



# A big data analytics based machining optimisation approach

Wei Ji<sup>1,2</sup>  · Shubin Yin<sup>2</sup> · Lihui Wang<sup>1</sup>

Received: 31 May 2018 / Accepted: 21 July 2018 / Published online: 25 July 2018  
© The Author(s) 2018

## Abstract

Currently, machine tool selection, cutting tool selection and machining conditions determination are not usually performed at the same time but progressively, which may lead to suboptimal or trade-off solutions. Targeting this issue, this paper proposes a big data analytics based optimisation method for enriched Distributed Process Planning by considering machine tool selection, cutting tool selection and machining conditions determination simultaneously. Within the context, the machining resources are represented by data attributes, i.e. workpiece, machining requirement, machine tool, cutting tool, machine conditions, machining process and machining result. Consequently, the problem of machining optimisation can be treated as a statistic problem and solved by a hybrid algorithm. Regarding the algorithm, artificial neural networks based models are trained by machining data and used as optimisation objectives, whereas analytical hierarchy process is adopted to decide the weights of the multi-objective optimisation; and evolutionary algorithm or swarm intelligence is proposed to perform the optimisation. Finally, the results of a simplified proof-of-concept case study are reported to validate the proposed approach, where a Deep Belief Network model was trained by a set of hypothetic data and used to calculate the fitness of a genetic algorithm.

**Keywords** Big data analytics · Machining optimisation · Hybrid algorithm · Deep belief network · Genetic algorithm

## Introduction

Modern industries are characterised by smartness and intelligence in manufacturing, as identified in Industry 4.0 and Made in China 2025 where automation and digitalisation are the key elements. In modern manufacturing, machining being highly nonlinear is the most complex process. In order to decompose the complexity, a method named Distributed Process Planning (DPP) is used to divide the machining process planning into supervisory planning, execution control and operation planning (Wang et al. 2003). In the supervisory planning, generic processes are obtained based on machining features and machining knowledge, whereas the resource-specific processes including machine tool selection, cutting tool selection and machining conditions determination are carried out during the execution control and operation planning. Furthermore, due to the limitations of conventional methods, the optimisations of machine/cutter/cutting

parameter selections are usually performed separately, which ignores the relationships between them, constrains the solution space, and belongs to a local optimisation. Integrating the decision processes into one based on data science contributes to the global optimisation. However, it introduces several challenges: (1) the solution space is determined by the high dimensions of data; (2) it is difficult to establish high-accuracy objective functions; and (3) the needs in real industry are multi-objective. Big data analytics, used successfully in many areas including product lifecycle management, supply chain management and predictive maintenance (Wan et al. 2017), shed lights on these issues. In this research, based on DPP, big data analytics is applied to enrich its functionalities to select machine tool and cutting tool, and to optimise machining conditions simultaneously. The remainder of this paper outlines this approach and is organised as follows. A review of related work is provided in “[Related work](#)” section. Combining with big data analytics, the approach of enriched DPP is reported in “[Enriched distributed process planning \(DPP\)](#)” section, followed by a solution of enriched DPP in “[Solution strategy of enriched DPP](#)” section. The proposed approach is then validated by a simplified proof-of-concept case study in “[A simplified case study](#)” section. Finally, a brief discussion on the presented approach is given in “[Con-](#)

✉ Wei Ji  
weiji@kth.se

<sup>1</sup> Department of Production Engineering, KTH Royal Institute of Technology, Stockholm, Sweden

<sup>2</sup> Department of Mechanical Engineering, Harbin University of Science and Technology, Harbin, China

clusions” section, together with the contributions and future work of this research.

## Related work

Nowadays, machine tool selection is considered a minor issue in process planning due to limited number of machine tools on a given shop floor; however, in cloud manufacturing environments, there are numerous machine tools that can be selected. For cutting tool selection and machining conditions determination, two common approaches exist: (1) in most of reported process planning methods, cutting tool is regarded as a standard machining resource and its parametrical optimisation is not considered, and (2) machining conditions are optimised after tool selection (Xu et al. 2011). In this case, the three simultaneous decision processes in process planning are treated sequentially, hindering the loss of both machining accuracy and efficiency.

Machining process optimisation started with mathematical model based methods that were popular in the 1990s. Chua et al. (1991) proposed a series of mathematical formulations to optimise the cutting conditions and to reduce the operation time. Yeo (1995) developed a multi-pass optimisation method for a CNC lathe, in which near-optimal solutions were obtained. Akturk and Avci (1996) presented a hierarchical method for a CNC machine tool. In their method, the mathematical models on system characterisation were established to minimise the total production cost. Experimental methods were applied with specific aims. Chen et al. (2011) proposed an experimental plan of a four-factor D-optimal design to obtain the optimal spindle speed, feed rate, cutting depth, and the status of lubrication concerning vibration and surface roughness in precision turning. Fernández-Valdivielso et al. (2015) presented an experiment based method to seek common feature of cutting tool with best performance in machining of superalloys in terms of surface integrity.

Machining optimisation, regarded as a search problem, was treated based on search algorithms in which tabu search (TS) is a popular one. Bretthauer and Cote (1997) presented a nonlinear programming method for capacity planning in a manufacturing system to determine the timing and size of capacity change. In their method, branch and bound, and outer approximation techniques were applied. Taiber (1996) proposed a set of modified algorithms from the field of combinatorial search problems, gradient projection method named as von Rosen, branch and bound algorithm, and shortest common super sequence algorithm, etc. The method was to assist the human planner in fulfilling machine tool and cutter selection, determination of the setup and process sequence, definition of tool paths and optimisation of cutting parameters. TS was applied to process planning, machining resource

selection, setup plan determination and operation sequencing (Li et al. 2003).

Genetic algorithm (GA), the most basic evolutionary algorithm, is a popular method used for optimisation. Chen and Tseng (1998) introduced a float encoding GA into machining conditions selection. Morad and Zalzal (1999) applied a GA to minimise the makespan, the total rejects and the total cost of production. Dereli and Filiz (2000) utilised a GA to obtain the optimal index positions on tool magazines. Hua et al. (2007) proposed a GA-based synthesis approach to archive machining scheme selection and operation sequencing optimisation for prismatic parts. Kondayya and Krishna (2012) used a non-dominated sorting genetic algorithm-II (NSGA-II) to optimise the cutting parameters during a CNC end-milling process. Manupati et al. (2016) proposed a modified block-based GA and modified NSGA to obtain the minimisation of both makespan and the variation of workload. A series of swarm intelligence (SI) based optimisation algorithms were applied to process planning. Guo et al. (2009) used a particle swarm optimisation (PSO) approach to obtaining operation sequence. Wen et al. (2014) proposed a honey bees mating based optimisation algorithm to optimise the process planning problems. In addition, multi-objective optimisation was employed due to many limitations of single-objective optimisation methods in the real machining process. Sardinias et al. (2006) proposed a GA-based multi-objective optimisation method to obtain the optimal cutting parameters during the turning process. Shin et al. (2011) introduced a multi-objective symbiotic evolutionary algorithm into flexible manufacturing system for solving process planning problems, where machine tool, sequence, and process are the three objectives. Li et al. (2018) proposed a multi-objective optimisation approach for tool path planning in freeform surface milling.

Expert system was developed to utilise the machining knowledge. Data classification methods were applied to optimising machining process. Data classification methods, e.g. decision trees and artificial neural networks (ANN), were also applied. Sluga et al. (1998) developed a decision tree based method to predict tool features, cutting geometry and cutting parameters a set of attribute values to improve and automate the tool selection and determination of cutting parameters. Thimm et al. (2001) proposed a datum hierarchy tree within graph theoretical approach to minimising machine and datum changes. Li et al. (2015) proposed a back propagation neural network model to predict the cutting parameters based on a set of mathematical objectives, e.g. machining time, energy consumption and surface roughness. Process planning was commonly treated as an NP-hard problem. Lian et al. (2012) applied an imperialist competitive algorithm (ICA) to find promising solutions with a reasonable computational cost. Their cases illustrated that the ICA was more efficient and robust than GA, SA, TS and PSO.

The hybrid methods were applied to relaxing the limitation of one single algorithm in the past decade, such as GA based, combined GA and PSO based. Wong et al. (2003) proposed a fuzzy expert system and GA to sequence machining process. Tiwari et al. (2006) presented an ant colony optimisation method to select the best process plan in an automated manufacturing environment. Venkatesan et al. (2009) developed a GA-based optimisation of weights applied to ANN for obtaining the best machining operation regarding marginal amount of time saving. Jayabal and Natarajan (2010) proposed a method of optimisation of thrust force, torque, and tool wear in drilling of coir fibre-reinforced composites, combining Nelder–Mead and GA methods. In their method, a nonlinear regression analysis was applied to establishing functions according to experimental data. A hybrid GA and intelligent search method was proposed by Salehi and Bahreininejad (2011), and it was applied to optimising machine tool, cutting tool and tool access direction for each operation. Pour (2018) proposed a hybrid algorithm based on time series analysis and wavelet transform to model surface roughness. Moreover, many efforts were also devoted to GA-based hybrid methods for optimisation of machining process. Petrovic et al. (2016) utilised PSO algorithm and chaos theory to optimise process plans, in which PSO was used in early stages of the optimisation process by implementing ten different chaotic maps that enlarged the search space and provided diversity. In addition, Rowe et al. (1996) reported an application of artificial intelligence in CNC grinding, including knowledge based and expert systems, fuzzy logic systems, and neural network systems. Within the context, setup time, process proving time and the extent of operator intervention could be improved. Arnaiz-González et al. (2016) used artificial neural networks to predict dimensional error on inclined surfaces machined by ball end mill. Their results showed that radial basis functions can predict better than multilayer perceptron in all cases.

From the literature, the optimisations in machining have been focused on the specialised cases and processes due to the limitations of traditional physical and experimental based methods in terms of high-dimensional data optimisation and high-accuracy optimisation. They lack the generality or capability for general process planning, which has recently caught attentions to the application of big data analytics in this area. Moreover, the major algorithms consist of mathematical methods, experimental methods, search methods, evolution based methods, and SI based methods. Nowadays, hybrid methods and big data analytics used in many areas are promising to provide high-accuracy solution strategies, in particular, for optimisation problems of high-dimensional data. Big data embraced by smart manufacturing (Kusiak 2017), as well as data collection, use and sharing are important gaps in innovation. Tao et al. (2018) shed light in a data-driven smart manufacturing framework. Recently, there

have been many articles reporting big data analytics in machining. Li et al. (2017) presented a data-driven approach combined with Deep Belief Network (DBN) to predicting the backlash error in machining centre. Liang et al. (2018) proposed a novel Cyber Physical System (CPS) and big data enabled machining optimisation system to optimise the energy in machining processes. Xu and Duan (2018) pointed out that CPS and big data are two keys for Industry 4.0 in the near future. Moreover, as reported by de Lacalle et al. (2006), their developed methods were used to detect potential milling problems associated with cutting force measurement, which demonstrated that the data in machining are abundantly enough to be used and mined. Therefore, big data analytics combined with hybrid algorithms shows potential for an integrated optimisation of machine tools, cutting tools and machining conditions. In the previous work, machining was considered for both scheduling by big data analytics (Ji and Wang 2017a) and machining optimisation by a proposed enriched process planning method in the conceptual level (Ji and Wang 2017b), which are limited by workpiece representation. To address the issue, an enriched DPP is proposed in this research, where the original DPP is selected to carry out the workpiece representation etc. and the machining process is considered as a whole to optimise the machining resources.

## Enriched distributed process planning (DPP)

### Concept of distributed process planning

A detailed system architecture of DPP is shown in Fig. 1, where supervisory planning, execution control and operation planning are the three major system modules. In this design, the execution control module is placed in-between the supervisory planning and operation planning modules, and looks after jobs dispatching (in the unit of setups) based on up-to-date monitoring data, availability of machines and scheduling decisions (Wang and Shen 2003; Wang et al. 2003). The distribution profile is a key feature. Combining with web-based knowledge sharing, dynamic scheduling, real-time monitoring and remote control, DPP can be embedded into web-based environment, which is named Web-DPP (Wang 2009, 2013). Towards the concept of cloud manufacturing, a Cloud-DPP was also developed as one of the applications of cyber-physical systems for more complex manufacturing environment (Wang 2014).

Within DPP, feature-based design and machining feature recognition are beyond the scope of this research. DPP assumes that machining features are already made available in the product data—they are either generated directly by using a feature-based design tool or recognised by a third-party machining feature recognition solution. Execution Control and Operation Planning in the original DPP

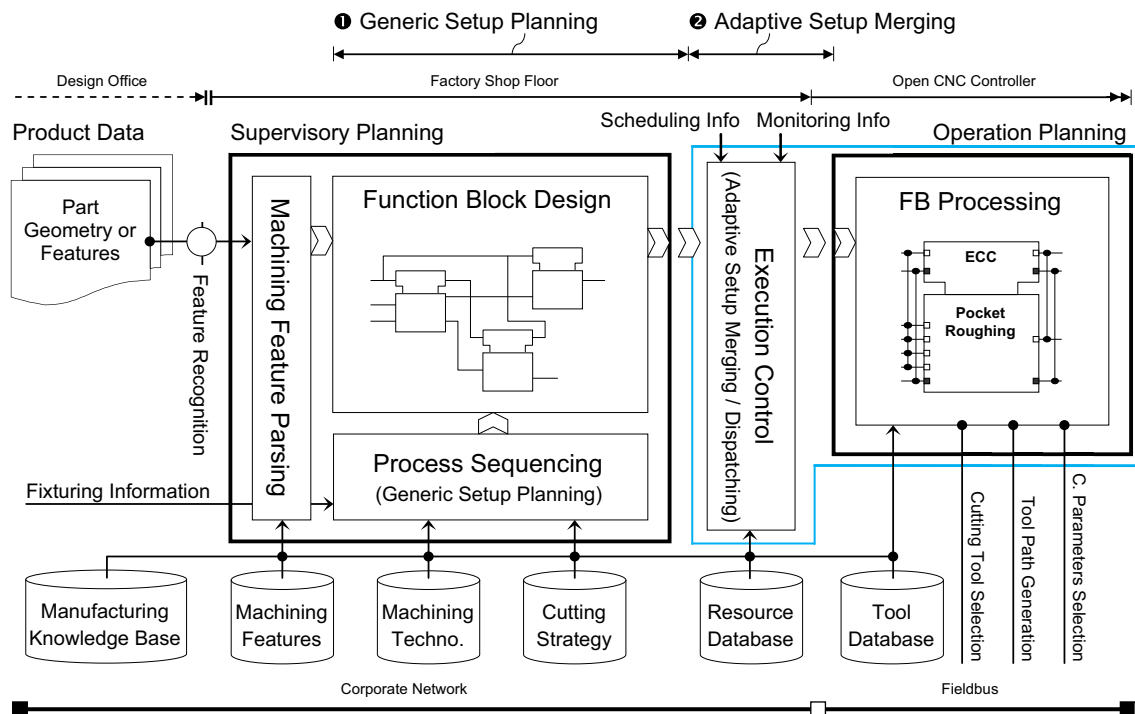


Fig. 1 System architecture of DPP (Wang et al. 2003)

do not consider the global machining process optimisation due to the complexity of relevant machining resources, i.e. machine tool, cutting tool and machining conditions. Targeting the issue, this paper proposes an enriched DPP combining DPP with big data analytics. Within the context, the suitable machine tool can be selected in Execution Control, and the suitable cutting tool and machining conditions will be chosen and optimised accordingly.

**Enriched DPP**

Machine tool selection, cutting tool selection and machining conditions determination are the major decision processes after Supervisory Planning in the original DPP (Fig. 1). They are treated separately by using the existing methods (experimental and physical), as shown in the left side of Fig. 2, where their relationships are ignored, resulting in the feasible sets of those elements being bounded. Targeting this issue, the proposed method is to address the whole process from customer orders to final parts, and to develop a generalised system in which those processes are merged, as shown in the right side of Fig. 2. The optimisation can therefore be treated as an integrated one after setting workpiece and machining requirements (“Task” in Fig. 2) towards a global optimal solution. Generally, there is no strict decision sequence between machine tool selection, cutting tool selection and machining conditions determination, and they are decided according to their availabilities.

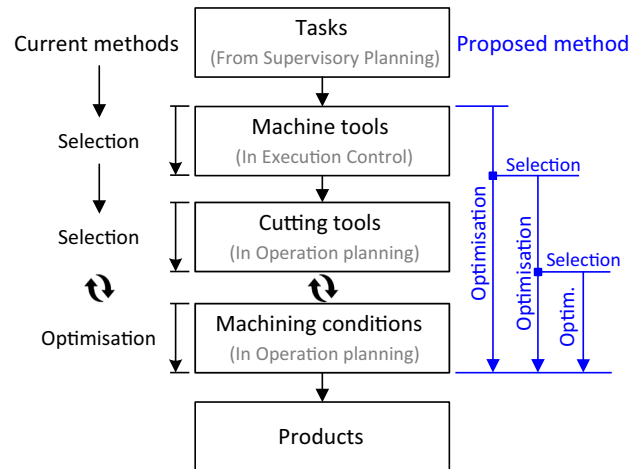
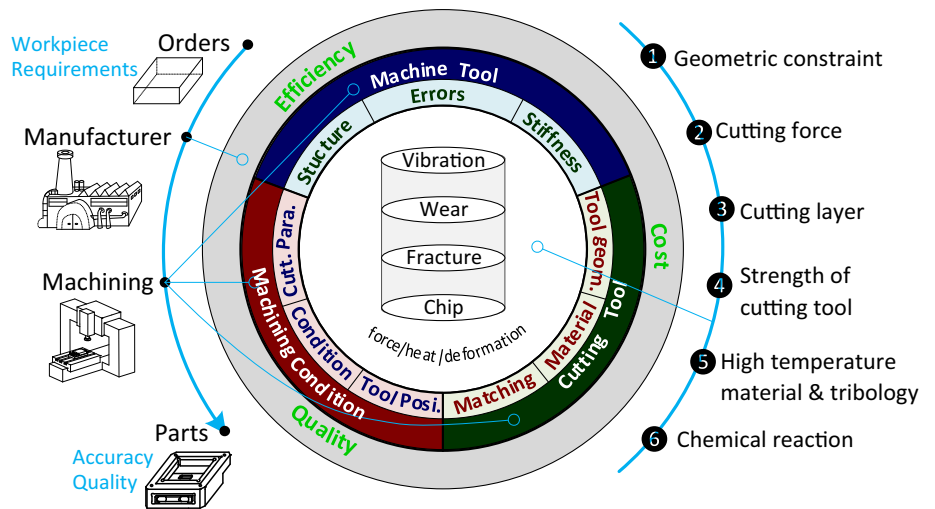


Fig. 2 Differences in cutting tool and machining conditions optimisation

**Big data in machining**

Big data are characterised by big variety, big volume and big velocity. In the manufacturing industry, data are generated along the lifecycle of products when orders are placed to the selected manufacturers, produced by machine tools, and delivered to customers, as shown in Fig. 3. A manufacturer concerns machining efficiency, cost and quality, whereas in machining, the right decisions of machine tool, machining conditions and cutting tools are the major objectives. The

**Fig. 3** Relationship between cutting tool and machining conditions

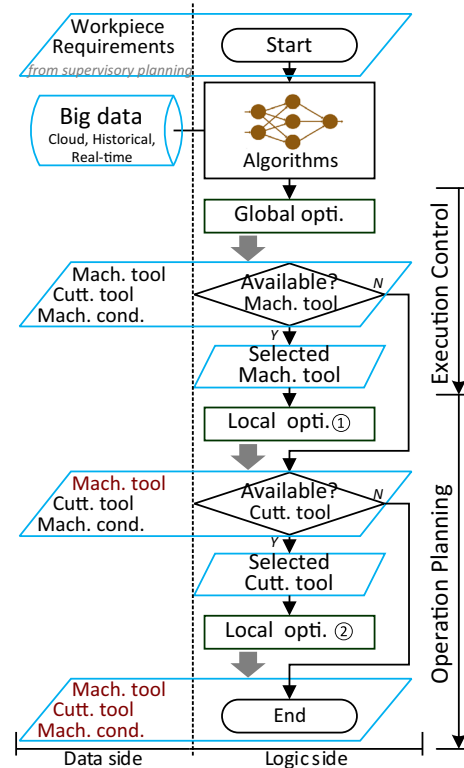


relationships between the machining related elements are also complex and include physical and chemical interactions, as shown in ①–⑥ of Fig. 3, and they refer to geometric constraint ①, cutting force ②, cutting layer ③, strength of cutting tool ④, high temperature material and tribology ⑤, and chemical process ⑥. Many parameters are needed to represent these elements. For example, the order is divided into workpiece and requirements. Machine tools are related to their structures, errors and stiffness. Cutting tool includes tool geometries, tool material, and the match between the geometry and the material. Machining conditions include cutting parameters, coolant types, tool position, and fixture properties. Performance factors related to force, heat and deformation during machining consist of chip control, machining dynamics, tool wear, and fracture. The final parts are also inspected in terms of accuracy and quality.

Considering the machining complexity, data in the process are of high dimensions. Usually, big data volume is required to guarantee the training accuracy. Machining processes are closely linked to the states of cutting tool and machine tool. Real-time data collection is needed for monitoring the machining processes, and regular data update is also essential for correcting the machine tool state.

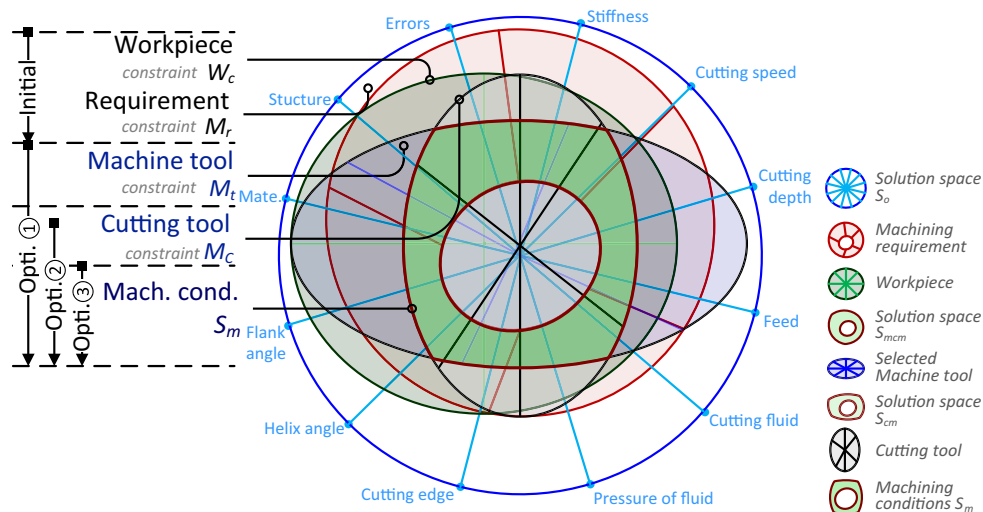
**Big data analytics based enriched DPP approach**

Based on the analysis in “Big data in machining” section, it is extremely difficult, if not impossible, to optimise the entire process from order to final part using the conventional methods; however, the process can be represented by data attributes. Along the direction, this research proposes a big data analytics based enriched DPP approach to addressing the optimisation process as a whole, where machine tool, cutting tool, and machining conditions can be decided altogether, as shown in Fig. 4. The optimisation is based on certain machining geometrical elements, such as machining features. The



**Fig. 4** Decision making framework of enriched DPP

workpiece and machining requirements are the initial input data, and DPP adopts a three-step optimisation scheme: (1) to globally optimise machine tool, cutting tool, and machining conditions; (2) to locally obtain the optimal cutting tool and machining conditions after selecting a machine tool; and (3) to locally optimise machining conditions after selecting a cutting tool. In general, the availability of machine tool is the lowest, followed by cutting tool or machining conditions; therefore, a machine is selected ahead of cut-



**Fig. 5** Optimisations of machining process and their solution spaces

ting tool selection and machining conditions determination. Within the context, each machining resource is represented by high-dimensional data attributes, which enables the patterns between the resources to be mined by the machine learning algorithms. As a result, new materials, new machine tools and/or new cutting tools can be optimised by providing the basic data regardless of the relevant machining experiments of those new machining resources.

## Solution strategy of enriched DPP

### Problem transformation in enriched DPP

The machining process is transformed into a statistic problem in the enriched DPP. The available parameters of machine tool, cutting tool and machining conditions constitute the original solution space  $S_o$  with multi-dimensional data, as shown in Fig. 5. There are three steps in optimisation in the enriched DPP, i.e., initialisation, optimisations ①, ② and ③. Initialisation specifies the workpiece and machining requirements which are two constraints to the original solution space, i.e. workpiece constraint  $W_c$  and machining requirement  $M_r$ . Current solution space  $S_{mcm}$  can be calculated by  $S_{mcm} = S_o \cap W_c \cap M_r$ . Then the global optimisation ① is carried out within space  $S_1$ , as a result of which machine tool, cutting tool and machining conditions are determined accordingly. The results provide a reference for machine tool selection. Once the optimal machine tool is chosen, cutting tool and the machining conditions are decided simultaneously. On the contrary, a substitute machine tool as a constraint  $M_t$  bounds solution space  $S_{mcm}$ . Then, the current feasible set  $S_{cm}$ ,  $S_{cm} = S_{mcm} \cap M_t$ , defines the updated solution space for execution of optimisation ② for selecting

cutting tool and machining conditions. Similar to the selection process of machine tool, the selected cutting tool, as a constraint  $M_c$ , bounds the feasible set  $S_{cm}$ , resulting in the machining conditions solution space  $S_m$ ,  $S_m = S_{cm} \cap M_c$ .

### A hybrid algorithm for enriched DPP

Hybrid algorithms can enhance optimisation performance by relaxing the limitations of each single algorithm. This paper proposes an optimisation method combining three types of algorithms, i.e. evolutionary algorithm (EA) or SI based optimisation, ANN based model trained by big data, and analytic hierarchy process (AHP) based weight decision. Within the context, EA is a generic population-based metaheuristic optimisation algorithm, and SI is a generic population-based metaheuristic optimisation algorithm. A global optimisation is carried out by EA or SI, and these algorithms refer to several steps, e.g. parameter selection (sample), operation, and criterion. An ANN is an interconnected group of nodes, akin to the vast network of neurons in a brain. Its model trained by big data through supervised learning is employed to obtain individual objective fitness with high accuracy, which plays a key role as oppose to the existing methods. The ANNs, trained by machining big data, are adopted to achieve good results for such high-dimension data attributes, and they are considered as a set of “equations” to derive solutions for the complicated optimisation problem. AHP is based on mathematics and psychology, and it helps decision makers find one that best suits their goal, and it was used to decide the weights in cutting tool evaluation in our previous works (Ji et al. 2017). Combining expert questionnaire, it is applied to calculating weights for multi-objective optimisation, and its pairwise comparison matrix is established by big data ana-

lytics in statistics level, in case that the higher accuracy can be obtained by comparing with the questionnaires.

## A simplified case study

A simplified case study is chosen to show the process of the proposed method as a proof of concept due to the unavailability of the real relevant data. A set of hypothetical data generated in-house, according to the real machining setups, are applied to training the DBN (Hinton 2009) model used to calculate the fitness of GA.

### Data attributes

The attributes of big data include the information regarding workpiece, machining requirements, machine tools, cutting tools, machining conditions, and machining results. Also, the details of the information cover all the factors affecting the machining process, as shown in Table 1, and they include:

- (1) Data attributes of workpiece: workpiece information includes the quantity of parts, part geometry, and part materials, etc. Here, the structural part geometry and quantity are represented by machining features and their quantity, whereas the curved or freeform parts are described by tool paths. Part materials refer to the material profiles with respect to the machinability of materials, e.g. hardness, brittleness, toughness and plasticity, etc., so that each material can be represented by a series of parameters.
- (2) Data attributes of machining requirements: machining requirements refer to the designed machining quality in terms of tolerances and surface integrity, e.g. dimensional tolerance, geometrical tolerance, roughness, white layer, and dark layer.
- (3) Data attributes of machine tool: machine tool information usually involves the number of machine tool, machine type, machine structure, structure and power of spindle, linear axis and rotational axis, errors of each axis, energy consumption of each component, and other related parameters. In this way, each machine tool can be represented by a set of parameters.
- (4) Data attributes of cutting tool: the cutting tool involves tool type, tool material, tool coating, and tool geometrical parameters.
- (5) Data attributes of machining conditions: the machining conditions includes cutting parameters, cutting coolant and its parameters, as well as fixture and its parameters.
- (6) Data attributes of machining results: machining results indicate the machining quality in terms of geometrical errors and surface integrity against the machining requirements, and also include the physical data of the cutting process consisting of cutting force, cutting vibration, etc. Such data are obtained through inspections of machined parts, both qualified and unqualified, and through sensor-based monitoring during machining process.

### Data generation

The attributes of the hypothetical data include:

- (1) Workpiece consists of machining feature and material information. Machining features includes 16 types, and according to the properties of the commonly used alloy steel, material information refers to Poisson's ratio (0.25, 0.35), Young's modulus (100, 500 GPa), elastic modulus (120, 240 GPa), shear modulus (40, 80 GPa), density (7, 9 g/cm<sup>3</sup>), and hardness (150, 300 HB);
- (2) Based on the requirements of dimension and surface integrity in milling operation, machining requirements refer to dimensional tolerance (IT5, IT10), surface roughness Ra (0.4, 6.3 μm), morphology Sa (1,5 μm), white layer depth (0.3, 10 μm), dark layer depth (1, 20 μm), hardened layer depth (1, 7 μm), and grain deformation (1, 15 μm);
- (3) Multi-axis CNC machine tools are the major equipment to perform machining. According to the commonly used types, machine tool data include machine tool structure (7 types), linear axis errors (1, 10 μm), rotational axis errors (2, 8''), and maximum spindle power (3, 20 KW);
- (4) Carbide cutting tools are selected since they are most popularly used in machining. Based on the available tool information, cutting tool data refer to tool radius (4, 6 mm), tool length (60, 100 mm), entrance angle (45, 90°), rake angle (0, 30°), flank angle (6, 20°), matrix material (11 types) and coat layer (8 types);
- (5) Except machining tool and cutting tool, fixture, coolant and cutting parameters are crucial in machining. To deal with those factors, machining conditions data are used, including cutting parameters, i.e. cutting speed (30, 150 m/min), feed rate (0.05, 0.5 mm/z), cutting depth (0.2, 0.5 mm) and cutting width (0.5, 1 diameter of cutting tool), the used coolant referring to fluid (6 types) and pressure (0, 20 bar), and the fixture involving constrained DOF (1 or 2), locator location errors (0.01, 0.1 mm) and locator installation error (0.01, 0.1 mm);
- (6) Machining results are obtained in form of dimension and surface integrity against the machining requirements. In this paper, surface roughness and white layer depth are the relevant indexes as an example to show how the proposed method works.

Table 1 A sample data sheet

No.	MFs	Volume/MF	Possion's rate	Young's modulus	Elastic modulus	Shear modulus	Density	Hardness
<i>Workpiece</i>								
1	Side	100	0.3	500	690	100	7.9	35 HRC
2								
No.	Tolerance/mm	Roughness $R_a$	Morpho.	White layer	Dark layer	Hardened layer	Grain deformation	
<i>Machining requirement</i>								
1	$\pm 0.01$	0.8 $\mu\text{m}$	–	5 $\mu\text{m}$	20 $\mu\text{m}$	15 $\mu\text{m}$	12 $\mu\text{m}$	
2								
No.	No.	Type	Structure	Errors in XYZ/ $\mu\text{m}$	Errors in ABC/ $^\circ$	Power in XYZ/kw	Spindle power	
<i>Machine tool</i>								
1	M1	4	XYZB	5/3/3	-0.354/-	2.8/3.0/2.5	26 kw	
2								
No.	Tool radius	Tool length	Entrance angle	Rake angle	Flank angle	Matrice	Coat	
<i>Cutting tool</i>								
1	30 mm	50 mm	45 $^\circ$	10 $^\circ$	10 $^\circ$	K2	TiCN	
2								
No.	Cutt. speed	Feed rate	Cutting depth	Cutting width	Pressure	Constrained DOF	Locators	
<i>Machining condition</i>								
1	50 m/min	0.15 mm/z	0.2 mm	15 mm	1	20 bar	XY	
2							X	
No.	Accuracy	Roughness $R_a$	Morphology	White layer	Dark layer	Hardened layer	Cutting force in XYZ/N	
<i>Machining results</i>								
1	0.007 mm	0.79 $\mu\text{m}$	–	4.5 $\mu\text{m}$	17.8 $\mu\text{m}$	12.6 $\mu\text{m}$	420.4/234.5/100.5	
2								



Based on the above data attributes and their ranges, 60,000 data were generated randomly by using Matlab®. The selected machining results, surface roughness and white layer depth, are labeled for training the DBN model.

### A detailed algorithm

Figure 6 depicts a hybrid algorithm of DBN and GA, where the DBN model trained by big data is employed to calculate the fitness, and GA is applied to performing the optimisation computation. DBN, a generative graphical model or alternatively a class of deep neural network, and composed of multiple layers of latent variables with connections between the layers but not between units within each layer, was developed as a solution to the problem of training multi-layer perceptions. The workpiece and machining requirements are set first, and (S1) their parameters combining the target factors are divided into two classifications: (1) proceeded to GA calculation processes, and (2) passed to the trained model. (S2) The former, as a population, is treated in the GA process, generating a new population. Then the latter filters the data which is used to train the DBN model. The population generated in S1 is decoded into a set of parameters which are imported to the trained model, obtaining a calculation fitness. (S3) The fitness is compared with the termination criterion, and once the computation is finished, the results provide a reference for machine tool selection, optimisations of cutting tools and machining conditions.

If the selected machine tool is not the optimal one, the parameters of the machine tool are sent back to the parameter classification process, and then the steps (S1), (S2) and (S3) are performed again until the optimised results of cutting tools and machining conditions can be obtained.

### Optimisation processes and results

A batch of parts is planned to be produced by a manufacturer, and its material parameters consist of 0.3 Poisson’s ratio, 250 GPa Young’s modulus, 200 GPa elastic modulus, 50 GPa shear modulus, 8 g/cm<sup>3</sup> density and 200 HB hardness, and its machining requirements involve Ra 1.6 surface roughness and 7 μm white layer depth; therefore the optimisation objectives are surface roughness and write layer depth, where their weights are 0.8 and 0.2, respectively.

The hypothetic data filtered by the machining requirements filter, are applied to training the DBN model. The DBN model is trained using a 35 × 100 × 100 × 100 × 2 neural network, in 100 training cycles, and a 0.2 learning rate that can be changed according to the training error. In this case, Matlab® is used for the computation in a computer with an Intel® Core™ CPU i7-3630QM 2.4 GHz, 8 GB memory, and graphics processing unit (GPU) NVIDIA GeForce GT 750 M. GPU is used to train the model in about 90 s. As

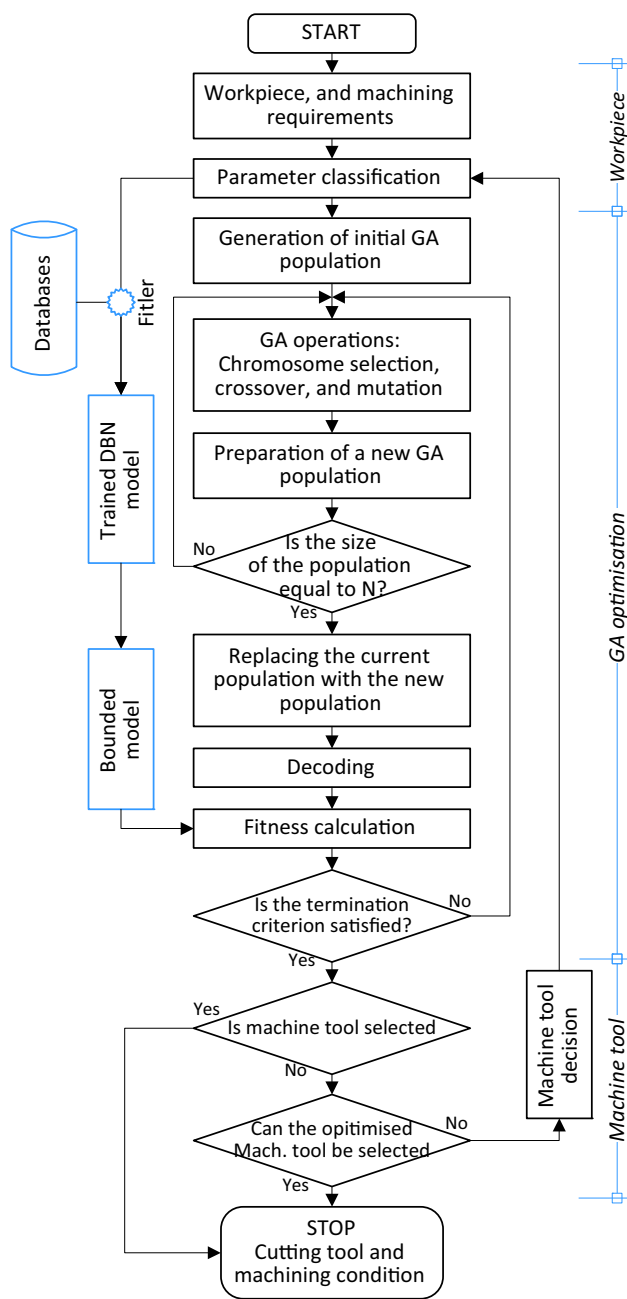
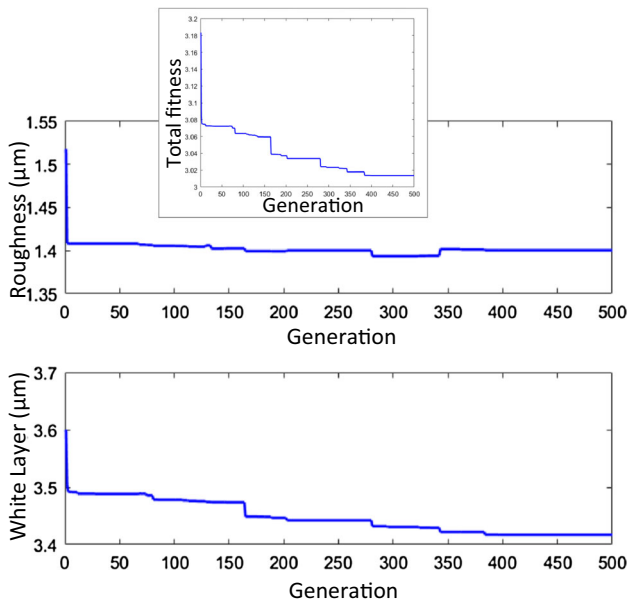


Fig. 6 A hybrid algorithm combining GA and DBN

a result of the training, the training error and test error of surface roughness are stabilised around 0.0525 and 0.075, respectively, and for white layer, the training error and test error are 0.0529 and 0.0924, respectively. Since the data are hypothetic, there is no need to carry out the questionnaire task for AHP, and therefore, both optimisations of the two objectives are performed together.

Consequently, the GA computation, 3–5 s each time, is performed to optimise machine tools, cutting tools and cutting parameters. The parameters of GA consist of (1) generations: 500, (2) population size: 100, (3) crossover probability:

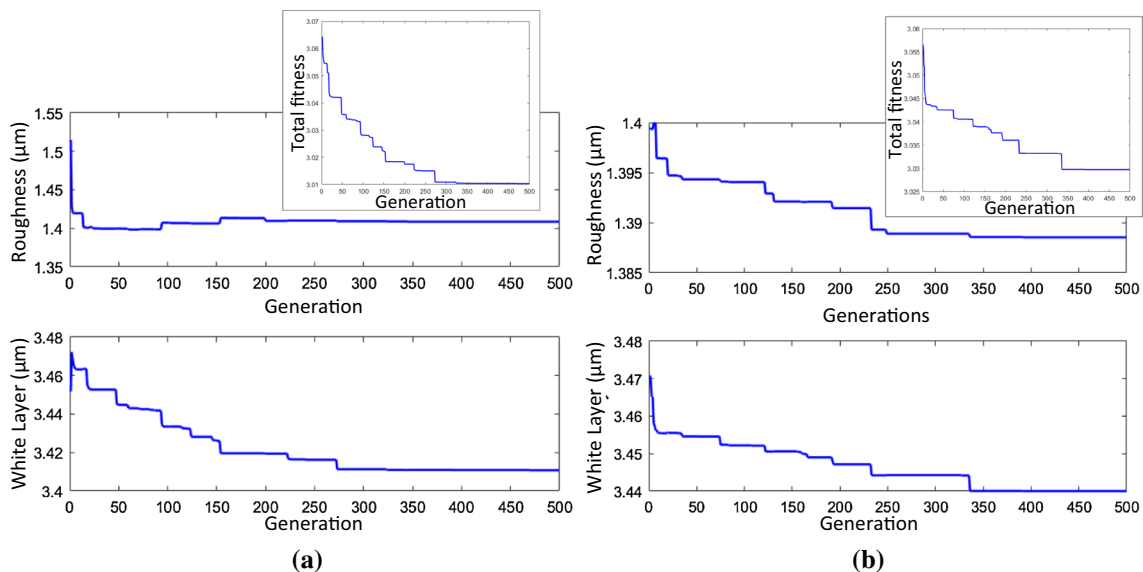


**Fig. 7** Calculation results including machine tool, cutting tool and cutting parameters

0.8, (4) mutation probability: 0.01, and (5) computation accuracy: 0.001. The fitness of machining requirements is stable after approximately 385 generations, as shown in Fig. 7, together with the optimised parameters of machine tool, cutting tool and cutting parameters: (1) machine tool:  $x$ ,  $y$ ,  $z$  and  $c$  structure with axis errors of  $1.003 \mu\text{m}$ ,  $5.50 \text{ mm}$ ,  $1.212 \text{ mm}$ , and  $2.0''$ , respectively, and approximately 12 KW spindle power; (2) cutting tool involves 4 mm tool radius, 60.00 mm tool length,  $45.176^\circ$  entrance angle,  $1.057^\circ$  rake angle and  $6.0^\circ$  flank angle, YG3 matrix material and

TiC-TiCN-Al<sub>2</sub>O<sub>3</sub>-TiN coat; and (3) machining conditions include 149.998 m/min cutting speed, 0.298 mm/z feed rate, 0.462 mm cutting depth, 75%-tool-diameter cutting width, oil based coolant with 4.216 Pa pressure, 1 constrained DOF, 0.10 mm locator location error and 0.1 mm locator installation error.

The optimised machine tool (if available), cutting tool geometries and cutting parameters are determined simultaneously. In case of no optimal machine tool, an alternative machine is selected, and its parameters are imported to GA. A machine tool, for instance, is equipped with  $10 \mu\text{m}$  error of  $x$ ,  $1 \mu\text{m}$  errors of  $y$  and  $z$ ,  $0.13''$  and  $0.62''$  errors of  $b$  and  $c$ , and a 3 KW maximum spindle power. Based on the selected machine tool and the initial workpiece parameters, the GA calculation is performed again. Figure 8a depicts the fitness of machining requirements after approximate 275 generations. Then the optimised results of tools and cutting parameters consist of: (1) cutting tool: 4 mm tool radius, 60.00 mm tool length,  $45.0^\circ$  entrance angle,  $0.242^\circ$  rake angle,  $6.002^\circ$  flank angle, YG3 matrix material and TiC-Al<sub>2</sub>O<sub>3</sub> coat, and (2) machining conditions: 149.973 m/min cutting speed, 0.30 mm/z feed rate, 0.50 mm cutting depth, 100% cutting width, oil-based coolant with 4.216 Pa pressure, 1 constrained DOF, 0.089 mm location error, and 0.1 mm installation error of locator. Similarly, a cutting tool, not the optimised one, was selected, and its parameters are 4 mm tool radius, 60 mm tool length,  $45^\circ$  entrance angle,  $7.5^\circ$  rake angle,  $9.7^\circ$  flank angle, YG3 matrix material and TiC-TiCN-TiC-TiCN-TiN coat. After setting the selected cutting tool, the third optimisation is carry out and its fitness is stabilised after 340 generations, as shown in Fig. 8b. The results of the optimised machining conditions are 149.991 m/min



**Fig. 8** Calculation results including cutting tool and cutting parameters. **a** Optimisation after machine tool selection, **b** optimisation after cutting tool selection

cutting speed, 0.300 mm/z feed rate, 0.50 mm cutting depth, 100%-tool-diameter cutting width, oil-based coolant with 4.216 Pa pressure, 1 constrained DOF, 0.1 mm location error, and 0.1 mm installation error of locator.

## Conclusions

Apart from the theory and empirical research to address rather specific issues, the proposed method is to solve a machining optimisation issue from the big data's perspective, where all machining resources are considered as a whole. It is also important to show the potential of the proposed method in real-world applications and encourage companies to fully utilise the benefits of big data. Although not possible to use industrial big data, the proposed method is validated based on reasonable hypothesis.

Different from the conventional optimisation methods in which machine tool, cutting tool and machining conditions are decided separately, this paper proposed a concept of big data analytics based enriched DPP combining the optimisations of the three as a whole. Within the context, each machining resource is represented by its data attributes. Once the data are ready, the patterns/relationships among the machining resources are mined with the help of the big data analytics algorithms, e.g. ANN, and they are used to optimise the relevant resources. Workpiece is regarded as the first constraint, and parts of machining requirements are regarded as optimised objectives, the relationship of which is handled by AHP. Based on the provided fitness by combining the mined pattern and AHP applied objective, an EA or SI algorithm is employed to obtain the optimal or near-optimal machining resources for selecting machine tool, cutting tool and machining conditions according to their availabilities. Each selected resource is regarded as a constraint bounding the solution space. This approach is validated by a simplified case study in which a GA and DBN hybrid algorithm is implemented. The main contributions of this work include:

- Transformation from real-world machining problems to statistic problems, allowing optimising machining process from the data analytic perspective;
- Extension of a feasible set by integrated optimisation of machine tool selection, cutting tool selection and machining conditions determination, making machining solutions more accurate compared with the conventional methods;
- Enhancing the functionality of the original DPP.

Nevertheless, this research is still at its early stage with limitations. Our future work is therefore planned as follows: (1) detailed machining resource representation; (2) data collection from real-world machining environment with challenges in data filtering, cleansing and storage; and (3)

real-time decision algorithms for efficient big data analytics. The result of these will be reported separately in the future.

**Acknowledgements** This project is supported by National Natural Science Foundation of China (Grant No. 51605121).

**Open Access** This article is distributed under the terms of the Creative Commons Attribution 4.0 International License (<http://creativecommons.org/licenses/by/4.0/>), which permits unrestricted use, distribution, and reproduction in any medium, provided you give appropriate credit to the original author(s) and the source, provide a link to the Creative Commons license, and indicate if changes were made.

## References

- Akturk, M. S., & Avci, S. (1996). An integrated process planning approach for CNC machine tools. *International Journal of Advanced Manufacturing Technology*, 12(3), 221–229. <https://doi.org/10.1007/BF01351201>.
- Arnaiz-González, Á., Fernández-Valdivielso, A., Bustillo, A., & López de Lacalle, L. N. (2016). Using artificial neural networks for the prediction of dimensional error on inclined surfaces manufactured by ball-end milling. *The International Journal of Advanced Manufacturing Technology*, 83(5), 847–859. <https://doi.org/10.1007/s00170-015-7543-y>.
- Bretthauer, K. M., & Cote, M. J. (1997). Nonlinear programming for multiperiod capacity planning in a manufacturing system. *European Journal of Operational Research*, 96(1), 167–179. [https://doi.org/10.1016/S0377-2217\(96\)00061-6](https://doi.org/10.1016/S0377-2217(96)00061-6).
- Chen, C.-C., Chiang, K.-T., Chou, C.-C., & Liao, Y.-C. (2011). The use of D-optimal design for modeling and analyzing the vibration and surface roughness in the precision turning with a diamond cutting tool. *International Journal of Advanced Manufacturing Technology*, 54(5–8), 465–478. <https://doi.org/10.1007/s00170-010-2964-0>.
- Chen, M. C., & Tseng, H. Y. (1998). Machining parameters selection for stock removal turning in process plans using a float encoding genetic algorithm. *Journal of the Chinese Institute of Engineers*, 21(4), 493–506. <https://doi.org/10.1080/02533839.1998.9670412>.
- Chua, M. S., Loh, H. T., & Wong, Y. S. (1991). Optimization of cutting conditions for multi-pass turning operations using sequential quadratic programming. *Journal of Materials Processing Technology*, 28(1–2), 253–262. [https://doi.org/10.1016/0924-0136\(91\)90224-3](https://doi.org/10.1016/0924-0136(91)90224-3).
- de Lacalle, L. N. L., Lamikiz, A., Sánchez, J. A., & de Bustos, I. F. (2006). Recording of real cutting forces along the milling of long complex parts. *Mechatronics*, 16(1), 21–32. <https://doi.org/10.16/j.mechatronics.2005.09.001>.
- Dereli, T., & Filiz, I. H. (2000). Allocating optimal index positions on tool magazines using genetic algorithms. *Robotics and Autonomous Systems*, 33(2–3), 155–167. [https://doi.org/10.1016/S0921-8890\(00\)00086-5](https://doi.org/10.1016/S0921-8890(00)00086-5).
- Fernández-Valdivielso, A., López de Lacalle, L. N., Urbikain, G., & Rodriguez, A. (2015). Detecting the key geometrical features and grades of carbide inserts for the turning of nickel-based alloys concerning surface integrity. *Proceedings of the Institution of Mechanical Engineers, Part C: Journal of Mechanical Engineering Science*, 230(20), 3725–3742. <https://doi.org/10.1177/0954406215616145>.
- Guo, Y. W., Mileham, A. R., Owen, G. W., Maropoulos, P. G., & Li, W. D. (2009). Operation sequencing optimization for five-axis prismatic parts using a particle swarm optimization approach.

- Proceedings of the Institution of Mechanical Engineers Part B-Journal of Engineering Manufacture*, 223(5), 485–497. <https://doi.org/10.1243/09544054JEM1224>.
- Hinton, G. E. (2009). Deep belief networks. *Scholarpedia*, 4(5), 5947.
- Hua, G. R., Zhou, X. H., & Ruan, X. Y. (2007). GA-based synthesis approach for machining scheme selection and operation sequencing optimization for prismatic parts. *International Journal of Advanced Manufacturing Technology*, 33(5–6), 594–603. <https://doi.org/10.1007/s00170-006-0477-7>.
- Jayabal, S., & Natarajan, U. (2010). Optimization of thrust force, torque, and tool wear in drilling of coir fiber-reinforced composites using Nelder–Mead and genetic algorithm methods. *International Journal of Advanced Manufacturing Technology*, 51(1–4), 371–381. <https://doi.org/10.1007/s00170-010-2605-7>.
- Ji, W., Shi, J., Liu, X., Wang, L., & Liang, S. Y. (2017). A novel approach of tool wear evaluation. *Journal of Manufacturing Science and Engineering*, 139(September), 1–8. <https://doi.org/10.1115/1.4037231>.
- Ji, W., & Wang, L. (2017a). Big data analytics based fault prediction for shop floor scheduling. *Journal of Manufacturing Systems*, 43, 187–194. <https://doi.org/10.1016/j.jmsy.2017.03.008>.
- Ji, W., & Wang, L. (2017b). Big data analytics based optimisation for enriched process planning: A methodology. *Procedia CIRP*, 63, 161–166. <https://doi.org/10.1016/j.procir.2017.03.090>.
- Kondayya, D., & Krishna, A. G. (2012). An integrated evolutionary approach for modelling and optimisation of CNC end milling process. *International Journal of Computer Integrated Manufacturing*, 25(11), 1069–1084. <https://doi.org/10.1080/0951192X.2012.684718>.
- Kusiak, A. (2017). Smart manufacturing must embrace big data. *Nature*, 544(7648), 23–25. <https://doi.org/10.1038/544023a>.
- Li, L., Deng, X., Zhao, J., Zhao, F., & Sutherland, J. W. (2018). Multi-objective optimization of tool path considering efficiency, energy-saving and carbon-emission for free-form surface milling. *Journal of Cleaner Production*, 172, 3311–3322. <https://doi.org/10.1016/j.jclepro.2017.07.219>.
- Li, L., Liu, F., Chen, B., & Li, C. B. (2015). Multi-objective optimization of cutting parameters in sculptured parts machining based on neural network. *Journal of Intelligent Manufacturing*, 26(5), 891–898. <https://doi.org/10.1007/s10845-013-0809-z>.
- Li, W. D., Ong, S. K., Lu, Y. Q., Nee, A. Y. C., Palade, V., Howlett, R. J., et al. (2003). A Tabu search-based optimization approach for process planning. *Knowledge-Based Intelligent Information and Engineering Systems, Pt 2, Proceedings*, 2774, 1000–1007.
- Li, Z., Wang, Y., & Wang, K. (2017). A data-driven method based on deep belief networks for backlash error prediction in machining centers. *Journal of Intelligent Manufacturing*. <https://doi.org/10.1007/s10845-017-1380-9>.
- Lian, K. L., Zhang, C. Y., Shao, X. Y., & Gao, L. (2012). Optimization of process planning with various flexibilities using an imperialist competitive algorithm. *International Journal of Advanced Manufacturing Technology*, 59(5–8), 815–828. <https://doi.org/10.1007/s00170-011-3527-8>.
- Liang, Y. C., Lu, X., Li, W. D., & Wang, S. (2018). Cyber physical system and big data enabled energy efficient machining optimization. *Journal of Cleaner Production*, 187, 46–62. <https://doi.org/10.1016/j.jclepro.2018.03.149>.
- Manupati, V. K., Chang, P. C., & Tiwari, M. K. (2016). Intelligent search techniques for network-based manufacturing systems: multi-objective formulation and solutions. *International Journal of Computer Integrated Manufacturing*, 29(8), 850–869. <https://doi.org/10.1080/0951192X.2015.1099073>.
- Morad, N., & Zalzal, A. (1999). Genetic algorithms in integrated process planning and scheduling. *Journal of Intelligent Manufacturing*, 10(2), 169–179. <https://doi.org/10.1023/A:1008976720878>.
- Petrovic, M., Mitic, M., Vukovic, N., & Miljkovic, Z. (2016). Chaotic particle swarm optimization algorithm for flexible process planning. *International Journal of Advanced Manufacturing Technology*, 85(9–12), 2535–2555. <https://doi.org/10.1007/s00170-015-7991-4>.
- Pour, M. (2018). Determining surface roughness of machining process types using a hybrid algorithm based on time series analysis and wavelet transform. *The International Journal of Advanced Manufacturing Technology*. <https://doi.org/10.1007/s00170-018-2070-2>.
- Rowe, W. B., Li, Y., Mills, B., & Allanson, D. R. (1996). Application of intelligent CNC in grinding. *Computers in Industry*, 31(1), 45–60. [https://doi.org/10.1016/0166-3615\(96\)00036-X](https://doi.org/10.1016/0166-3615(96)00036-X).
- Salehi, M., & Bahreininejad, A. (2011). Optimization process planning using hybrid genetic algorithm and intelligent search for job shop machining. *Journal of Intelligent Manufacturing*, 22(4), 643–652. <https://doi.org/10.1007/s10845-010-0382-7>.
- Sardinas, R. Q., Santana, M. R., & Brindis, E. A. (2006). Genetic algorithm-based multi-objective optimization of cutting parameters in turning processes. *Engineering Applications of Artificial Intelligence*, 19(2), 127–133. <https://doi.org/10.1016/j.engappai.2005.06.007>.
- Shin, K. S., Park, J. O., & Kim, Y. K. (2011). Multi-objective FMS process planning with various flexibilities using a symbiotic evolutionary algorithm. *Computers & Operations Research*, 38(3), 702–712. <https://doi.org/10.1016/j.cor.2010.08.007>.
- Sluga, A., Jermol, M., Zupanic, D., & Mladenic, D. (1998). Machine learning approach to machinability analysis. *Computers in Industry*, 37(3), 185–196. [https://doi.org/10.1016/S0166-3615\(98\)00098-0](https://doi.org/10.1016/S0166-3615(98)00098-0).
- Taiber, J. G. (1996). Optimization of process sequences considering prismatic workpieces. *Advances in Engineering Software*, 25(1), 41–50. [https://doi.org/10.1016/0965-9978\(95\)00084-4](https://doi.org/10.1016/0965-9978(95)00084-4).
- Tao, F., Qi, Q., Liu, A., & Kusiak, A. (2018). Data-driven smart manufacturing. *Journal of Manufacturing Systems*. <https://doi.org/10.1016/j.jmsy.2018.01.006>.
- Thimm, G., Britton, G. A., Whybrew, K., & Fok, S. C. (2001). Optimal process plans for manufacturing and tolerance charting. *Proceedings of the Institution of Mechanical Engineers Part B-Journal of Engineering Manufacture*, 215(8), 1099–1105. <https://doi.org/10.1243/0954405011519024>.
- Tiwari, M. K., Dashora, Y., Kumar, S., & Shankar, R. (2006). Ant colony optimization to select the best process plan in an automated manufacturing environment. *Proceedings of the Institution of Mechanical Engineers Part B-Journal of Engineering Manufacture*, 220(9), 1457–1472. <https://doi.org/10.1243/09544054JEM449>.
- Venkatesan, D., Kannan, K., & Saravanan, R. (2009). A genetic algorithm-based artificial neural network model for the optimization of machining processes. *Neural Computing and Applications*, 18(2), 135–140. <https://doi.org/10.1007/s00521-007-0166-y>.
- Wan, J., Tang, S., Li, D., Wang, S., Liu, C., Abbas, H., et al. (2017). A manufacturing big data solution for active preventive maintenance. *IEEE Transactions on Industrial Informatics*, 13(4), 2039–2047. <https://doi.org/10.1109/TII.2017.2670505>.
- Wang, L. (2009). Web-based decision making for collaborative manufacturing. *International Journal of Computer Integrated Manufacturing*, 22(4), 334–344. <https://doi.org/10.1080/09511920802014912>.
- Wang, L. (2013). Machine availability monitoring and machining process planning towards Cloud manufacturing. *CIRP Journal of Manufacturing Science and Technology*, 6(4), 263–273. <https://doi.org/10.1016/j.cirpj.2013.07.001>.
- Wang, L. (2014). Cyber manufacturing: Research and applications. In *Proceedings of the TMCE* (pp. 39–49). Budapest.

- Wang, L., Feng, H.-Y., & Cai, N. (2003). Architecture design for distributed process planning. *Journal of Manufacturing Systems*, 22(2), 99–115.
- Wang, L., & Shen, W. (2003). DPP: An agent-based approach for distributed process planning. *Journal of Intelligent Manufacturing*, 14, 429–439.
- Wen, X. Y., Li, X. Y., Gao, L., & Sang, H. Y. (2014). Honey bees mating optimization algorithm for process planning problem. *Journal of Intelligent Manufacturing*, 25(3), 459–472. <https://doi.org/10.1007/s10845-012-0696-8>.
- Wong, T. N., Chan, L. C. F., & Lau, H. C. W. (2003). Machining process sequencing with fuzzy expert system and genetic algorithms. *Engineering with Computers*, 19(2–3), 191–202. <https://doi.org/10.1007/s00366-003-0260-4>.
- Xu, L. D., & Duan, L. (2018). Big data for cyber physical systems in industry 4.0: A survey. *Enterprise Information Systems*, 7575, 1–22. <https://doi.org/10.1080/17517575.2018.1442934>.
- 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*, 24(1), 1–31. <https://doi.org/10.1080/0951192x.2010.518632>.
- Yeo, S. H. (1995). A multipass optimization strategy for CNC lathe operations. *International Journal of Production Economics*, 40(2–3), 209–218. [https://doi.org/10.1016/0925-5273\(95\)00052-1](https://doi.org/10.1016/0925-5273(95)00052-1).