate Kriging opportunities in predicting the performance of production lines. Finally, an application in a crankshaft production line is presented, and the results indicate that the proposed approach can achieve higher prediction accuracy than the other methods.

1. Introduction

A production line, also known as a transfer line or a flow line, is one of the most important and common types of manufacturing systems employed for high-volume low-variety production of industrial components. Unlike flexible manufacturing systems (FMS) or manufacturing cells (FMC), production lines play a significant role in processing the main products of a plant, and usually require high capital investment. They are often organized with a predetermined sequence of equipment [1] and intermediate buffers arranged in a serial structure and connected by a material handling system. Because of high sensibility to failures, the improvement of production lines' availability is an important issue for the designers or operators to resolve. It is therefore necessary to research the methods for evaluating or predicting the availability of production lines.

A review focusing on availability analysis techniques of production lines is given as follows, which can be divided into three groups: exact analytical methods, approximate analytical methods and simulation. Exact analytical methods, which are generally on the basis of queueing models and Markov chain [2], can obtain the exact solutions of steady-state probability, thus providing insight into the qualitative performance of production lines. But they are only suitable for small lines (no more than three-stage) because of the state explosion problem.

Based on the two-stage exact models, approximate analytical methods are developed for lines with more machines. The most representative methods are decomposition and aggregation ap-

proaches. The common intention of decomposition methods is to decompose an N-machine system into a cluster of N-1 subsystems which contain two pseudo-machines and one original buffer, and to get the result by solving simultaneous equations [3]. Gershwin [4] developed an efficient decomposition method for synchronous lines and it can get accurate results with the help of the Dallery-David-Xie (DDX) algorithm [5]. Afterwards, Burman [6] modified the continuous model for asynchronous lines and presented a new accelerate DDX (ADDX) algorithm. Other extensions of decomposition method can be found in [7-9]. Compared with decomposition methods, aggregation methods are more straightforward and simpler. The general idea is to aggregate the original line into one unique equivalent machine by iteratively replacing a two-machine one-buffer subline [10]. Meerkov and his group have obtained some achievements on aggregation methods. Detailed descriptions are summarized in [11], and a corresponding software called PSE Toolbox is developed. Although approximate analytical methods have been fully investigated, there is a limitation for wide application: all the distributions have to be limited to special forms, such as geometric or exponential.

Compared with analytical approaches, simulation can build the models of production lines at any requested level of detail [12] without being restricted by assumptions such as specified distributions. Consequently, discrete event simulation (DES) has been proved to be the ideal tool for exhibiting the dynamics of complex manufacturing process, and meanwhile, the accuracy can be controlled. Because of the applicability and practicability, DES has been widely used for predicting and optimizing the performance of production lines [13-15]. Furthermore, some authors [16-18] investigated the real production lines by means of case studies. A comprehensive discussion of many important aspects of discrete event simulation is given by Law [19] from fundamentals to applications. In addition, many commercial software packages (e.g., Flexsim, Witness, Plant Simulation, Arena, etc.) have been designed specifically to simulate manufacturing systems, thus increasing the popularity of simulation in recent years. Despite the modelling flexibility and great ease of use, DES is usually time-consuming, particularly at the initial design stage when lots of system parameters are indeterminate. Although high-performance computers are developed, a lot of computing time and resources are still necessary to obtain statistically significant results.

To overcome the limitations of the above methods, an integrated simulation-surrogate model methodology is presented to predict the availability of production lines in this paper. For reasons of generality, our research is limited to the analysis of discrete serial-parallel lines with unreliable workstations and finite intermediate buffers. The rest of the article is organized as follows. Production line description, assumptions and symbols are described in Section 2. Section 3 proposes a general DES to simulate the production process and obtain the production lines' availability. Section 4 outlines a surrogate model combined with LHS and Kriging model for prediction. An application in a rough machining production line for crankshaft is presented in Section 5. Finally, conclusions and prospects close the paper in Section 6.

2. Production line model, assumptions and symbols

2.1 Production line description

A discrete production line is often organized with workstations connected in product-flow layout and separated by intermediate buffers. A workstation may be composed of several machines in series/parallel or just one machine (in this case the terms "workstation" and "machine" are used interchangeably). The graphical-based structural model of an N-workstation production line is given in Fig. 1, where workstations are represented by squares and buffers are represented by circles. In a production line, workpieces from outside enter the system from the first workstation WS_1. Each workpiece is processed by WS_1 within operation time T_1, after which it is transferred to the first buffer B_1. Then it moves in the direction of arrows until it is finished by the last workstation WS_N, and exits the system.
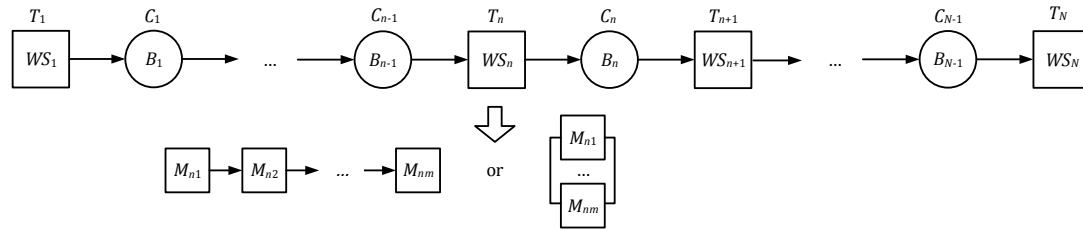
Fig. 1 The block diagram of an *N*-workstation production line

In the real production operations, the machines always experience random breakdowns. Consequently, failure of one workstation may affect all other workstations upstream and downstream. This complex phenomenon is generally regarded as perturbation propagation [11], which makes the analysis of the production line difficult. To limit the propagation of disruptions, buffers are usually placed between workstations. In fact, buffers are capable to provide continuity by means of saving parts from the upstream subsystem and releasing parts to the downstream subsystem. The buffer inventory can provide a period of isolation time [20] for maintenance actions before the buffer becomes empty or full without bringing down the entire system immediately. From this point of view, buffers alleviate this mutual interference by decoupling adjacent workstations from "rigid" connection to "elastic" connection.

Except the failures of machines, another reason for line inefficiency is the workstations' interference: starvation and blocking. A workstation is called *starved* when its upstream buffer is empty (buffer inventory is 0). It is said to be *blocked* when its downstream buffer is full (buffer inventory is its maximum capacity C_n). Taking WS_n as an example, starvation is the phenomenon that when WS_n has finished a workpiece it is forced to wait because B_{n-1} is empty. When a workstation is *up*, it is said to be *busy* when it is processing a workpiece, and is said to be *idle* when it is either starved or blocked. Generally, uptimes (including busy times and idle times) and downtimes exhibit statistical regularity, and can be expressed as independent and identically distributed random variables. Thus, the state is summarized as follows:

$$\text{the state of } WS_n = \begin{cases} up \begin{cases} busy \\ idle \begin{cases} starved \\ blocked \end{cases} \end{cases} \\ down \end{cases}$$

As mentioned above, the parallel machines are simplified to one workstation in this research (see Fig. 1). They are generally used to balance the production line, and have the same operation as well as configuration in most situations. They are therefore assumed to have identical operation time as well as parameter distributions. Thus the workstation's operation time equals to the machines' operation time divided by the number of parallel machines. Moreover, we defined "degradation ratio" as the production capacity coefficient of the workstation when one of the parallel machines is down. For example, the degradation ratio is 2/3 when the workstation contains three parallel machines. Without loss of generality, let it be 0 in the series case.

2.2 Assumptions

The following additional assumptions are also used:

(a) As the supply and storage of production line are beyond the scope of this research and they are considered to be infinite. In other words, WS_1 is never starved and WS_N is never blocked.

(b) Scheduled downtimes such as breaks, meetings, and preventative maintenance are not concerned in this paper.

(c) Operation time of each workstation which contains transfer time and setup time is constant because most gantry robots and machines are controlled by predetermined NC code.

(d) Failures don't destroy workpieces. Therefore, the workpieces remain at the machines during maintenance, and processing resumes when the machines are up.

(e) The two-parameter Weibull distribution is employed to model uptimes and downtimes. Although the proposed DES model is appropriate for any distributed workstations, the Weibull distribution is one of the most common distributions in reliability engineering and can be conveniently transformed into exponential distribution, which is widely used in analytical approaches.

2.3 Symbols of system parameters

It is necessary to present a summary of the symbols as well as their explanations used in this research. The system input and output parameters are listed in Table 1 and Table 2, respectively.

Table 1 The input parameters of model

Notations	Explanation of the notations
N	Total number of workstations
T_n	Operation time of WS_n, and $\boldsymbol{T} = (T_1, T_2, \cdots, T_N)^{\mathrm{T}}$
C_n	Maximum capacity of B_n (including the space at WS_{n+1}), and $\boldsymbol{C} = (C_1, C_2, \cdots, C_{N-1})^{\mathrm{T}}$
D_n	Degradation ratio of WS_n, and $\boldsymbol{D} = (D_1, D_2, \cdots, D_N)^{\mathrm{T}}$, $D_n \in [0, 1)$
ut_{ni}	Uptime of WS_n before the ith failure, and $UT_n = \{ut_{n1}, ut_{n2}, \cdots, ut_{ni}, \cdots\}$
dt_{ni}	Downtime of WS_n during the ith failure, and $DT_n = \{dt_{n1}, dt_{n2}, \cdots, dt_{ni}, \cdots\}$
α_n	Scale parameter of Weibull distribution for WS_n, and $\boldsymbol{\alpha} = (\alpha_1, \alpha_2, \cdots, \alpha_N)^{\mathrm{T}}$
β_n	Shape parameter of Weibull distribution for WS_n, and $\boldsymbol{\beta} = (\beta_1, \beta_2, \cdots, \beta_N)^{\mathrm{T}}$
t_{sim}	Simulation time
t_{warmup}	Warmup period
r	Number of independent replications of the simulation

Table 2 The output parameters of model

Notations	Explanation of the notations
$k_n(t)$	Inventory level of B_n at time t, and $0 \le k_n(t) \le C_n$
$ct_n(t)$	Cumulative busy time of WS_n at time t
$op_n(t)$	Output of WS_n at time t
$s_n(t)$	State of WS_n at time t, and $s_n = D_n$ during downtime, $s_n = 1$ during busy time, $s_n = $ "*starved*" or "*blocked*" during idle time
OP	Total amount of output, and $OP = op_N(t_{\mathrm{sim}})$
$OPh(t)$	Hourly output of system at time t
$MTBF$	Mean time between failures, and $MTBF_n = \bar{u}t_n$
$MTTR$	Mean time to repair, and $MTTR_n = \bar{d}t_n$
A	Availability of the whole line, defined as the probability of system being processing. Actually, in the steady state, if no failures occur, the system can process a workpiece in every bottleneck operation time T_{max}, which is the maximum operation time of workstations. Thus, the total processing time of system is $OP \times T_{\mathrm{max}}$, and $A = OP \times T_{\mathrm{max}}/t_{\mathrm{sim}}$.

3. Discrete event simulation

In this section a general DES is developed under Matlab environment to simulate the manufacturing process of production lines. The simulation model can provide real-time information on operating characteristics, and evaluate the availability of production lines with various input parameters. Meanwhile, it has a good compatibility with subsequent surrogate model programs.

3.1 Simulation process

The flow chart of the proposed DES is shown in Fig. 2 and the main steps are described as follows:

(a) Set the input parameters for the DES model: N, $\boldsymbol{T}$, $\boldsymbol{C}$, $\boldsymbol{D}$, t_{sim}, and t_{warmup}.
(b) Initialization of $k_n(0)$, $ct_n(0)$ and $op_n(0)$ for every buffer and workstation with the default values all zero.
(c) Generate sample sets UT_n and DT_n with required distribution by Monte Carlo technique for every workstation. Accumulate and sort these data in chronological order until t_{sim} ends.

(d) Scan the real-time state for every workstation and buffer at each time unit by means of a nested loop, of which the outer loop runs from $t = 1$ to $t = t_{sim}$ and the inner loop runs from $n = 1$ to $n = N$. In this process, discriminate WS_n between up and down according to the samples obtained from step (c). If WS_n is up, discriminate WS_n between busy and idle according to $k_{n-1}(t\text{-}1)$ and $k_n(t\text{-}1)$. Then record $k_n(t)$, $ct_n(t)$, $op_n(t)$ and $s_n(t)$.

(e) Calculate OP, OPh and A at the end of simulation.
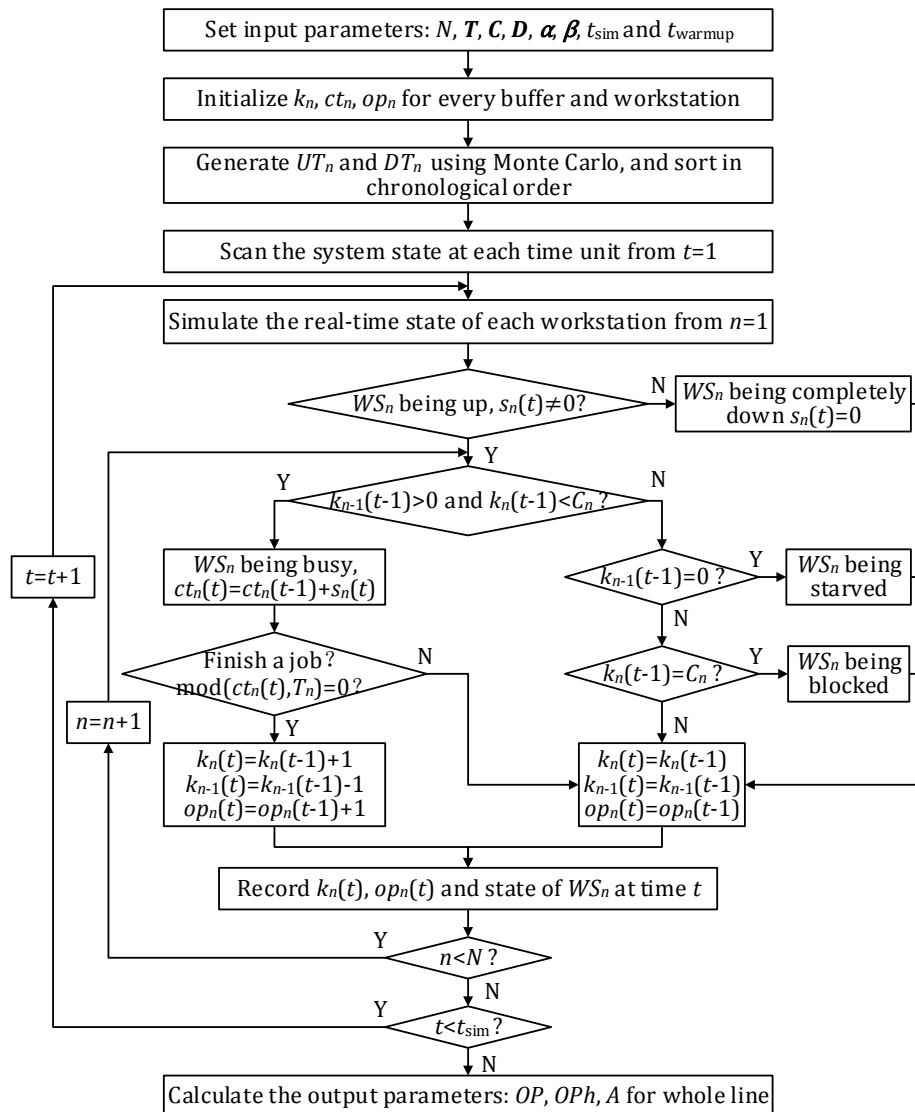
Fig. 2 The flow chart of DES model for production lines

3.2 Validation

The validation of proposed DES model was conducted by comparing the results with Plant Simulation software in different lines. We investigated 4 cases as follows, with the configuration details in Table 3:

Case 1: Synchronous series line with same workstations;
Case 2: Synchronous series line with different workstations;
Case 3: Asynchronous series-parallel line, also be described in Section 5;
Case 4: Asynchronous series-parallel line with different buffers and Weibull workstations.

We performed $r = 100$ independent repeated trials with each of simulation length $t_{sim} = 1440$ h (3 m × 30 d × 16 h). The average results, which are summarized in Table 4, indicate that the percentage errors of OP are extremely small. That means the proposed DES model is practicable.

**Table 3** Configuration details of 4 cases

Parameters		Case 1	Case 2	Case 3	Case 4
N		5	5	5	5
T (s)		(200,200,200,200,200)	(200,200,200,200,200)	(180,200,190,180,190)	(180,200,190,180,190)
K		(10,10,10,10)	(10,10,10,10)	(10,10,10,10)	(10,15,10,15)
D		(0,0,0,0,0)	(0,0,0,0,0)	(0,0,0.5,0,0)	(0,0,0.5,0,0)
UT (h)	α	(400,400,400,400,400)	(400,600,350,300,450)	(400,600,350,300,450)	(400,600,350,300,450)
	β	(1,1,1,1,1)	(1,1,1,1,1)	(1,1,1,1,1)	(1.2,0.8,1.2,0.8,1.2)
DT (h)	α	(2,2,2,2,2)	(1.5,2,2.5,3,1.5)	(1.5,2,2.5,3,1.5)	(1.5,2,2.5,3,1.5)
	β	(1,1,1,1,1)	(1,1,1,1,1)	(1,1,1,1,1)	(0.8,1.2,0.8,1.2,0.8)

Table 4 Simulation results of Plant Simulation and the DES in proposed approach

Parameters	Case 1			Case 2			Case 3			Case 4		
	Matlab	PS	Error, %	Matlab	PS	Error, %	Matlab	PS	Error, %	Matlab	PS	Error, %
OP	25351	25369	-0.0710	25294	25301	-0.0277	25480	25486	-0.0235	25505	25511	-0.0235

3.3 Warm-up

The start-up or initial transient problem is a common problem in simulation process. In order to ensure that observations can represent steady-state behaviour, the warming up or initial-data deletion technique is often suggested. In this research, a graphical procedure proposed by Welch is employed to choose the warm-up period (see [19]). This procedure can smooth out the plot of observations based on r independent replications of the simulation and moving average with w (where w is the *window*, a parameter to adjust the smoothness). Output parameters $OPh(t)$ is selected as the observation, because A is closely related to OP. Taking case 3 as an example, the moving averages for $OPh(t)$ with w = 50 h are shown Fig. 3. From the plot we chose a warmup period of t_{warmup} = 80 h (5 d × 16 h).
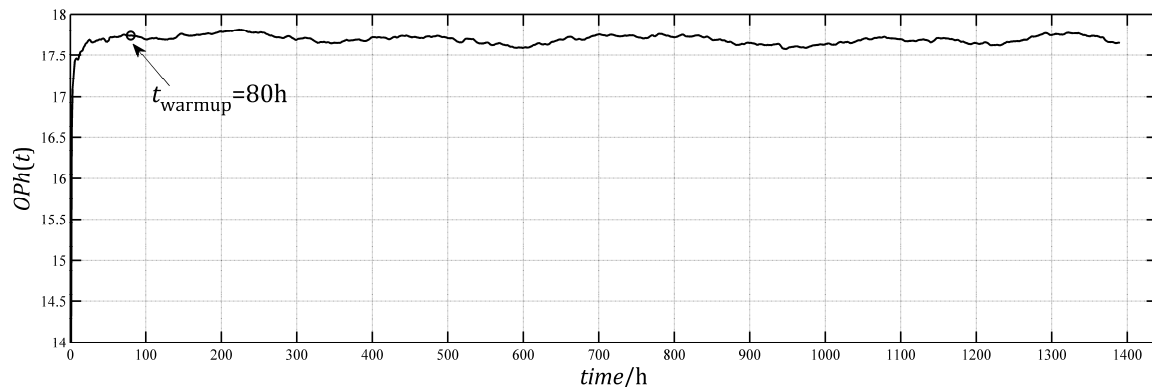
Fig. 3 Moving averages for $OPh(t)$ with w = 50 h

4. Prediction based on Surrogate Model

To increase the efficiency, design of experiments (DOE) [21] technique is usually integrated into the simulation, which also be referred to as computer experiments or simulation experiments [22]. This technique can help us to explore the relationship between input parameters (factors) and performance measures (responses) with the least amount of simulating. Another use of DOE is to construct a surrogate model, also known as metamodel or response surfaces, which is a simplified model of the simulation model for representing the quantitative relationship between factors and responses [23]. Fig. 4 illustrates the relationship of different models. Surrogate model provides one approach to predict the responses from a limited set of simulated factor-level configurations. In this section, a prediction method based on surrogate model, which combines Latin hypercube sampling (LHS) and Kriging model, is presented to predict the availability of production lines.
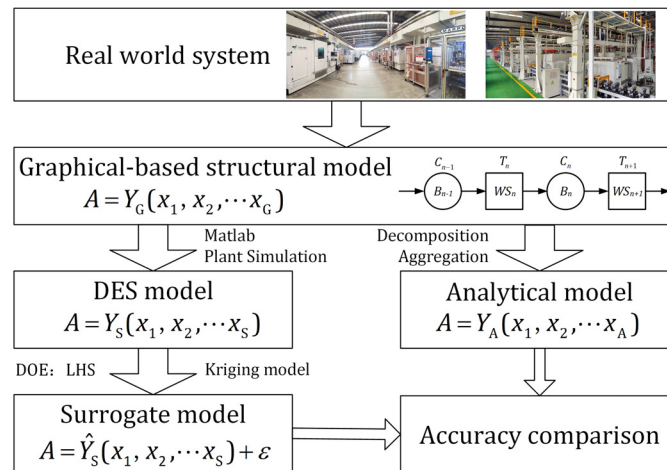
Fig. 4 The relationship of different models

4.1 Latin hypercube sampling

As a kind of space filling design, LHS is widely used for computer experiments because its stratification property allows it to cover the input domain uniformly in a relatively small sample size. A LHS with p sample points in q dimensions is written as an $p \times q$ matrix $\boldsymbol{X} = [\boldsymbol{x}_1, \boldsymbol{x}_2, \cdots, \boldsymbol{x}_p]^{\mathrm{T}}$, in which each column represents a factor and each row $\boldsymbol{x}_j = [x_{j(1)}, x_{j(2)}, \cdots, x_{j(q)}]$ represents a sample. Firstly, a LHS divides each dimension into p equal levels and gets p^q grids. Then select p of them to make that exactly one is selected at each level. Finally, generate one point randomly in every grid and get p sample points. To get better uniformity, various criterions are available to optimize the design.

In this paper, the LHS experiment plan was created by Matlab function *lhsdesign* with *maxmin* criterion (maximize minimum distance between points) in 40 iterations. For example, Fig. 5 shows a LHS with 100 sample points for variation of workstations' reliability parameters, which is in the range of $MTBF \in [-200, 200]$ h and $MTTR \in [-1, 1]$ h.
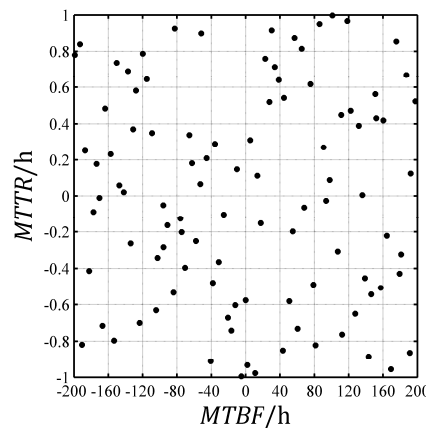

Fig. 5 A 100 × 2 LHS experiment plan

4.2 Kriging model

Kriging is a popular interpolation methodology for computer experiments to construct a cost-effective model as a surrogate to the tedious and time-consuming engineering simulation. Actually, the Kriging is the best linear unbiased interpolation and has the ease of immediate validation by measuring its uncertainty [24]. Compared with traditional polynomial regression, Kriging can give better global predictions because it assumes that the prediction errors are correlated, i.e., gives more weight to 'neighbouring' observations [25].

In Kriging model, the simulation output at design point $\boldsymbol{x}$ is defined as:

$$Y(\boldsymbol{x}) = \boldsymbol{b}(\boldsymbol{x})^{\mathrm{T}} \boldsymbol{\lambda} + Z(\boldsymbol{x}) \tag{1}$$

This model is built by adding up two terms: The first term, which represents the global trend, is a linear combination where $\boldsymbol{b}(\boldsymbol{x})$ is a vector of given basis functions of $\boldsymbol{x}$, and $\boldsymbol{\lambda}$ is a vector of unknown coefficients need to be estimated from the simulation results. The second one $Z(\boldsymbol{x})$ is a local bias expressed by a second-order stationary random process with mean zero and covariance $\mathrm{Cov}[Z(\boldsymbol{x}), Z(\boldsymbol{x}')] = \sigma^2 R(\|\boldsymbol{x} - \boldsymbol{x}'\|; \boldsymbol{\theta})$, where σ^2 can be elucidated as the variance of $Z(\boldsymbol{x})$ for all $\boldsymbol{x}$, and R is the spatial correlation function (SCF) that depends on the 'distances' $\|\boldsymbol{x} - \boldsymbol{x}'\|$ (Euclidean norm). For simplification, the form $R(\|\boldsymbol{x} - \boldsymbol{x}'\|; \boldsymbol{\theta}) = \prod_{v=1}^{q} R(|x_{(v)} - x'_{(v)}|; \theta_{(v)})$ is used for studying q-dimensional problems. The parameters σ^2 and $\boldsymbol{\theta}$ are estimated from the experimental data.

There are two steps to build a surrogate model by Kriging: modelling and prediction. Firstly, model $Y(\boldsymbol{x})$, i.e., estimate the unknown parameters based on the simulated points. Secondly, predict the availability of production lines for given points. We performed these tasks by using a Matlab toolbox called DACE (Design and Analysis of Computer Experiments) [26]. The basis functions $\boldsymbol{b}(\boldsymbol{x})$ are commonly expressed as polynomials of orders d = 0, 1 or 2. For SCF, the toolbox provides seven types and we choose three of them: exponential, Gaussian and spherical, which are defined as follows, respectively:

$$R_{\exp}(|x - x'|; \theta) = \exp(-\theta|x - x'|) \tag{2}$$

$$R_{\mathrm{Gauss}}(|x - x'|; \theta) = \exp(-\theta|x - x'|^2) \tag{3}$$

$$R_{\mathrm{sph}}(|x - x'|; \theta) = 1 - 1.5\xi^2 + 0.5\xi^3, \qquad \xi = \min\{1, \theta|x - x'|\} \tag{4}$$

4.3 Prediction accuracy

The goodness of prediction fit for different models is measured by their mean absolute percentage error ($MAPE$), root-mean-square error ($RMSE$) and R-square (R^2), expressed as:

$$MAPE = \frac{1}{p} \sum_{j=1}^{p} \left| \frac{y_j - \hat{y}_j}{y_j} \right| \times 100\% \tag{5}$$

$$RMSE = \sqrt{\frac{1}{p} \sum_{j=1}^{p} (y_j - \hat{y}_j)^2} \tag{6}$$

$$R^2 = 1 - \frac{\sum_{j=1}^{p} (y_j - \hat{y}_j)^2}{\sum_{j=1}^{p} (y_j - \bar{y})^2} \tag{7}$$

where y_j is the simulated value and $\hat{y}_j$ is the forecast value. For the first two indicators, a value closer to zero means a better fit. With respect to R^2, it can take any value between zero and one, and the higher the value, the better accuracy of predictions will be.

5. Case study

In this section, a rough machining production line before heat treatment for car crankshaft is given to illustrate application of the proposed method. It is organized with five workstations, among which the third one is composed of two parallel machines (see Fig. 6). The operating content is described in Table 5, and the design parameter is the same as Case 3 in Section 3.2.
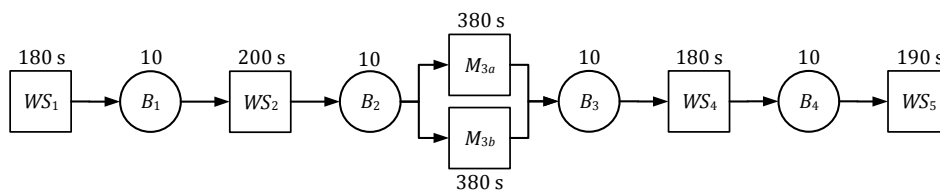

Fig. 6 The block diagram of crankshaft production line

Table 5 The operating content of crankshaft production line

No.	Operating content
OP1	Milling two end faces and location seam, drilling center hole
OP2	Turning rear and front journals
OP3	Finishing turning rear and front journals, main journals and undercuts
OP4	Milling pin journals, undercuts and outer rings of balancers
OP5	Drilling vertical/oblique oil holes and chamfers

To investigate how equipment management influences the performance of the whole line, the availability was predicted at different levels of workstations' reliability and maintainability. We made r = 100 independent replications of the proposed DES with t_{sim} = 1440 h and t_{warmup} = 80 h for p = 100 sample points obtained via experiment plan in Section 4.1. These average simulated values of A are shown by black scatter points in Fig. 7. Then we built the Kriging model from above simulation result, and made predictions for grid points with step sizes of $MTBF$ = 40 h and $MTTR$ = 0.2 h. Another simulation was performed for these grid points as verification group by Plant Simulation software. Fig. 7 displays the predictions obtained from Kriging model with second-order basis functions and Gaussian correlation function. We can see that the surface basically overlaps that of simulated values, which demonstrates that the prediction accuracy is high enough. Meanwhile, the response surface proves that A is an increasing function of $MTBF$ and a decreasing function of $MTTR$.
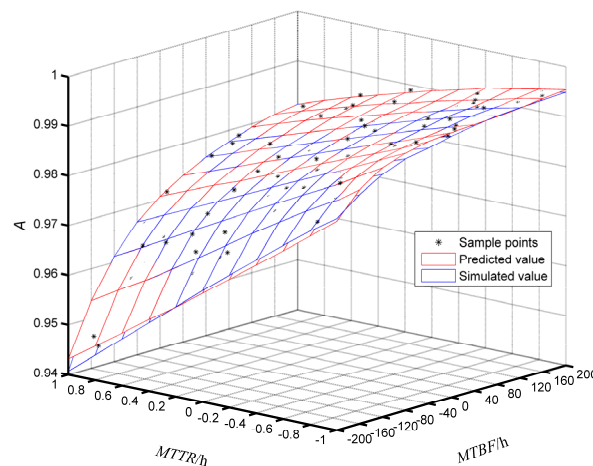
Fig. 7 Response surface of availability A

We also compared various Kriging models with four other methods to show the accuracy of the proposed method. They are moving least square (MLS) method with Gaussian weight function [27], Matlab function *griddata* with *v4* method, aggregation approximate method [11] and semi-analytical simulation [28]. The first two methods interpolate the same sample points and act like Kriging model. The results (listed in Table 6) indicate that Kriging models have the lowest $MAPE$, $RMSE$ and the highest R^2, which means that the proposed method provides better predictions than the other four methods. The methods combining simulation and surrogate model are generally better than analytical methods. Furthermore, for Kriging models, Gaussian correlation function delivers better performance, and the other two perform basically the same. The ideal basis function is second-order polynomial.

Table 6 Prediction accuracy of availability A

Method	$MAPE$ ($\times 10^{-2}$)			$RMSE$ ($\times 10^{-4}$)			R^2		
	$d = 0$	$d = 1$	$d = 2$	$d = 0$	$d = 1$	$d = 2$	$d = 0$	$d = 1$	$d = 2$
Exponential Kriging	6.9253	6.7465	5.4850	10.835	9.4687	7.2949	0.99034	0.99262	0.99562
Gaussian Kriging	4.2619	4.3389	**4.1437**	5.8173	5.9956	**5.7200**	0.99721	0.99704	**0.99731**
Spherical Kriging	6.6904	6.3328	5.6112	10.596	8.7454	7.3571	0.99076	0.99370	0.99554
Gaussian MLS	46.527	17.823	6.5864	67.015	23.145	8.8208	0.63028	0.95590	0.99359
Griddata with v4		8.5439			12.922			0.98625	
Aggregation approximate		65.843			76.846			0.51385	
Semi-analytical simulation		11.139			15.356			0.98059	

6. Conclusion and further research

In this paper, a new procedure has been proposed to predict the availability of discrete production lines with unreliable workstations and finite buffers. The main advantages are its flexibility and property that it is not restricted by specified distributions. This procedure consists of two phases: DES and surrogate model. The first phase involves conducting computer experiments to obtain the availability of production lines under different system parameters with the help of DOE. These input parameters can describe production lines with different structures, operation times, uptimes, downtimes and buffer capacities. The second phase aims to predict the availability based on the simulated results by surrogate model. It is constructed by combining Kriging model and LHS.

A case study of crankshaft production line has shown that the proposed method was practical. It provided better predictions than other interpolation methods, approximate analytical method and semi-analytical simulation. For Kriging model, Gaussian correlation function and second-order basis function predicted with the best performance. We also got the response surface of whole line availability versus workstations' *MTBF* and *MTTR*. The main contribution of this paper are developing a new DES under Matlab environment and importing Kriging model to the domain of modeling the availability of production lines. Although the proposed method was flexible and accurate, the downside is that its efficiency is still lower than analytical methods. Since the method is based on DES, the time-consuming problem will inevitably emerge. Even though LHS and Kriging help a lot, there is still room for improvement, especially for the long lines.

In the future, we will investigate the prediction of other performance measures with more input parameters for higher dimensions, such as speed losses, defect losses or process failure. Moreover, sensitivity analysis methods should be helpful to gain a deeper understanding of the relationships between factors and responses. It would also be interesting to observe the performance of other DOE techniques and surrogate models.

Acknowledgement

References

[1] Lv, Y., Zhang, J., Qin, W. (2017). A genetic regulatory network-based sequencing method for mixed-model assembly lines, *Advances in Production Engineering & Management*, Vol. 12, No. 1, 62-74, doi: 10.14743/apem2017.1.240.

[2] Govil, M.K., Fu, M.C. (1999). Queueing theory in manufacturing: A survey, *Journal of Manufacturing Systems*, Vol. 18, No. 3, 214-240, doi: 10.1016/S0278-6125(99)80033-8.

[3] Dallery, Y., Gershwin, S.B. (1992). Manufacturing flow line systems: A review of models and analytical results, *Queueing Systems*, Vol. 12, No. 1-2, 3-94, doi: 10.1007/BF01158636.

[4] Gershwin, S.B. (1987). An efficient decomposition method for the approximate evaluation of tandem queues with finite storage space and blocking, *Operations Research*, Vol. 35, No. 2, 291-305, doi: 10.1287/opre.35.2.291.

[5] Dallery, Y., David, R., Xie, X.-L. (1988). An efficient algorithm for analysis of transfer lines with unreliable machines and finite buffers, *IIE Transactions*, Vol. 20, No. 3, 280-283, doi: 10.1080/07408178808966181.

[6] Burman, M.H. (1995). *New results in flow line analysis*, PhD Thesis, Operations Research Center, Massachusetts Institute of Technology, USA.

[7] Xia, B., Xi, L., Zhou, B. (2012). An improved decomposition method for evaluating the performance of transfer lines with unreliable machines and finite buffers, *International Journal of Production Research*, Vol. 50, No. 15, 4009-4024, doi: 10.1080/00207543.2011.587842.

[8] Xia, B., Xi, L., Zhou, B., Du, S. (2013). An efficient analytical method for performance evaluation of transfer lines with unreliable machines and finite transfer-delay buffers, *International Journal of Production Research*, Vol. 51, No. 6, 1799-1819, doi: 10.1080/00207543.2012.713137.

[9] Colledani, M., Gershwin, S.B. (2013). A decomposition method for approximate evaluation of continuous flow multi-stage lines with general Markovian machines, *Annals of Operations Research*, Vol. 209, No. 1, 5-40, doi: 10.1007/s10479-011-0961-9.

[10] Li, J., Blumenfeld, D.E., Huang, N., Alden, J.M. (2009). Throughput analysis of production systems: Recent advances and future topics, *International Journal of Production Research*, Vol. 47, No. 14, 3823-3851, doi: 10.1080/00207540701829752.

[11] Li, J., Meerkov, S.M. (2009). *Production Systems Engineering*, Springer, New York, USA, doi: 10.1007/978-0-387-75579-3.

[12] Kuhn, W. (2006). Digital factory – Simulation enhancing the product and production engineering process, In: *Proceedings of the 2006 Winter Simulation Conference*, Monterey, USA, 1899-1906.

[13] Betterton, C.E., Cox, J.F. (2012). Production rate of synchronous transfer lines using Monte Carlo simulation, *International Journal of Production Research*, Vol. 50, No. 24, 7256-7270, doi: 10.1080/0207543.2011.645081.

[14] Dhouib, K., Gharbi, A., Ayed, S. (2009). Simulation based throughput assessment of non-homogeneous transfer lines, *International Journal of Simulation Modelling*, Vol. 8, No. 1, 5-15, doi: 10.2507/IJSIMM08(1)1.111.

[15] Dhouib, K., Gharbi, A., Ayed, S. (2008). Availability and throughput of unreliable, unbuffered production lines with non-homogeneous deterministic processing times, *International Journal of Production Research*, Vol. 46, No. 20, 5651-5677, doi: 10.1080/00207540701294635.

[16] Padhi, S.S., Wagner, S.M., Niranjan, T.T., Aggarwal, V. (2013). A simulation-based methodology to analyse production line disruptions, *International Journal of Production Research*, Vol. 51, No. 6, 1885-1897, doi: 10.1080/00207543.2012.720389.

[17] Heshmat, M., El-Sharief, M.A., El-Sebaie, M.G. (2013). Simulation modeling of automatic production lines with intermediate buffers, *Journal of Engineering Sciences,* Vol. 41, No. 6, 2175-2189.

[18] Supsomboon, S., Hongthanapach, K. (2014). A simulation model for machine efficiency improvement using reliability centered maintenance: case study of semiconductor factory, *Modelling and Simulation in Engineering*, Vol. 2014, 1-9, doi: 10.1155/2014/956182.

[19] Law, A.M. (2015). *Simulation modeling and analysis,* 5th edition, McGraw-Hill, New York, USA.

[20] Macchi, M., Kristjanpoller, F., Garetti, M., Arata, A., Fumagalli, L. (2012). Introducing buffer inventories in the RBD analysis of process production systems, *Reliability Engineering & System Safety*, Vol. 104, 84-95, doi: 10.1016/j.ress.2012.03.015.

[21] Montgomery, D.C. (2013). *Design and analysis of experiments,* 8th edition, Wiley, New Jersey, USA.

[22] Barton, R.R. (2013). Designing simulation experiments, In: *Proceedings of the 2013 Winter Simulation Conference*, Washington, USA, 342-353, doi: 10.1109/WSC.2013.6721432.

[23] Kleijnen, J.P.C. (2015). *Design and analysis of simulation experiments,* 2nd edition*,* Springer International Publishing, Switzerland, doi: 10.1007/978-3-319-18087-8.

[24] Kleijnen, J.P.C. (2009). Kriging metamodeling in simulation: A review, *European Journal of Operational Research*, Vol. 192, No. 3, 707-716, doi: 10.1016/j.ejor.2007.10.013.

[25] Matta, A., Pezzoni, M., Semeraro, Q. (2012). A Kriging-based algorithm to optimize production systems approximated by analytical models, *Journal of Intelligent Manufacturing*, Vol. 23, No. 3, 587-597, doi: 10.1007/s10845-010-0397-0.

[26] Lophaven, S.N., Nielsen, H.B., Søndergaard, J. (2002). *DACE: A Matlab Kriging toolbox, Version 2.0,* Informatics and Mathematical Modelling, Technical University of Denmark, Lyngby, Denmark.

[27] Jia, J. Moving Least Square(MLS2D), from *http://cn.mathworks.com/matlabcentral/fileexchange/34547-moving-least-square-mls2d-*, accessed January12, 2012.

[28] Jia, Y., Yang, Z., Li, G., Du, X., He, J., Wang, L., Li, Q. (2016). An efficient semi-analytical simulation for availability evaluation of discrete production lines with unreliable machines, In: *2016 International Conference on System Reliability and Science*, Paris, France, 115-121, doi: 10.1109/ICSRS.2016.7815849.

This page intentionally left blank.

Calendar of events

- 2nd International Conference on Design and Production Engineering, Birmingham, Birmingham, UK, August 21-22, 2017.

- International Conference on Materials and Intelligent Manufacturing, Singapore, 21-23 August, 2017.

- 8th International Conference on Advances in Engineering and Technology, Budapest, Hungary, September 6-7, 2017.

- 11th International Conference on Advanced Materials & Processing, Edinburgh, Scotland, September 7-8, 2017.

- 8th CEBU International Conference on Recent Trends in Engineering and Technology, Cebu, Philippines, September 21-22, 2017.

- 5th International Conference on Industrial Engineering, Robotics & Engineering Materials, Kuala Lumpur, Malaysia, September 27-28, 2017.

- 3rd International Conference and Exhibition on Automobile Engineering, Berlin, Germany, September 28-29, 2017.

- 2nd International Conference on Advanced Production and Industrial Engineering, Delhi, India, October 6-7, 2017.

- 2017 International Conference on Material Engineering and Manufacturing – EI Compendex, Scopus, Chengdu, China, October 9-11, 2017.

- LEAP HR: Manufacturing, Cleveland, United States of America, October 17-18, 2017.

- 4th World Congress on Robotics and Artificial Intelligence, Osaka, Japan, October 23-24, 2017.

- 31st European Simulation and Modelling Conference, Lisbon, Portugal, October 25-27, 2017.

- 4th International Conference on Mechanical, Materials and Manufacturing, Atlanta, United States of America, October 25-27, 2017.

- 28th DAAAM International Symposium, Zadar, Croatia, November 8-11, 2017.

- International Conference on Mechatronics, Automation and Intelligent Materials, Paris, France, November 13-14, 2017.

- 7th International Conference on Innovative Trends in Engineering, Technology, Computers and Applied Sciences, Tokyo, Japan, November 23-24, 2017.

- 14th International Conference on Functional Energy Materials, Dallas, Texas, USA, December 6-7, 2017.

- IEEE International Conference on Industrial Engineering and Engineering Management, Singapore, December 10-13, 2017.

- The 7th Advanced Functional Materials and Devices – SCOPUS, EI Compendex, Havana, Cuba, December 13-15, 2017.

- 11th International Conference on Data Mining, Computers, Communication & Industrial Applications, Kuala Lumpur, Malaysia, December 14-15, 2017.

- 10th International Conference on Advances in Mechanical Design and Manufacturing, Giza, Egypt, December 18-20, 2017.

- 9th International Conference on Mechatronics and Manufacturing, Phuket Island, Thailand, January 27-29, 2018.

- 2018 International Conference on Robotics and Intelligent System, EI Compendex, Scopus, Amsterdam, Netherlands, February 21-23, 2018.

- The 8th International Conference on Key Engineering Materials, Osaka, Japan, March 16-18, 2018.

- 2nd International Conference on 3D Printing Technology and Innovations, London, UK, March 19-20, 2018.

- The 2018 International Conference on Robotics and Intelligent Control, Nagoya, Japan, April 27-30, 2018.

Notes for contributors

General

Articles submitted to the *APEM journal* should be original and unpublished contributions and should not be under consideration for any other publication at the same time. Manuscript should be written in English. Responsibility for the contents of the paper rests upon the authors and not upon the editors or the publisher. Authors of submitted papers automatically accept a copyright transfer to *Production Engineering Institute*, *University of Maribor*. For most up-to-date information on publishing procedure please see the *APEM journal* homepage *apem-journal.org*.

Submission of papers

A submission must include the corresponding author's complete name, affiliation, address, phone and fax numbers, and e-mail address. All papers for consideration by *Advances in Production Engineering & Management* should be submitted by e-mail to the journal Editor-in-Chief:

Miran Brezocnik, Editor-in-Chief
UNIVERSITY OF MARIBOR
Faculty of Mechanical Engineering
Production Engineering Institute
Smetanova ulica 17, SI – 2000 Maribor
Slovenia, European Union
E-mail: editor@apem-journal.org

Manuscript preparation

Manuscript should be prepared in *Microsoft Word 2007* (or higher version) word processor. *Word .docx* format is required. Papers on A4 format, single-spaced, typed in one column, using body text font size of 11 pt, should not exceed 12 pages, including abstract, keywords, body text, figures, tables, acknowledgements (if any), references, and appendices (if any). The title of the paper, authors' names, affiliations and headings of the body text should be in *Calibri* font. Body text, figures and tables captions have to be written in *Cambria* font. Mathematical equations and expressions must be set in *Microsoft Word Equation Editor* and written in *Cambria Math* font. For detail instructions on manuscript preparation please see instruction for authors in the *APEM journal* homepage *apem-journal.org*.

The review process

Every manuscript submitted for possible publication in the *APEM journal* is first briefly reviewed by the editor for general suitability for the journal. Notification of successful submission is sent. After initial screening, and checking by a special plagiarism detection tool, the manuscript is passed on to at least two referees. A double-blind peer review process ensures the content's validity and relevance. Optionally, authors are invited to suggest up to three well-respected experts in the field discussed in the article who might act as reviewers. The review process can take up to eight weeks. Based on the comments of the referees, the editor will take a decision about the paper. The following decisions can be made: accepting the paper, reconsidering the paper after changes, or rejecting the paper. Accepted papers may not be offered elsewhere for publication. The editor may, in some circumstances, vary this process at his discretion.

Proofs

Proofs will be sent to the corresponding author and should be returned within 3 days of receipt. Corrections should be restricted to typesetting errors and minor changes.

Offprints

An e-offprint, i.e., a PDF version of the published article, will be sent by e-mail to the corresponding author. Additionally, one complete copy of the journal will be sent free of charge to the corresponding author of the published article.

Contents