Published in final edited form as:
Proc Winter Simul Conf. 2015 December ; 2015: 2100–2111. doi:10.1109/WSC.2015.7408324.

INTEGRATING DATA ANALYTICS AND SIMULATION METHODS TO SUPPORT MANUFACTURING DECISION MAKING

Deogratias Kibira,
Department of Industrial and Systems Engineering, Morgan State University, 1700 E Cold Spring Ln, Baltimore, MD 21251, USA

Qais Hatim,
Department of Industrial and Manufacturing Engineering, Pennsylvania State University, University Park, State College, PA 16801, USA

Soundar Kumara, and
Department of Industrial and Manufacturing Engineering, Pennsylvania State University, University Park, State College, PA 16801, USA

Guodong Shao
Systems Integration Division, Engineering Laboratory, National Institute of Standards and Technology, 100 Bureau Drive, Gaithersburg, MD 20899, USA

Abstract

Modern manufacturing systems are installed with smart devices such as sensors that monitor system performance and collect data to manage uncertainties in their operations. However, multiple parameters and variables affect system performance, making it impossible for a human to make informed decisions without systematic methodologies and tools. Further, the large volume and variety of streaming data collected is beyond simulation analysis alone. Simulation models are run with well-prepared data. Novel approaches, combining different methods, are needed to use this data for making guided decisions. This paper proposes a methodology whereby parameters that most affect system performance are extracted from the data using data analytics methods. These parameters are used to develop scenarios for simulation inputs; system optimizations are performed on simulation data outputs. A case study of a machine shop demonstrates the proposed methodology. This paper also reviews candidate standards for data collection, simulation, and systems interfaces.

1 INTRODUCTION

The manufacturing environment is characterized by continuously changing conditions that affect processes, operations, and priorities. Therefore, evaluating a manufacturing system performance to decide course of action is a challenging task. To monitor performance,

DISCLAIMER
No approval or endorsement of any commercial product by the National Institute of Standards and Technology is intended or implied. Certain commercial software systems are identified in this paper to facilitate understanding. Such identification does not imply that these software systems are necessarily the best available for the purpose.

today's smart manufacturing systems are installed with ubiquitous sensors and other smart systems that are collecting large volumes and varieties of data. The collected data has also issues of veracity, certainty, and validity for intended purpose. Furthermore, the data are interrelated and influenced by many factors. Traditional data analysis methods alone, including simulation, fail to transform this high-volume, continuously streaming data into knowledge for decision support. Data analytics methods are being advanced and applied to understanding how to utilize the high-volume, high-variety data that is being collected from today's manufacturing systems. Data analytics methods, especially data mining, have been targeting important areas in manufacturing such as product quality (Skormin et al. 2002), production planning and scheduling (Chen 2001), and manufacturing process optimization (Gröger et al. 2012; Zheng et al. 2014). Data mining is the process of identifying knowledge hidden in large amounts of data and can be useful to support decision making. Considering the wide range of possible system behaviors that depend on inputs, data mining tools can uncover important parameters that are associated with a given type of behavior. The discovered associations between inputs and behavior can further be analyzed using simulation models to determine the parameter settings that result in the best system performance. As a consequence, better decisions can result when data mining is integrated with simulation models.

Traditionally, decision makers use simulation models to represent a real-world system in a virtual environment, and to test and evaluate the system's performance under different operating conditions. Applying a simulation analysis approach involves collecting data and developing a model using an appropriate simulation software tool (Banks et al. 2009). Evaluations are done based on performance indicators such as capital investments, asset utilization, and environmental impacts (Dudas et al. 2009). The selected indicators largely depend on the performance objectives of the organization and may be different for each simulation study. Because simulation users often need to select system inputs from the large number of possible alternatives, simulation are often combined with optimization methods.

Optimizations apply mathematical techniques for modeling real-world problems and solve problems based on specific objectives to produce actionable recommendations. Brady and Yellig (2005) proposed two approaches for integrating simulation with optimization. The first one is to construct an external optimization framework around the simulation model. The second one is an internal approach to investigate the relationships and interactions among system variables within the simulation model. The tracking features within the tools can be used for the purpose. We use the first approach in this paper.

In summary, we note two issues for using the large volume of collected data to improve the performance of a manufacturing system with simulation. The first one is to determine important parameters affecting the required performance from the data. The second is to determine the best input settings of the parameters to optimize the process. The collected data contains intricate dependencies, which requires automated tools to extract useable information. In this paper we propose a methodology utilizing the strengths of data mining, simulation, and optimization for decision guidance in manufacturing systems. Data mining methods first extract those parameters and variables that affect system performance. We then use the identified parameters and associated data as simulation inputs to predict system performance for defined scenarios. Subsequently, optimization methods are used to determine the best parameter settings, from alternatives generated by the simulation that lead to actionable recommendations. We believe that the synergistic effect of data mining, simulation, and optimization can support manufacturing decision making in the face of big data and system complexity.

The rest of the paper is organized as follows: Section two reviews related work, Section three describes the proposed methodology. Section four shows how the methodology can be used for a machining job shop. Section five concludes the paper and discusses the future work.

2 RELATED WORK AND STANDARDS

This section reviews the existing work and information standards related to the proposed methodology of this paper. Simulation provides an accurate projection of manufacturing system behavior. However, determining the set of inputs that optimize system performance is challenging because simulation optimization necessitates that the decision maker fully understands both the optimization approach and the underlying stochastic processes (Andradóttir 1998). Researchers such as Skoogh et al. (2010) published the GDM-Tool for processing input-streaming data with the purpose of enabling the reuse of simulation models. This tool does not process input data for optimizing defined system performance. Secondly, the large volume of data, the number of possible input parameters, and the variety of their interactions make it difficult to choose the best combination of data inputs relevant for the desired objectives. Data mining uses techniques such as classification, clustering, association, and sequential pattern discovery to discover knowledge hidden in large volumes of data. Recently, researchers have recognized the potential benefit of integrating data mining, simulation, and optimization (Better et al. 2007). Data mining methods, applied to manufacturing data, discover knowledge and patterns in the data and relationships between the data that can be represented in simulation models (Alnoukari et al. 2010).

Previous work in integrating data mining and simulation include software project management (Garcia et al. 2008). In this application, the authors use an association rule mining algorithm to build a model that relates management policy attributes to quality, time, and effort in software development. The applications of data mining in simulation modeling are classified into two modeling types (1) micro-level modeling, which uses data mining techniques on historical data to tune input parameters and (2) macro-level modeling that uses the data mining techniques to analyze data to reveal patterns that could help better model the overall behavior of the system (Remondino et al. 2005). In this paper, we use the latter approach and use the discovered patterns as inputs to simulation and optimization models to obtain input parameter values that provide optimal system performance.

Optimizations are done by formulating problems using operations research methods including metaheuristics and mathematical programming (Olafsson et al. 2008). Carson and Maria (1997) categorized optimization methods into gradient-based search methods, stochastic optimization, response surface methodology, heuristic methods, and statistical methods. For manufacturing, simulation-based optimization methods include response surface, direct search, perturbation analysis, and evolutionary algorithms (Azadivar 1992; Paris et al. 2001). Tools have been developed for analysis of simulation output data (Bogon et al. 2012). This process is classified external optimization, in that it is done outside the simulation model. Simulation tools also incorporate algorithms to provide optimization capability.

Implementing the methodology with multiple methods and tools requires standards. Data and system interface standards are the foundation for information representation, model composition, and system integration. Standards are used to measure, collect, represent, and exchange the data relevant to data analytics, simulation, and production. Currently, different data formats are used in industry. Sample standards for manufacturing systems at different levels follow (Jain and Shao 2014):

- ISA-95 is developed for the integration of enterprise and control systems under coordination efforts by the International Society for Automation (ISA) (ANSI 2010).
- The OAGIS standard, from the Open Applications Group, establishes integration scenarios for a set of applications including enterprise requirements planning (ERP), manufacturing execution system (MES), and Capacity analysis (OAGIS 2014). While OAGIS does not cover full enterprise objects, it is focused on the required models for data exchange.
- Business to Manufacturing Markup Language (B2MML) is a set of eXtensible Markup Language (XML) schemas that implement the data models in the ISA-95 standard. B2MML enables businesses to integrate their Manufacturing Execution System (MES) solutions with their Enterprise Resource Planning (ERP) systems.
- Core Manufacturing Simulation Data (CMSD) is a standard to help achieve simulation applications interoperability (SISO 2012). CMSD enables exchanging shop floor simulation data with manufacturing applications such as ERP, Master Production Schedule, and MES.
- MTConnect is a middleware standard that enables the real time, automated data extraction from numerically-controlled machine tools using the XML standard (AMT 2013).
- Emerging Data analytics standard: PMML is a data mining standard developed by the Data Mining Group (DMG), an independent, vendor-led consortium. PMML describes the exchange of statistical and data mining models. With PMML, it is easy to develop a model on one system using one application and deploy the model on another system using another application (DMG 2014).

3 PROPOSED METHODOLOGY

This section describes the methodology illustrated in Figure 1. The first step is formulating the problem and specifying high-level performance objectives, indicators, and metrics. This is followed by acquiring domain knowledge of the manufacturing system, processes, performance indicators, and metrics. Next, a conceptual model needs to be developed for understanding the requirements for modeling, simulation, and analysis. Then, data analytics methods need to be applied to the data collected to extract parameters and developing scenarios for inputs to the simulation model. Actionable recommendations are obtained through simulation optimizations. Each step of the methodology is described next.

Two features distinguish this methodology from traditional approaches (1) input of a large volume and variety of constantly streaming data collected from the system using smart devices, and (2) using association and classification methods of data mining to determine important parameters associated with given performance indicators. The indicators can differ with every industry or occasion. As indicated in the introductory section, traditional simulation approaches would fail to be applied to this type data.

3.1 Formulate the Problem

Formulate the problem by receiving problem input data from the real world, identify the system or processes of interest and specify performance goals by defining indicators and metrics at a high level. Identify relevant resources, products, and activities. System conditions, constraints, and decision variables should also be defined.

3.2 Acquire Domain Knowledge

Acquire or obtain from domain experts, knowledge related to the problem including performance indicators, metrics, conditions, and targeted goals. If the goal is agility performance, for example, the user would research on the relationship between agility and collectable data. The user would also study factors that define and determine agility performance.

3.3 Design a Conceptual Model

Develop a conceptual model, which is a simplified representation of the identified problem. It provides the right level of abstraction that satisfies the modeling objectives and focuses on the metrics of concerns. It helps modelers better understand the problem and prepare for modeling and analysis. When designing a conceptual model, the following typical questions need to be answered to help users abstract the problem and plan the detailed modeling (1) What are the components (systems/processes) that need to be modeled?, (2) What are the inputs and outputs of each component?, (3) What are the relationships between components?, (4) What are the metrics and indicators?, and (5) What are the data requirements for the metrics? The conceptual models help identify requirements for data collection.

3.4 Collect Data

Collect raw data using various devices and methods such as sensors, bar codes, vision systems, meters, and radio frequency identification (RFID). Gröger et al. (2012) classified data into manufacturing process data and operational data. Process data is made up of execution data; i.e., machine and production events recorded by the MES. Process data from machine tools include processing time, idle time, loading time, energy consumption, machine setting, tool, and tear down time. MTConnect is one standard that can be used for this purpose. Operational data mainly encompasses Computer-aided design (CAD), Computer-aided Process Planning (CAPP), and ERP data. For data storage, Structured Query Language (ISO/IEC 2011) is one means of storing and retrieving data. The data is represented in neutral format such as XML.

3.5 Use Data Analytics Methods

Select appropriate data analytics methods that should (1) use the collected data to identify parameters that are related to defined performance, (2) be adaptable to different data and performance objectives, and (3) perform the data analysis.

Data mining methods are used because the complexity of the shop floor data makes it difficult to establish analytical relationships between the input variables and performance measures. Choosing the appropriate data mining method depends on the particular problem. For example, association methods should be used to determine whether there is a relationship between two data sets. Classification methods should be used to identify specific characteristics or attributes of a data set and to determine whether a new data item belongs to a group that exhibits these attributes (Better et al. 2007). Our approach is to first define performance indicators and use the association method to determine, from the collected data, the particular parameters that impact the performance indicator. Each performance objective or sets of objectives form distinct groups. These objectives are defined before the data mining process and the corresponding groups are known a priori. The determination of the relevant data type acts as a data preparation for input to the simulation model.

If y is the performance indicator, we can represent y as a function $y = f(x, w)$, where $x = (x_1, x_2, x_3, \dots x_d)^T$ denotes the set of parameters that impact energy use and w denotes the weight of the parameters.

3.6 Perform Simulation Modeling and Optimization

Construct the simulation and optimization models, incorporating sufficient detail to evaluate performance. There are a number of commercial simulation tools available on the market. In performing optimization, we need to define the decision variables, x and optimization criteria. Also, define constraints and restrictions on values of decision variables.

In example of optimizing energy consumption:

If

$F(x)$ = function that expresses the total energy consumption

$A(x)$ = matrix of production needs for products

b = minimum requirements for each product

L_{min} = lower limit

L_{max} = upper limit

The formulation would be as follows:

$$
\begin{aligned}
&\text{Minimize} \quad F(x)\\
&\text{Subject to} \quad A(x) \geq b\,(\text{constraints})\\
&\qquad\qquad \mathrm{L}_{\min} \leq x \leq \mathrm{L}_{\max}
\end{aligned}
$$

The optimization model can also use any optimization tools supplied with simulation software. Simulation quantifies the impact of the inputs used to run the system. By making several runs of different inputs and what-if scenarios, the tools systematically compare the results of each current run with those of past runs to decide on a new set of input values until the optimum is gradually approached. The CMSD standard can be used to model the input data for the simulation modeling.

3.7 Derive Actionable Recommendations

Interpret and translate the output from the optimizations into actionable recommendations that can be executed on the manufacturing system. The users also need to check if the recommended actions conflict with already perceived knowledge about the system and resolve this conflict.

4 CASE ANALYSIS FOR IMPLEMENTING THE METHODOLOGY

This section describes how the methodology was demonstrated using a machining job shop. It is a simplified setting to showcase the steps of the methodology and does not include master data from the ERP system. This section (1) describes the production process, (2) defines performance objectives and, (3) describes how the proposed methodology was applied to achieve the performance objectives.

The job shop produces a variety of custom-designed metal products. The shop floor consists of a number of machine tools including a turning lathe, a mill, a drill press, and a boring machine. When an order is received, the users can decide to focus on any or all of these performance objectives (1) minimize costs (e.g., labor, cutting tool, and energy costs), (2) minimize resource usage (e.g., material, energy, and water), and (3) maximize productivity. Each part has a process plan. However, the sequencing of orders or of parts at a machine or a station can vary depending on the users' objectives. Some machines can perform more than one process. The choice of a machine for a process will produce different impacts on resource (materials and energy) consumption and processing time.

The machines can have different setup parameter settings such as feed rate, cutting speed, and depth of cut. These also affect cycle time, production rate, cost, and resource consumption. Figure 2 shows the production flow through the shop. Data are collected on resources, products, environment, and decision rules. Because of multiple objectives and large volume of data collected, it is impossible to determine the optimal combination of sequence, machines and settings, or batch size without a tool or a systematic methodology to identify and optimize these parameters according to the required performance objective.

Formulate the problem

The problem is formulated as follows.

Objectives: optimize materials and energy consumption and productivity

Decision: obtain optimal process plan (including machines and machine settings) for manufacturing parts

Conditions/situation: consider that multiple machines can be chosen to perform an operation, multiple settings for a machine; and variable impacts can occur depending on the selected machines and settings.

Acquire domain knowledge

The following knowledge was needed before modeling: machining processes, energy consumption in machining, production scheduling in job shops, sequencing, costing of manufacturing processes, performance indicators and metrics, and performance data.

Design conceptual model

Based on the knowledge of the defined problem, a high-level conceptual model is developed to highlight the relationship between inputs and outputs. The information needs are: product design, process routes, product material, mapping product design and material to a process, machines and tools, machine setting, and a performance indicator that drives the selections above.

Collect data

Data is collected from the machines as production orders flow through the shop. The attributes of the production order are:

- product type (sub-attributes: design features, material),
- manufacturing equipment (sub-attributes: machine type for an operation, machine settings, tool, machine energy use, machine process time),
- production planning (sub-attributes: batch size, sequencing rule, part routing), and
- performance data (sub-attributes: energy consumption, production cost, production time).

Use data analytics' methods

We use association rules' techniques from data mining to discover the parameters (attributes) that have significant impact on the defined performance. For this demonstration we discover that for a given material, the parameters that affect energy consumption are (1) the machine, (2) diameter of cutter, (3) number of teeth on cutter, (4) depth of cut and, (5) feed rate.

Perform simulation modeling and optimization

We construct a discrete event simulation model of the machine shop using a simulation software tool to predict performance. For energy consumption we evaluate how a given machine and cutting tool affect the energy use without caring about other indicators. The main simulation modules are part arrival, data requirements for the part and process, the part routing to various machines, part exit, and statistics generation. Instead of a separate optimization tool, actionable recommendations are obtained by using optimization capability provided by using OptQuest that is optimization package integrated with Arena. OptQuest uses heuristics known as Tabu search, integer programming, neural networks, and scatter search for seeking within the control (input) space and converges to an optimal solution. The user controls the possible ranges of input variables and defines the objective and sets-up inputs for OptQuest. The CMSD standard can be used to model the input data for the simulation modeling. Table 1 shows the scenarios used in this simplified case. The table also displays the resulting impacts from various system inputs.

Derive actionable recommendations

We execute the simulation model for processing a part product that requires the processes: facing, grooving, threading, spot drilling, and final drilling. Each process is associated with a resource set (R); i.e., machine (designated M) and a tool (T). Three cases have been considered: predefined process plan for the features' production sequence, relaxation on the operational order for some features, and unspecified process plan. In the predefined case, each process has a predetermined machine and cutting tool, determined to optimize a given performance objective. In case of minimum-energy-utilization objective, the machines selected are those that perform the process with minimum energy consumption. In the unspecified case, a machine is selected according to a priority rule such as machine with minimum number of parts waiting.

For each of the three cases, different process plans are tested and for each combination (production and process planning) impacts on two key performance indicators (KPIs): energy consumption and production time. Table 1 shows the energy consumption and production time data for different scenarios of process plans. The resource column shows options of machine and tools for a process; while the indicator columns show the resulting impacts. The table shows the tool-tip energy while the production time displays only the total processing time on the machines. Table 1 shows that the choice of sequence plan, operation, and resource influences the performance indicator. By resource we refer to the machine tool and cutting tool used. The results are summarized in Table 2 where the optimum inputs and settings can be selected visually. The minimum energy consumption is obtained by selecting resources $R_2R_3R_4R_6R_9$.

5 DISCUSSION AND FUTURE WORK

This paper has introduced a methodology that integrates data analytics, simulation, and, optimization to analyze large volumes of data for the purpose of improving decision making. Data mining extracts information - such as patterns and statistical distributions – that provides inputs to a simulation model. We use this model to develop different manufacturing scenarios and to compute various performance metrics. We then use optimization techniques to search for best input selections for those metrics. We demonstrated how to use the methodology using a case study for identifying a process plan that optimizes production cost.

Implementing this methodology requires standards that are relevant for the following purposes (1) data collection, (2) data representation, (3) model composition, and (4) system integration. Candidate standards include MTConnect, PMML, CMSD, and ISA-95. OAGIS (OAGIS 2014) can integrate applications including ERP, MES, and capacity analysis but it is more emphasized at the enterprise level. ISA-95 is more emphasized at the operations level. Further, OAGIS and ISA-95 standards were not intended to provide interfaces with simulation systems nor with each other. Future work is needed for these two standards to support simulation integrations both at shop floor level and between different planning levels in a manufacturing company. On the other hand, CMSD is developed especially for integrating simulation systems applications with other manufacturing applications. It is a candidate standard for interoperability with simulation models. More standardization efforts are needed especially for data collection, where data collected is still limited to machine tool data, representation and data mining.

For further development of this methodology, future work includes the definition and description of a framework for data collection and interface for input to data mining and simulation tools; investigation of data mining standards for the methodology; the requirements analysis for extension of existing standards for interfacing between data mining tools, simulations, optimization, and manufacturing system monitoring tools; and conducting industrial case studies to further validate the proposed methodology.

Acknowledgments

This effort has been sponsored in part under the cooperative agreement No.70NANB12H284 between NIST and Pennsylvania State University and under cooperative agreement No. 70NANB13H153 between NIST and Morgan State University. The work described was funded by the United States Government and is not subject to copyright.

References

Alnoukari M, El Sheikh A, Alzoabi Z. An Integrated Data Mining and Simulation Solution. Handbook of Research on Discrete-event Simulation Environments Technologies and Application. 2010; 16:359–380.

AMT. [Accessed February 2, 2015] Getting Started with MTConnect: Monitoring Your Shop Floor – What’s In It For You?. AMT - The Association for Manufacturing Technology. 2013. http://www.mtconnect.org/media/39437/gettingstartedwithmtconnectshopfloormonitoringwhatsinitforyourevapril4th-2013.pdf

Andradóttir, S. Handbook of Simulation: Principles, Methodology, Advances, Applications, and Practice (Chapter 9). New York: John Wiley & Sons; 1998.

ANSI. ANSI/ISA-95.00.01: Enterprise-Control System Integration - Part 1: Models and Terminology. American National Standards Institute; 2010.

Azadivar, F. A Tutorial on Simulation Optimization. In: Swain, JJ.Goldsmith, D.Crain, RC., Wilson, JR., editors. Proceedings of the 1992 Winter Simulation Conference; New Jersey: Institute of Electrical and Electronics Engineers, Inc; 1992. p. 198-204.

Banks, J., Carson, JS., II, Nelson, BL., Nicol, DM. Discrete-Event Simulation System. 5. Prentice-Hall International Series in Industrial and Systems Engineering; 2009.

Better M, Glover F, Laguna M. Advances in Analytics: Integrating Dynamic Data Mining with Simulation Optimization. IBM Journal of Research and Development. 2007; 51:477–488.

Bogon, T., Timm, IJ., Lattner, AD., Paraskevopoulos, D., Jessen, U., Schmitz, M., Wenzel, S., Spieckermann, S. Towards Assisted Input and Output Data Analysis in Manufacturing Simulation: The EDASim Approach. In: Laroque, C.Himmelspach, J.Pasupathy, R.Rose, O., Uhrmacher, AM., editors. Proceedings of the 2012 Winter Simulation Conference. New Jersey: Institute of Electrical and Electronics Engineers, Inc; 2012. p. 257-269.

Brady, T., Bowden, R. The Effectiveness of Generic Optimization Routines in Computer Simulation Languages. Proceedings of the Industrial Engineering Research Conference; Dallas, Texas. 2001.

Brady, T., Yellig, E. Simulation data mining: A New Form of Computer Simulation Output. In: Kuhl, ME.Steiger, NM.Armstrong, FB., Joines, JA., editors. Proceedings of the 2005 Winter Simulation Conference. New Jersey: Institute of Electrical and Electronics Engineers, Inc; 2005. p. 285-289.

Carson, Y., Maria, A. Simulation Optimization: Methods and Applications. In: Andradottir, S.Healy, KJ.Withers, DH., Nelson, BL., editors. Proceedings of the 1997 Winter Simulation Conference; New Jersey: Institute of Electrical and Electronics Engineers, Inc; 1997. p. 118-126.

Chen IJ. Planning for ERP Systems: Analysis and Future Trend. Business Process Management Journal. 2001; 7:374–386.

Data Mining Group (DMG). [Accessed March 2, 2015] PMML v4.2.1. 2014. http://www.dmg.org/

Dudas, C., Hg, AHC., Bostron, H. Information Extraction From Solution Set Of Simulation-Based Multi-Objective Optimization Using Data Mining. 7th International Industrial Simulation Conference; Loughborough, UK. 2009. p. 65-69.

Garcia M, Roman I, Penalvo F, Bonilla M. An Association Rule Mining Method for Estimating the Impact of Project Management Policies on Software Quality, Development Time and Effort. Expert Systems with Applications. 2008; 34:522–529.

Gröger, C., Niedermann, F., Mitschang, B. Data Mining-Driven Manufacturing Process Optimization. Proceedings of the World Congress on Engineering; 2012; London, U.K. 2012. p. 4-6.

[Accessed July 16, 2015] ISO/IEC SQL Part 1: SQL Framework. http://www.jtc1sc32.org/doc/N2151-2200/32N2153T-text_for_ballot-FDIS_9075-1.pdf

Jain, S., Shao, G. Virtual Factory Revisited for Manufacturing Data Analytics. In: Tolk, A.Diallo, SD.Ryzhov, IO.Yilmaz, L.Buckley, S., Miller, JA., editors. Proceedings of the 2014 Winter Simulation Conference; New Jersey: Inst. of Electrical and Electronics Engineers, Inc; 2014. p. 887-898.

Li X, Olafsson S. Discovering Dispatching Rules Using Data Mining. Journal of Scheduling. 2005; 8:515–527.

MTConnect Institute. [Accessed April 3, 2015] http://www.mtconnect.org

OAGi. [Accessed November 23, 2014] Open Application Group’s Integration Specification (OAGIS), Edition 10.0. http://www.oagi.org/dnn2/DownloadsandResources/OAGIS100PublicDownload.aspx

Olafsson S, Li X, Wu S. Operations Research And Data Mining. European Journal of Operational Research. 2008; 187:1429–1448.

Paris J, Pierreval H. Dealing with Design Options in the Optimization of Manufacturing Systems: An Evolutionary Approach. International Journal of Production Research. 2001; 39:1081–1094.

Reinhart, G., Gloneggera, M., Festnera, M., Egbersa, J., Schilpa, J. Adaption of Processing Times to Individual Work Capacities in Synchronized Assembly Lines Technologies and Systems for Assembly, Quality, Productivity and Customization. In: Hu, J., editor. Proceedings of the 4th CIRP Conference on Assembly technologies and Systems; Ann Arbor, Michigan. 2012.

Remondino, M., Correndo, G. Data Mining Applied to Agent Based Simulation. Proceedings of the 19th European Conference on Modeling and Simulation; Riga, Latvia. 2005.

Shao, G., Jain, S., Shin, S. Data Analytics Using Simulation for Smart Manufacturing. In: Tolk, A.Diallo, SD.Ryzhov, IO.Yilmaz, L.Buckley, S., Miller, JA., editors. Proceedings of the 2014 Winter Simulation Conference; New Jersey: Institute of Electrical and Electronics Engineers, Inc; 2014. p. 2192-2203.

SISO. SISO-STD-008-01-2012: Standard for Core Manufacturing Simulation Data – XML Representation. Simulation Interoperability Standards Organization; Orlando, F L: 2012.

Skoogh, A., Michaloski, J., Bengtsson, N. Towards Continuously Updated Simulation Models: Combining Automated Raw Data Collection and Automated Data Processing. In: Johansson, B.Jain, S.Montoya-Torres, J., Yucesan, E., editors. Proceedings of the 2010 Winter Simulation Conference; New Jersey: Institute of Electrical and Electronics Engineers, Inc; 2010. p. 1678-1689.

Skormin VA V, Gorodeski I, Popyack LJ. Data Mining Technology for Failure Prognostic of Avionics. IEEE Transactions - Aerospace and Electronic Systems. 2002; 38:388–403.

Zheng, L., Zeng, C., Li, L., Jiang, Y., Xue, W., Li, J., Shen, C., Zhou, W., Li, H., Tang, L., Li, T., Duan, B., Lei, M., Wang, P. Applying Data Mining Techniques to Address Critical Process Optimization Needs in Advanced Manufacturing. Proceedings of the 20th ACM SIGKDD International Conference on Knowledge Discovery and Data Mining; 2014. p. 1739-1748.

Biographies

DEOGRATIAS KIBIRA is with Morgan State University. His current research is in developing performance assurance methodologies for smart manufacturing systems. He has PhD in Manufacturing Engineering from the University of New South Wales. His e-mail address is deogratias.kibira@.morgan.edu.

QAIS HATIM holds the Dual Degree Doctoral of Industrial Engineering and Operation Research in the Department of Industrial Engineering at The Pennsylvania State University, USA, Class of 2015. He was also a guest researcher in the Life Cycle Engineering Group in NIST's Systems Integration Division of the Engineering Laboratory. His current research is combining statistics, optimization, and data analytics with manufacturing in order to build robust models that are feasible for implementation in real life situations. His email is qaiszakry@gmail.com.

GUODONG SHAO is a computer scientist in the Life Cycle Engineering Group in NIST's Systems Integration Division of the Engineering Laboratory. His current research topics include modeling, simulation, and analysis; data analytics; and optimization for Smart Manufacturing. He served as a member of the WSC Board of Directors and on the editorial board of the International Journal on Advances in Systems and Measurements. He holds a PhD in IT from George Mason University. His email address is gshao@nist.gov.

SOUNDAR KUMARA is a Professor in Industrial and Manufacturing Engineering Department at Penn State University where he also holds a joint appointment in the School of Information Sciences and technology and the Department of Computer Science and Engineering. His email is skumara@psu.edu.

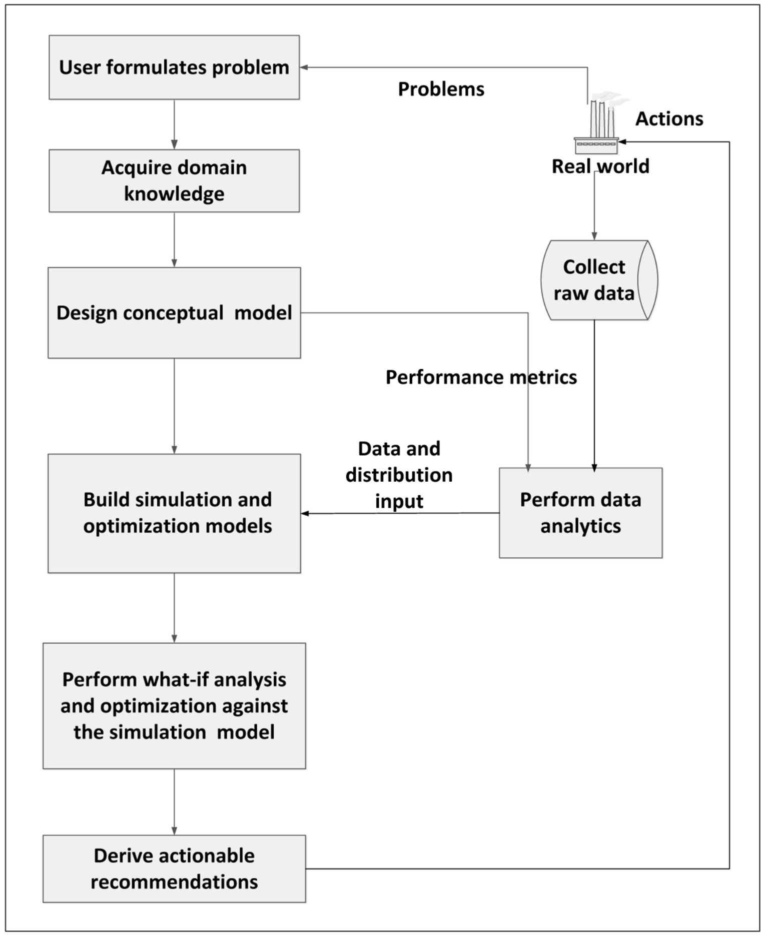

Figure 1.

Procedure for data analytics and simulation optimizations.

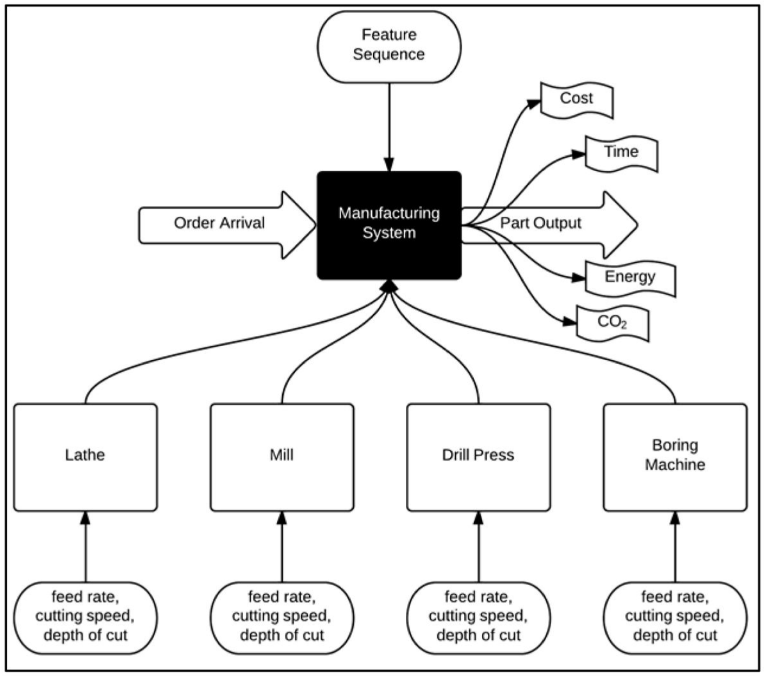

Figure 2.
Production flow through a machining shop.

Table 1

Impacts of selected resources on performance indicators.

Feature Sequence Plan	Operation	Resource R_i	Sustainability Indicator	Productivity Indicator
			Machining Energy (kWh)	**Production time (h)**
Predefined Feature Sequence Plan	Facing	R_1= M1-T1	19.901	0.215
		R_2= M2-T5	16.205	0.014
	Grooving	R_3= M2-T4	16.205	0.014
	Threading	R_4= M1-T2	5.970	0.064
	Spot Drill	R_6= M1-T3	5.307	0.057
		R_7= M3-T7	6.336	0.292
	Drill	R_6= M1-T3	13.267	0.143
		R_9= M4-T9	8.817	0.183
Partially Defined Feature Sequence Plan	Facing	R_2= M2-T5	16.205	0.014
	Grooving	R_3= M2-T4	16.205	0.014
	Threading	R_4= M1-T2	7.793	0.060
	Spot Drill	R_6= M1-T3	6.927	0.053
	Drill	R_6= M1-T3	17.318	0.132
Undefined Feature Sequence Plan	Facing	R_2= M2-T5	16.205	0.014
	Grooving	R_3= M2-T4	16.205	0.014
	Threading	R_4= M1-T2	7.793	0.060
	Spot Drill	R_6= M1-T3	6.927	0.053
	Drill	R_6= M1-T3	17.318	0.132

NIST Author Manuscript
NIST Author Manuscript

Table 2

Summary of process plans for different feature sequences when minimizing energy consumption.

Feature Sequence Plan	Process Plan PP_j	Facing	Grooving	Threading	Spot Drill	Drill
Predefined Feature Sequence Plan	PP_1	R_1	R_3	R_4	R_6	R_6
	PP_2	R_1	R_3	R_4	R_7	R_6
	PP_3	R_1	R_3	R_4	R_6	R_9
	PP_4	R_1	R_3	R_4	R_7	R_9
	PP_5	R_2	R_3	R_4	R_6	R_6
	PP_6	R_2	R_3	R_4	R_7	R_6
	PP_7	R_2	R_3	R_4	R_6	R_9
	PP_8	R_2	R_3	R_4	R_7	R_9
Partially-Defined Feature Sequence Plan	PP_1	R_2	R_3	R_4	R_6	R_6
Undefined Feature Sequence Plan	PP_1	R_2	R_3	R_4	R_6	R_6